Open Data and Big Data Programs in Local Government

Policy Analysis

MPA Research Report

Submitted to

The Local Government Program
Department of Political Science
The University of Western Ontario

Susan Chase
July 2014

Abstract

The Open Government movement provides the opportunity for government to be more transparent and accountable to the public. The Open Data program is a component of the broader Open Government initiative intended to improve accessibility of information and to provide for greater participation and engagement of the community. There are many benefits to be realized with the Open Data program but there are challenges to be met before the program can be implemented. This paper will identify the issues around policy development and determine if policies are a response to the new technology trend, to changing social values, or as a response to traditional views and regulations on government information and service delivery. The paradox between the need for privacy and the protection of personal information and the support for the principle of openness and transparency creates a unique challenge for policy makers to overcome. The social impacts and inputs create a contradictory storyline that adds to the complexity of policy development.

This paper examines the issues in developing a policy framework that addresses the contradictory nature of this program. Through a literature review, a survey of large municipalities in Ontario and in-depth interviews, the research attempts to identify if there is a formal process that can deal with the complexities and evolving nature of this program or if government is avoiding the issue and simply delivering a program in their goal to achieve transparency. The ethical questions that arise during the analysis of open data programs pose interesting challenges for policy makers in local government. The analysis investigates if there are particular factors responsible for the program to be placed on the formal policy agenda. It looks to identify the factors that are associated with policy development for a evolving program or if there is in fact a lack of policy development as a result.

In the end, it appears there is a definite lack of policy development and that this is likely due to the quickly changing technology and opportunities that technology brings but also due to the emerging social issues being presented by the different generational beliefs and values. Recommendations are provided to help mitigate the risk and yet there are consistently new concerns expressed that requires a re-examination of the policy. The report proves to be a cautionary tale concerned with local government moving too quickly and without regard to the risks because of trying to meet the incessant demand for openness and accountability.

Table of Contents

Introduction

Governments are being held to higher standards and scrutiny than ever before. Open Government programs are a means to provide greater access to information, to improve transparency of government operations and decision making, to encourage greater public participation in government and to keep government accountable to the public and to regulators. A significant driver of this demand for openness is the infiltration of technology in both business and personal lives. We live in the information age and the knowledge economy. People now have immediate access to the information they want such as their personal correspondence, transit schedules, and stock market updates, from wherever it might be stored. People expect that government information should be just as accessible online as it is through any other means.[1] Local government collects a great deal of data about municipal operations, services and residents. Reports and documents are available online but this information is typically read-only and is produced with the message government determines. This research paper will attempt to identify the issues around the development of policy for the open data and big data programs in municipal governments. It will also investigate if Open Data policies are a response to a new technology trend, to changing social values, or a response to traditional views and regulations on government information and service delivery.

As part of the new public management values, government actors want to collaborate more with the citizen and engage others in the decision making process to improve performance standards and increase customer focus.[2] Citizens want information about how their government operates and to hold government accountable to make the best decisions possible. Several municipalities have developed open data programs to share the information online in order to make it available to whoever would like to use it. A principle of open data is that the data is

---

[1] The Institute for Citizen-Centred Services http://www.iccs-isac.org/research/citizens-first/citizens-first-6/?lang=en

[2] Tindal, C.R., Nobes Tindal, S. *Local Government in Canada*, Seventh Edition (Canada: Nelson Education, 2009), 288

digital, open format, and flexible so that it can be freely shared and used in different computer-based applications. This emerging trend of openness is seen by many as being of prime importance to truly open government and yet it seems almost contradictory in relation to the Privacy Commissioners recommendations to maintain privacy and security of data. The question is how to make information available to the public yet still maintain the privacy by design principles and eliminate the risk of releasing data inappropriately. Are municipalities analyzing the data they have in order to identify trends or linkages? What purpose is appropriate and who should determine how the data is to be used? Should government be responsible for simply publishing the data and not for monitoring its use or outputs? The development of a clear policy on the collection and storage of the data, how the data can be used and how the data is to be published and updated is required. Yet the open data environment is complex and evolving daily. Technology is changing so rapidly that what is viable data one day is not the next. There are many benefits to this program but there are also several concerns that should be considered before governments move towards implementation. A thorough evaluation of all the potential outcomes and risks of the program needs to be investigated and then incorporated into a policy framework. This will help ensure the promise of open government while also protecting the integrity and privacy of the information collected.

This research paper illustrates surveyed municipalities that have open data programs to identify if they also have policies that mitigate these risks, and reviews the challenges in developing these policies to publish data as quickly as possible. The research identifies what municipalities are doing regarding these programs, the implications of the programs, and then identifies recommendations for policies and guidelines for the public sector. Municipalities are cautious and risk adverse in nature, yet open data programs expose the municipality in many positive and negative ways. An example is evident in the recent proceedings in the United States regarding the National Security Agency (NSA) and 'WikiLeaks' and their attempt to respond to the ever-changing reaction from their citizens. People reacted strongly to the CBC

(Canadian Broadcasting Company) reports that the Canadian federal government is accessing data and information about persons using telecommunications technologies directly through the carriers. The carriers are refusing to disclose the type of information the government requested.[3] Some say this information and data should be widely available and yet others say it is private and no one including government should be able to access this level of personal data. Private sector businesses use big data programs extensively and primarily for marketing or inventory management. Government is slower to implement big data programs but it is unclear as to whether this is due to resource constraints or whether it is due to the interpretations of the Municipal Freedom of Information Protection of Privacy Act (MFIPPA) and protecting the privacy of the data. There is a need for further study on the implications of acquiring and analyzing the vast amounts of data collected by government actors. It is important to ensure frameworks are in place to not only protect an individuals' privacy but also their right to consent or not consent to the use of their personal data.

Background

Local municipal governments are facing more demands for new or improved programs and services than there are available resources to deploy. Municipal administrators look for quantifiable tools that help determine what government services ought to move forward and which programs should be cancelled, during various decision making processes. Technology influences and improves government operations and service delivery. To list a few: more efficient use of resources such as automated payroll systems; safety measure improvements through traffic signal systems; environmental conservation solutions through smart meters; and improved communications and customer engagement through social media. Don Tapscott of the Martin Prosperity Institute recently said,

---

[3] CBC news reports Retrieved from http://www.cbc.ca/news/politics/telecoms-refuse-to-release-information-on-private-data-given-to-feds-1.2626286

"we need to open up by releasing data to the world. [We] could engender self-organizing networks involving the private sector, NGO's, academics, foundations, individuals and other government agencies to create public value [in doing so]. This has nothing to do with outsourcing, but rather [its change] in the division of labour in society about how we create services and public value overall." [4]

Implementing an Open Data program is typically included as a component of an 'Open Government' initiative, often captured in a municipal strategic plan. The City of Ottawa's Accountability and Transparency Policy itemizes Open Government as a key initiative. The municipal website (Ottawa.ca) states that "as part of [their] commitment to improving citizen engagement and enhancing transparency and accountability to its residents, the City provides public access to its data. Open Data is data that is made freely available to the public in machine readable format without restrictions from copyright, patents or other mechanisms of control." [5] Many different levels of government are implementing Open Data programs as people expect greater transparency from their public administrators.

Both private and public sector organizations are moving forward in this digital age to find ways to enhance service levels, to better inform and engage citizens and compete on a global scale. The face of municipal government is changing and will continue to do so because of technology. It is important to understand that with use of technology, networks and programs such as open data initiatives, there is more risk. Government administrators are no longer the only ones with the knowledge of service delivery performance once the data is published freely. They can now be held accountable by the public far beyond those who complain that their garbage is not picked up. Through the use of technology and data analysis, people can constructively criticize that their municipality is not planning efficient routes for garbage trucks and thus the cost to taxpayers is more due to these inefficient operations. Open data programs are an instrument for local government to appear transparent to the public. Government

---

[4] Tapscott, D. "Rob Ford? John Tory? Olivia Chow? Here's what matters." *The Toronto Star*: Big Ideas in partnership with Martin Prosperity Institute. March 14, 2014
[5] City of Ottawa "Accountability and Transparency Policy," November 28, 2007 Retrieved from http://ottawa.ca/en/city-hall/your-city-government/policies-and-administrative-structure/accountability-and-transparency (accessed June 2014)

administrators are providing high levels of disclosure to the public by allowing them to view and analyze data the government collects regarding service delivery. People are able to view the data, manipulate it into meaningful outputs as a means to conduct their own research, develop apps that are more user-friendly or applicable to the resident, or satisfy their own questions and curiosity. Some municipalities hold competitions that encourage the public to create apps later published on the municipal website for all to use. This model empowers citizens with knowledge and more detailed information about government services than ever before. There are many groups that vary in size, who are proponents of open data programs. Organizations such as 'civic access,'[6] the 'Sunlight Foundation,'[7] and the 'Open Government Partnership'[8] strongly urge the publication of government data. There are equally as many who are concerned with this movement to publish government-sourced data as well. Privacy commissioners, security organizations, and government administrators are more reluctant to publish data due to the risks of private data being released, the linking of datasets to produce incorrect or inappropriate conclusions or the commercialization of publicly funded assets.

Big Data programs are internal programs whereby technology professionals present the data collected in various databases and correlate it in order to provide decision-making analysis. There are varying definitions of big data available. The Gartner Group defines it as "high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making."[9] A further explanation refers to "the data itself and the efforts to leverage it for knowledge, information and insight."[10] Private sector businesses regularly conduct business this way, marketing or focusing service offerings to the personal needs of the customer. Loyalty cards

---

[6] Civic Access, http://civicaccess.ca/ (accessed June 2014)
[7] The Sunlight Foundation, http://sunlightfoundation.com/ (accessed June 2014)
[8] The Open Government Partnership, http://www.opengovpartnership.org/ (accessed July 2014)
[9] Gartner Research, *IT Glossary* retrieved from: http://www.gartner.com/it-glossary/big-data/ (accessed May 2014)
[10] Graham, K and Daley, M. *Big Data and Local Government: Turning potential into Practice* Municipal Monitor Spring 2014, 15

are one mechanism that businesses use in order to aggregate and use this data. A customer signs up for a loyalty card thereby giving permission for the information to be gathered. The business now can track all purchases and inquiries made by the customer to better understand the customer's buying trends. Big data analysis can lead to changing the way government does business in the public sector. It can identify trends about how the community is changing and adapt services to meet the new requirements. Customer service could be focused on the individual customer needs or the segmentation of the customers. This would allow for very large amounts of data to be analyzed based on the different segments of the customer. This provides valuable "insight about the preferences and behaviours of each of the customer segments."[11] This data analysis provides opportunities for government to further improve service delivery and yet ethical considerations must be given to the security, privacy and transparency of these programs in order to avoid the "Big Brother" threat or the police-state of too much government monitoring. Government is slower to implement big data programs but it is unclear as to whether this is also due to resource constraints or whether it is due to the interpretations of MFFIPPA and protecting the privacy of the data.

The risk of open data and/or big data extraction and analysis becoming the norm is an issue, which will result in changing the intent to protect privacy. The Privacy Commissioner of Ontario stated clearly the intent of the MFIPPA and yet, some government agencies still move forward with combining datasets in order to analyze the information in a consolidated way. The program is a mechanism government organizations can use to provide a free and publically accessible product, allowing end-user's the ability to form their own opinions based on raw data as opposed to information that may have political or administrative bias. The program is based on the presumption that government data is a public asset, acquired with public funds that should be available for people to access, analyze and reuse. The Information and Privacy Commissioner for British Columbia recently reviewed their Open Government project and stated

---

[11] *Ibid.*, 16

"Open government is of vital importance to our democracy. Citizens need information about government's actions and decisions to hold governments to account, engage in informed debate, and participate in democratic processes."[12] Many different levels of government are implementing Open Data programs as people expect greater transparency from their public administrators. The United States federal government issued an executive order regarding an open data policy[13] and in 2012, the Canadian federal government announced an action plan to increase access to Open Data.[14] Different provinces and a number of Canadian municipalities are also implementing Open Data programs. The Canadian Federal Government, "[defines] this stream of activity as making raw data available in machine-readable formats to citizens, governments, not-for-profit and private sector organizations to leverage it in innovative and value-added ways."[15] The program is based on the presumption that government data is a public asset that should be available for people to access, analyze and reuse.

The Open Data program is intended to improve accountability by making municipal data readily available to those outside the organization that may want to use it for different purposes. The data is provided in a raw machine-readable format so that end-users can assume no manipulation occurred. This also provides end-users the ability to transform available data to the most appropriate format for their needs. Open Data enhances transparency by improving accessibility for people that wish to identify issues or trends by analyzing available data. Access to this data allows end-users' to form their own opinions based on raw data, opposed to published information that may have a political or administrative bias. Governments will often publish transit schedules yet users expect more information, such as if the buses are actually running on time. Open Data programs provide access to the data at no charge to anyone

---

[12] Denham, E. *Evaluating the Government of British Columbia's Open Government Initiative,* (Victoria, BC, CAN Office of the Information and Privacy Commissioner for British Columbia, 2013.) 4
[13] The Government of the United States of America, The White House, Open Government Initiative. http://www.whitehouse.gov/open (accessed March 2014)
[14] The Government of Canada - Open Data Plan http://data.gc.ca/eng (accessed June 2014)
[15] *ibid.*

interested in analyzing or in creating and sharing applications. Stakeholders will use the data in different ways and for different purposes but the key consideration for administrators is making the data available in an easy and accessible format.

Open data programs are not just about technology, it is more about a new way of government, a new way of sharing information and offering services. Many people believe that open data programs are just for technology 'geeks' or researchers to use and access to information but open data is more about information sharing and proving that government is open and transparent. It can provide for more efficient service delivery and interaction between the government and the citizen. This leads to a two-fold benefit as government spends time trying to interpret what the residents want through communication channels.

Technology can enable many improvements in delivering government services. To list a few: more efficient use of resources such as automated payroll systems; safety measure improvements through traffic signal systems; environmental conservation solutions through smart meters; and improved communications and customer engagement through social media. There is a negative side to open data programs however. The first is the need to protect the privacy of the personal data and the security of the data. Open data programs make government more vulnerable to criticism and susceptible to questioning so administrators need to be prepared for this challenge. Some question the releasing of the data at no cost when it clearly costs tax dollars to collect the data initially and then to make it available through online means. The dynamics of accessing and providing government data and information is rapidly changing as cities' administration models transition from older, more traditional government models, to new public management in response to new urbanism movement.

New urbanism is defined as a "movement that calls for the development of compact neighbourhoods containing community facilities and stores, with transit close by" that was intended to prevent further sprawl from the city core and to preserve communities and

neighbourhoods.[16] From the millennial generation to the new creative economy, cities are challenged to be more open and more participatory than ever before while still maintaining the integrity and sanctity of government regulations and security. There is a need for further study on the implications of acquiring and analyzing the vast amounts of data collected by government actors. It is important to ensure frameworks are in place to not only protect individuals' privacy but also their right to consent or not consent to the use of their personal data.

Literature Review

The research paper is intended to study the challenges of public policy development for local government administrators regarding open data and big data programs. This is a relatively new program for the public sector and specific research is limited. Published research provides many benefits for the development and delivery of these programs, particularly for the improvement of government transparency and accountability. There are several legal and regulatory frameworks that are relevant to open and big data programs, including the Municipal Freedom of Information and Protection of Privacy Act (MFIPPA) governing municipalities, the Health Information Protection of Privacy Act (HIPPA) which governs the health care sector, and the Freedom of Information and Protection of Privacy Act (FIPPA) that governs all other organizations. Provincial and Federal Privacy Commissions in Canada support these policy frameworks, yet different interpretations of these frameworks exist. Each municipality has its own policies regarding the protection of records and information. Some e-government policies cover open data or big data and yet it appears that the level of expertise and capacity does not identify all the risks and concerns in every organization. Government actors are very aware of the implications of the Municipal Freedom of Information and Protection of Privacy Act and the overall theme of "Privacy by Design" as developed by the Ontario Government Office of the information and Privacy Commissioner and as a result are typically cautious when posting

---

[16] Tindal and Nobes Tindal, *Local Government in Canada*, 76

information to public facing websites. The demand for open data creates new pressures and attitudes towards publishing data. The research is limited as to whether or not appropriate policy frameworks are in place to reduce risks of compromising privacy and security. A report submitted to the House of Commons, researched "the efforts and measures taken by social media companies to protect the personal information of Canadians." The findings suggest that individuals willingly provide their personal information to social media sites but "they may not fully understand the way their information is used, or the associate[d] privacy risks."[17] This report makes recommendations to the House that guidelines and frameworks be developed that will help to educate individuals about their rights to privacy and to provide access to a person's own data when needed. The report further identifies the challenges of maintaining pace with technology and providing greater authority to the Commission to manage and enforce the principles of personal privacy with private sector companies, particularly social media.

Big data provides other challenges for privacy and security policies as noted in the article, "Big Data and Local Government" by Graham and Daley.[18] Governments hold a vast amount of data, all collected to effectively manage and deliver the services needed in each community. People willingly provide this data as they sign up for recreation programs, borrow a library book, or register for an event. People also must provide information if they want to live in that community through examples such as submitting taxes, water payments, or a vote during an election. Government is starting to recognize the value of this data through analytics that allow them to monitor trends, identify requirements and to forecast changes in a community's social patterns or issues. Graham notes controversially that such data analytics become almost Orwellian, where 'big brother' is watching over all the time. Graham notes that "Big data has the power to dramatically change how customers receive services, and ultimately to improve their

---

[17] Canada, Standing Committee on Access to Information, Privacy and Ethics. Privacy and Social Media in the Age of Big Data. Ottawa, ON, CAN: Canada. Parliament. House of Commons, May 8 2012, 3
[18] Graham and Daley, *Municipal World*, Spring 2014

experience with local government" however it also identifies bigger issues such as privacy and security of the individual.[19]

The Information and Privacy Commissioner of Ontario states that limits are required for data collection that specify the purpose, collection and use of personal information otherwise the risk of the data being used for other purposes or though re-identification becomes very high.[20] The report emphasizes the need for the individual to be able to control who accesses their personal information and for what purpose. The Commissioner identifies the risk of "ubiquitous mass surveillance […] detailed profiling and power imbalances which may ultimately lead to discrimination. The report refers to FIPPA, section 41(1)(b) and MFIPPA, section 31(b), which state that: "An institution shall not use personal information in its custody of under its control except, (b) for the purpose or for which it was obtained or compiled or for a consistent purpose." The report further states that use or disclosure is consistent only if "the individual might reasonably [expect] such a use or disclosure."[21] This appears to be a clear delineation of the use of data except some organizations freely and openly combine data through analytics in order to improve service levels or publish data, recognizing they have removed personally identifiable data elements. If enough datasets are published, it becomes possible to re-link data together to form more closely identifiable patterns of information.

The Organization for Economic Co-operation and Development (OECD) revised its guidelines for Data Protection Principles for the 21st Century in March 2014.[22] These guidelines suggest the responsibility for protecting data should shift from the individual that provides the data to their government or organization that receives the data, and that they be

---

[19] *ibid*

[20] Freedom of Information and Protection of Privacy Act, RSO 1990, s. 41(1)(a) and Government of Ontario Municipal Freedom of Information and Protection of Privacy Act
R.S.O. 1990, CHAPTER M.56 Section 31

[21] Cavoukian, A. Information and Privacy Commissioner of Ontario, Canada. *The Unintended Consequences of Privacy Paternalism.* March 5, 2014. 9

[22] Cate, F.H., Cullen, P., Mayer-Schonberger, V. Data Protection Principles for the 21st Century: Revising the 1980 OECD Guidelines. The Organisation for Economic Co-operation and Development (OECD). Oxford Internet Institute, University of Oxford. March 2014

more transparent how the personal data will be used when receiving it. The authors argue that people have more understanding of how their data is to be used and can make their own informed decisions about whether or not to provide this information. The benefits realized were weighed against the risk of threats to privacy.

Research shows examples of de-identified published data and yet when enough datasets are available, there are ways to re-identify or personalize the information. One example provided by Alistair Croll in the article "Big data is our generation's civil rights issue, and we don't know it," is when the City of London, England, published data that unintentionally lead to a user generating racial boundary maps.[23] Croll provided another example where American Express used a customer's purchasing history to adjust credit limits. Despite having an excellent credit rating, American Express restricted the customer's credit limits because "other customers who have used their card at the establishment where you recently shopped have a poor repayment history." One customer's credit limit was impacted due to the trend analysis of other customers. This sort of data analytics begin to border on discrimination or as Croll notes, compromising one's civil rights.

The Ontario government published a paper entitled "Open by Default" in March 2014, which an advisory team created the terms of reference which include providing "advice and guidance to the Minister of Government Services on the development and implementation of Open Government initiatives in Ontario and act as a sounding board for ideas and suggestions from the public."[24] This report encourages greater openness by government and provides many recommendations on how to achieve more open government, which includes increased public participation, more open government information, identifying data as a public asset and recommendations for long term programs. The interesting piece in this document is the conflict

---

[23] Croll, A. Big Data is our generation's civil rights issue, and we don't know it. Posted August 2, 2012, http://radar.oreilly.com/alistairc (accessed March 2014)
[24] Government of Ontario, The Open Government Engagement Team, *Open by Default*, June 2014 https://www.ontario.ca/government/open-default-new-way-forward-ontario (accessed June 2014)

with the standards established by the various Privacy Commissioners. They conclude that the different legislations that protect privacy are outdated and cumbersome. One of the recommendations is to 'reform' the Freedom of Information and Protection of Privacy Act to be better aligned with the open by default principles.

The article "Big Data's Big Unintended Consequences" by Wigan and Clarke identifies the risk of publishing big data without a policy framework to protect people's rights to privacy while still being able to take advantage of the innovation and knowledge that comes from the analysis of data, particularly government data.[25] They provide many examples of how data can be helpful but also provide examples of risks and concerns.

The Privacy Commissioner of British Columbia produced a report supporting several recommendations to help shape open government policy frameworks, which uses the FIPPA as the core principles. The report "Evaluating the Government of British Columbia's Open Government Initiative," provides several recommendations including the need for an external advisory board to provide advices on the data that should be published and that "government should incorporate access by design principles" in a change from the original "privacy by design" principles.[26] The report identifies the need for open government and the publication of data and yet within frameworks to protect the privacy of individual's. This report provides strong background information supporting the need for policies and guidelines as governments move forward with open and big data programs.

There have been many books written extolling the need for open government, transparency and citizen engagement. Books such as "Grown up Digital" are reference material for the different generational needs and understanding of technology and personal security.[27] Tapscott writes about the 'millennial' generation who are more willing to share personal information online than the baby boomer generation. He writes that the 'net generation' "expect

---

[25] Wigan, M.R., Clarke, R., *Big Data's Unintended Consequences* IEEE Computer Society, June 2013
[26] Denham, E. *Evaluating the Government of British Columbia's Open Government Initiative*.
[27] Tapscott, D. *Grown up Digital* (New York, U.S.A.: McGraw-Hill) 2009

to collaborate with politicians … they want to be involved directly: to interact with them, contribute ideas, scrutinize actions, work to catalyze initiatives..."[28] These expectations will change the way government operates and communicates with the public. Tapscott also wrote "Macrowikinomics" in which he discusses how technology and innovation is changing the way we work, live, govern and create.[29]   It leads to the need for more open and engaged government and how government administrators and politicians need to be aware of the changes. This supports the new creative economy which drives many of the decisions and strategies made by local governments in urban environments. Moving away from the industrial or manufacturing workplace, municipal actors are redefining their strategies to attract new economic principles based on knowledge and creative sectors. Richard Florida writes about this creative economy as well. He writes in "Ontario in the Creative Age" that "the current economic transformation is as big and as challenging as the transformation from agriculture to industry. Our economy is shifting away from jobs based largely on physical skills or repetitive tasks to ones that require analytical skills and judgment." Florida contends that cities must change in order to survive. He states that "[c]ompetitive advantage and prosperity will go to those jurisdictions that can best prepare themselves and adapt to this long-run trend."[30]

Research into new urbanism and its impact on open government and open data programs begins back in the early 1980's with planners such as Jane Jacobs identifying the need for communities to be more self-sustaining and connected with the people. Richard Florida took this further with the creative community concept. The former Chief Information Officer of the United States, Vivek Kundra, now a Shorenstein Center Fellow of Harvard University published a paper called "Digital Fuel of the 21st Century: Innovation through Open Data and the Network Effect", which furthered the discourse. Kundra points out the value of

---

[28] *ibid.* 244

[29] Tapscott, D., Williams, A.D. *Macrowikinomics* (Toronto: Penguin Group, 2012)

[30] Florida, R., Martin, R. "*Ontario in the Creative Age*," Martin Prosperity Institute, February 2009 http://martinprosperity.org/2011/01/01/ontario-in-the-creative-age/ (accessed June 2014)

data in this information age and economy through the paper.[31] He states that "data is power" and that "open data is the fuel that powers important decisions at each level of society – from government, to business, to community to households." He also claims that the power of data will help to improve accountability of government, fight government corruption and enhance government services. He further states that open data programs will "change the default setting of government to open, transparent and participatory." [32]

In the book "Citizenville", the former Mayor of San Francisco discusses how cities are changing and how city administrators need to adapt to meet these changing needs.[33] He notes

> "opening up data will (1) create more trust in government, (2) provide opportunities for entrepreneurs, and (3) create jobs. There's a fourth benefit: Open data empowers people. The boss always had more information than the employee. The elected official always had more information than the voter. The police had more information than the citizens. The director had more information than the people he ruled. It used to be 'big brother is watching you,' now its 'we, the citizens are watching you." [34]

He also states that "we have entered an age when technology is enabling people to do whatever they want, whenever they want. Government, too, must offer that option." [35] Newsom's theory is that people will become more engaged in their government as they become empowered with the information made available to them through the use of technology and tools such as open data.

The literature review provides examples of the benefits of open and big data programs as well as identifies the risks and threats to personal privacy. The research indicates that organizations are quickly moving towards openness and transparency through the programs and yet there appears to be a contradiction in policy development and the actions being taken.

---

[31] Kundra, V. "*Digital Fuel of the 21st Century: Innovation through Open Data and the Network Effect.*" Harvard University, John F. Kennedy School of Government. June 2012
[32] *Ibid.*
[33] Newsom, G., Dickey, L. "*Citizenville: How to take the Town Square Digital and Reinvent Government."* (New York: Penguin Group, 2013)
[34] *ibid.33*
[35] *Ibid.*161

New Urbanism

New urbanism is a movement developed to address the issues created by urban sprawl, which caused the growth of suburbs and the decline of some inner cities. The goal of the new urbanists is to promote all community activity to occur within the neighbourhood encouraging economic growth, environmental preservation as well as developing a social environment that creates a healthy and vibrant community. New urbanism "values compact urban form that encourages pedestrian activity and minimizes environmental degradation; encourages social, economic, and land use diversity as opposed to homogeneity; connects uses and functions; has a quality public realm that provides opportunities for interaction and exchange; offers equitable access to goods, services, and facilities; and protects environmental and human health. New urbanism is against disconnected, automobile-dependent, land consumptive, environmentally degrading, single-use, homogenous, inequitable, and inaccessible urban places." [36] Open government and the implied use of technology is also very important in the development of the new 'creative city' or the city as defined by new urbanism theories. Cities evolved over the years because of different technological innovations shaping their development, such as the early industrial revolution, Fordist manufacturing/mass production age and now a city's growth is more influenced by information and communications technology. Richard Florida believes that the leading cities will be those that take advantage of the knowledge industries and knowledge workers with innovative new businesses being developed and encouraged. These knowledge and information workers are part of the new creative class which is critical for the continued social and economic well-being of cities in today's information technology and communication age. [37] An open government program that utilizes technology in such a way as to make information available, provide online access to services and engages through technology will

---

[36] Talen, Emily. *"New Urbanism." The Encyclopedia of Housing.* Ed. Andrew T. (Carswell. 2nd ed. Thousand Oaks, CA: SAGE Publications, Inc., 2012. 513-18. *SAGE knowledge.* Web) 19 July 2014

[37] Florida, R. *Startup City: The Urban Shift in Venture Capital and High Technology* report. Martin Prosperity Institute, Rotman School of Management, University of Toronto, March 2014

help the development of the creative city. The value of open data is not just for those who are technically skilled to write the applications and reports but it is also for those who will use those applications in their daily lives and to create an understanding of government operations and information. David Eaves, Canadian open government activist, said that "historically we [did not] build libraries for an already literate citizenry. We built libraries to help citizens become literate. Today we build open data portals not because we have a data or public policy literate citizenry, we build them so that citizens may become literate in data, visualization, coding and public policy." [38] This new creative city is dependent on information, raw data and technology and open data is a key part of this new dynamic.

Net Generation

Don Tapscott first identified the 'Net Generation' in his book "Growing up Digital." He identified this generation as those people born between the years 1977 - 1997. They grew up with technology readily available to them and they view it as an integral part of their world. They adapt easily to technology and expect there can be a technological solution to any problem. Tapscott notes this generation is more likely to turn on their computer than they will a television and that they are very active participants in events, particularly online events. [39] The Net Generation is involved in government which is evident from the 2008 presidential election in the United States. Obama's campaign included many online efforts to engage young people in the election. This campaign is often used as an illustration of the successful use of technology, in particular online engagement and participation. Tapscott notes this generation wants to be involved, interact, contribute and scrutinize their government. The "will insist on integrity from politicians" and this can be achieved through Open Government programs. [40] A challenge when

---

[38] Eaves, D. *Learning from Libraries: The Literacy Challenge of Open Data* http://eaves.ca/2010/06/10/learning-from-libraries-the-literacy-challenge-of-open-data/ (accessed May 2014)
[39] Tapscott, *Grown up Digital*, 21
[40] *Ibid.* 244

considering this generation is their lack of concern about privacy and protection of their personal information.

Tapscott writes that the younger generation, are accustomed to using technology and having the information readily available to them, want to be more active in decisions being made that affect them. They may not want to become politicians but they want to contribute to the development of new policies and to assist in government decision making. Through the use of technology tools, and analysis of available data, the net generation will want to influence government services and programs.[41] The Net Generation brings different expectations of accessibility and availability of information, particularly from government yet they also have very different expectations on personal privacy. They are more willing to share personal information online and less concerned with security risks. These changing social values present a new challenge to policy makers as privacy and security become less important, in correlation with more of the net generation entering the work force. The older generation (referred to as 'baby boomers') and their concern for privacy becomes less of an issue as they leave the workforce. These different social values for the different generations become a significant factor in open government initiatives and in policy development. The challenge is ensuring the policy provides enough protection to satisfy some of the public and yet is open enough for those who are not as concerned, all while still providing some form of security.

Challenges, Risks, & Ethical Considerations

Government collects a great deal of data in its normal business operations and it tends to justify increased data collection and analysis because of security concerns. A prime example is the recent National Security Agency (NSA) scandal in the United States. The challenge is in establishing a framework that prevents the drawing of inferences from the data, which results in basic generalizations rather than hard evidence. This can create negative circumstances for

---

[41] Tapscott, *Grown up Digital*, 308

individuals if they happen to be caught in the analysis. There are instances of people being denied financial services or government services based on these assumptions.[42] Policies can be developed to manage the open data and big data programs, in particular around issues such as the protection of personal information. Yet technology is changing so rapidly that it is not possible for the policy makers to keep pace. A policy is developed in line with MFIPPA to protect the use of data today but a new technology is developed that allows a manipulation of the data in such a way that it correlates data providing new outputs that is not foreseen before. Evolving technology creates new tools that can integrate and correlate datasets that might result in outcomes that are not anticipated and may or may not be provided in true context as well. The dilemma is how to create a policy when all the tools are changing so quickly and when the participants in the program change often.

As important as it is to publish data to share knowledge, information and prove transparency, local government is responsible for ensuring all data in its custody is secure. The appropriate measures need to be taken to provide the utmost privacy and protection of personal information, as data is published. The published information is important as it establishes trust and credibility in the organization and the administrators. The more knowledge people have, the more engaged they can be in government decision making. The data is acquired, published and maintained through the use of public funds and publicly funded resources and yet, individuals and companies are able to use this data for their own commercial and/or marketing purposes and for financial gain. Government does not typically permit the private sector to resell other publicly funded resources or infrastructure and yet, the open data program encourages this practice. Local government provides many services to the public, some of which the public pays for directly and some through their property taxes. Most proponents of the open data program say that providing the data for free is important because it is for the 'public good.' The Canadian Government's Open Data program defines open data as "Open

---

[42] Wigan and Clarke, *Big Data's Unintended Consequences*

data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share alike."[43] The intent is that open data can be and should be used by anyone, in any way. The usual stipulation is that credit be given to the organization who originally published the data. This very definition and intent eliminates the concern about the resale of a publicly funded asset.[44]

The cost of open data programs is a consideration not often understood. Most people think that to publish data in an open format is a very simple, copy process and yet the process can be quite complex. There are resource constraints and sustainability issues but open data is meant to be provided at no cost to the public. The government actors determined that the social benefits outweigh the costs to develop and maintain the program, in defining open data as a public good. It is believed that residents and businesses who paid the taxes to their local government are the beneficiaries of the open data program as are the application developers and commercial businesses. The benefits all return to the community either in wages, local spending or in the increased growth of the local knowledge economy. The cost to develop and sustain the program is not always considered particularly when programs are developed ad hoc.

There are issues regarding the raw data being presented without context. A recent example was Google reporting of the spread of influenza. The report was considered overstated because key data was missing. Data was collected based on the number of online inquiries about flu-related terms. There was no context as to whether people actually had the flu themselves; perhaps they were searching for someone else or any other purpose. There was no contextual information collected to validate the data.[45] If the data is collected with consent and with the appropriate context, then it would be much more valuable in that it provides more

---

[43] Government of Canada, Open Data Program. http://data.gc.ca/eng/open-data-101   (accessed May 2014)

[44] Bates, J. Manchester Metropolitan university, The Journal of Community Informatics Volume 8, No 2 (2012) http://ci-journal.net/index.php/ciej/article/view/845/916#foot1   (accessed June 2014)

[45] Cavoukian, A. Stewart, D., Dewitt, B., Information and Privacy Commissioner, Ontario, Canada *. Using Privacy by Design to Achieve Big Data Innovation Without Compromising Privacy, June 10, 2014*

relevant information that could be further analyzed. This is also one of the concerns local government actors express about publishing open data. The publication of credit card use by administrators provides interesting information that requires context to be fully understood. A recent example was the publication of supposedly excessive use of a credit card by a senior employee of the City of Toronto. The published data is simply the credit card expenditures and one employee's card use appeared excessive, especially when compared to others. Once questioned, it was explained that this senior employee used their city-issued credit card to pay for expenses to assist residents when an emergency evacuation occurred in a part of the City due to a massive fire in a residential area. The costs were for emergency supplies, accommodations, food and other purchases fully warranted by the situation.[46] This example shows the need for context to explain the data and yet, publishing the context or explanation is not easily achieved alongside the data and is not typically included on open data sites.

Another consideration is the data is collected for one purpose and then used for another and this is interpreted by some as being contrary to the MFIPPA. In section 31(b), the Act states: "An institution shall not use personal information in its custody or under its control except, (b) for the purpose for which it was obtained or compiled or for a consistent purpose."[47] The Privacy Commissioner of Ontario states "in determining whether a use is 'consistent'" with the primary purpose, section 33 of MFIPPA provide that a use or disclosure will be considered consistent only if "the individual might reasonably have expected such a use or disclosure."[48] She further notes "when a consistent purpose cannot be established, Ontario institution may still use the personal information in their custody or control if the person to who the information relates identified that information and consent to its use."[49] Some municipalities interpret this to mean that if the government collects the data for one business purpose such as program

---

[46] City of Toronto interview by author, July 11, 2014
[47] Government of Ontario Municipal Freedom of Information and Protection of Privacy Act R.S.O. 1990, CHAPTER M.56 Section 31
[48] Cavoukian, *The Unintended Consequences of Privacy Paternalism*, 9
[49] *Ibid.9*

registration, then they cannot use that same data for any other purpose such as customer service. Other municipalities interpret this differently. They believe that if the municipality collects the data for municipal business, then it is available to be used for other purposes, such as overall customer service. This proves to be a fundamental challenge for some when considering what data can be provided. If consent is not provided for the data to be used for purposes other than why it is originally collected, then it should not be published as open data.

Further to the example of purpose, another risk is the requirement for consent for the use of the data. The Privacy Commissioner clearly states that an individual must give consent for the use of their data and it can only be used for the purpose for which it is collected. A recent example highlights this challenge.   During the recent ice storm in the Town of Newmarket, the Communications Department wanted to access various databases used by other departments to collect contact information from residents who were likely impacted by the storm and the clean-up measures. This information is considered to be private and not available for communications purposes because the residents did not consent to their information being used for any purpose other than the original intent. The information is provided for other purposes such as program registration, service inquiries or to acquire a licence. The Legal department advised the Communications Department that they needed to ask residents for their permission to create an emergency contact information database that residents could freely choose to participate in or not.[50]   The challenge is how to ensure consent is received for the data to be published in an open data program when it was originally collected for a specific operational purpose.

Another risk is the potential to establish links between datasets that can lead to inaccurate or inappropriate conclusions, profiling potentially leading towards discrimination is an example. In London, England, during a presentation on mapping the city, with layers of various types of data from bicycle commuters to public transit use, a reporter identified a potential

---

[50] Author's own files February 2014

problem. One of the layers displayed the most common surnames in the City, geographically represented on the map. This data is readily available, non-personal and non-identifiable to an individual and presented some interesting reading. The problem became apparent once the data was mapped and showed what could be considered as racially profiling neighbourhoods.[51] This was clearly not the intent when the data the administration originally published the data, yet it became an unexpected and undesirable outcome.

Another risk is re-connecting datasets to expose different outcomes or to identify private information. Most datasets contain unique identifiers in order to distinguish the record from another with similar information. These unique identifiers make datasets very easy to connect with other datasets so most organizations will remove this identifier before publishing online as a way to de-identify the data.[52] The challenge is that studies show that even without this unique identifier, very little information is required to re-identify individuals, especially with the greater number of datasets made available.[53] Typically inferences are made about the data but given the amount of data available through the vast variety of devices, it becomes relatively simple to re-identify or at least come very close to identifying an individual.[54]

The NSA incident regarding the United States Government monitoring of its citizens as well as others like the Canadian government requesting personal information of customers of the telecommunication carriers, are examples of governments accessing information in a concealed and secretive manner and are examples of the risk in publishing government open data. Less intrusive examples are most government agencies also collect information that allows for the tracking of transportation patterns, transit use, shopping patterns, and others. These highlight citizens' habits and preferences and can provide important trend analyses.

---

[51] Croll, A. *Private Data in Public Ways* . http://radar.oreilly.com/2012/08/big-data-is-our-generations-civil-rights-issue-and-we-dont-know-it.html Posted August 2, 2012
[52] Cavoukian, *The Unintended Consequences of Privacy Paternalism*, 15
[53] Wigan, M.R. "*Owning Identify – one or many – do we have a choice?*" IEEE Technology and Society, vol. 29. No. 2, 2010 page 33-38
[54] Cavoukian, A. *"A Primer on Metadata: Separating Fact from Fiction."* Information and Privacy Commissioner, Ontario, Canada July 2013, 4

Information is collected from questionnaires, surveys, registration forms and online activities. It is apparent this information is collected in order to better serve the customer based on their needs, but there is a point where the monitoring becomes viewed as government interference leading to the threat or perception of restricting freedom of speech and movement. Some people view any government review of public behaviour as intrusive and should not take place under any circumstance. Others see it as an inevitable outcome of today's use of technology and may provide their data for improved government services.

Another consideration is that once the data is made available, it will be analyzed and assessed in new ways, producing greater scrutiny of operations. There is the possibility of questions and possible gaps in the data presented that may cause concern among the community. One community group may discover another received funding and will then question why they do not, for example. It is possible service problems may be identified upon the analysis that will have to be addressed by the organization. The moral of the public sector employees can be negatively impacted by the increased scrutiny of the work they are doing, especially concerning when context is not available as possible explanations. These suggestions should be brought forward in a constructive manner so the municipality can identify new processes or improve services as opposed to intending to demean municipal workers.

Most data published in Open Data programs are about properties, land use, economic issues or general government programs. Much of the data in government systems is about specific people and this personally identifiable data must be kept private. It is considered inappropriate to publish the names of children participating in a recreation program or the actual account balances of people or local businesses, for example. Personally identified data is defined as the person's name, address, credit card information and the internet protocol (IP) address of their computing device. Lifestyle information is also considered private. This includes information such as race, religion, relationship stats, sexual orientation, political affiliations, friends and family members. Behavioural data is also considered sensitive and

should be protected. This includes the types of sites usually accessed, the amount of time spent on specific websites, and habits captured through loyalty programs.[55] The protection of these types of data is what needs to be considered through policy development and processes so there is no risk of identifying individuals through the Open Data program.

Benefits of the Program

There are key outcomes that support the implementation of an Open Data program in spite of the risks identified.   One outcome is that the stakeholders will be able to access data and applications that help them understand local government processes in such a way that improves transparency around decisions and service delivery. Stakeholders are also able to hold government accountable for the use of resources as they will be able to report on and analyze the data independently rather than being reliant on information that may be provided with a political bias. The data can also be interpreted by different stakeholder groups in the context of their own environment, be it personal, business or community based, which will result in new information being created and may be viewed as value-added or areas requiring further analysis. As more data becomes available and as more applications are created and shared, stakeholders become more engaged as they start using the data and information in ways that best serves their own needs and the needs of their community. These outcomes all lead to enhanced citizen engagement in local government. Enhanced engagement means that elected officials and administrators will be able to access more fulsome and potentially meaningful information in regards to what their community expects, providing clear direction on what public services are needed and where resources ought to be allocated.

Many municipalities analyze data as they manage their programs. Analyzing participation in a recreation program will determine whether or not to continue offering the

---

[55] Bradshaw, S., Harris, K., Zeifman, H. *Big Data, Big responsibilities: Recommendations to the Office of the Privacy Commissioner on Canadian Privacy Rights in a Digital Age*. Waterloo, Ontario: Centre for International Governance Innovation http://site.ebrary.com/lib/uwo/Doc?id=10746713&ppg=1 (accessed July 2014)

program, which is accomplished by reviewing the available data. Reviewing the data in the water usage system helps to identify if there are problems such as leaks or breaks in the infrastructure. Reviewing the data of bylaw enforcement programs helps local administrators to determine if there is a growing trend in a particular infraction and perhaps further officers are required. By-law enforcement and administration use aerial photography to identify backyard pools to ensure compliance with local bylaws. This is small scale analysis and the information tends to be used internally.

There are as many benefits to the open data program as there are risks. A policy framework will establish strong processes that will influence how data is identified; collected and published and that will meet the demands for openness and transparency, yet still provide security and protection of personal information. The challenge is in having the policy developed when the program seems to be moving forward anyway. Administrators should not be ignoring the risks of the program as they attempt to respond quickly to the demands to publish and yet it is difficult to create a policy in an environment that seems to be changing daily.

Research Methodology

For the purpose of this paper, the research focuses on the larger urban cities in Ontario with a population of greater than 180,000.[56] Only large urban municipalities were contacted as they seem to be more advanced in the movement towards open government, likely due to the pressures of new urbanism and redevelopment pressures. As this topic is currently evolving and changing almost daily, the collected data reflects a moment in time. The surveys were sent to the lead of the open data program if that is apparent through the website analysis. Where this information is not available, the head of the Information Technology (IT) department were contacted as it appears that IT most often leads the project. Recipients were asked to forward

---

[56] Population counts were collected from Statistics Canada, Population and dwelling counts, for Canada, provinces and territories, and census subdivisions (municipalities), 2011   Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=302&SR=1&S=51&O=A&RPP=9999&PR=35&CMA=0

the email to a more appropriate person if necessary. All data was collected and interviews conducted in June to July 2014.

Preliminary research was conducted via an analysis of the City's websites. This analysis showed that of the twelve municipalities studied, nine published open data programs. All nine also published some form of policy but these are terms of use policies as opposed to a framework for publishing or protecting data. Three reference the Privacy by Design principles in their methodology and six reference MFIPPA as governing law. Interestingly, a review of the published datasets shows all municipalities publish only property related information, transportation, budget, or 311 call centre statistical related data at most. None publish any data that can correlate to customer based datasets or programs. This may be a clear indication of the concern about accidentally publishing personal information. One of the recommendations of the Open by Default report and supported by the Ontario Privacy Commissioner is the requirement to publish the inventory of the types of personal information that the government collect but does not publish or release for privacy, security or other legal reasons.[57] This is seen as an additional way to protect personal information while still providing open access. Table 1 summarizes the preliminary findings.

Table 1

---

[57] Government of Ontario, The Open Government Engagement Team, *Open by Default*, June 2014 https://www.ontario.ca/government/open-default-new-way-forward-ontario (accessed June 2014), 34 recommendation 2-3

City	Population	Open Data Program Y/N	Policy online	Privacy by Design	MFIPPA FIPPA PHIPA	Apps online?	Data Types Published
Toronto	2,615,060	Y	Y	Y *	Y	Y	Property 311 stats
Ottawa	883,391	Y	Y	N	N	Y	Property 311 stats
Mississauga	713,443	Y	Y	N	N	N	Mostly PDF Property, Planning
Brampton	523,911	Y	Y	Y	Y	N	Transit Energy consumption
Hamilton	519,949	Y	Acceptable Use	N	Y	N	GIS only
London	366,151	Y	Y	N	N	Y	Property
Markham	301,709	N	n/a	n/a	n/a	n/a	n/a
Vaughan	288,301	N	n/a	n/a	n/a	n/a	n/a
Kitchener	219,153	Y	Y	N	Y	N	Property
Windsor	210,891	Y	Y	N	Y	N	Property, budget, 311 stats
Richmond Hill	185,541	N	n/a	n/a	n/a	n/a	n/a
Oakville	182,520	Y	Y Use licence	N	N	N	Property, budget, schedules

A survey was sent to the nine municipalities who published Open Data Programs once the preliminary research was completed. The survey asked if a follow-up interview by telephone would be possible. A copy of the survey is attached in Appendix A. Of the seven who responded to the survey, six offered to be interviewed for further details. These interviews provided a broader understanding of the issues and barriers facing the implementation of the open data program. A list of the questions asked during the interviews is attached in Appendix B. Interviews were conducted with the municipalities who fit into the large category and yet do not have a published open data program to understand their reasoning for not moving forward with the program. Interestingly, the three municipalities who do not offer open data programs

are all lower tier municipalities in York Region. An open data program does exist at York Region, so they were included as part of the interview process.

Findings and Analysis

The focus of the research study is to examine the development of policy around the evolving open data programs for local government. The intent is to identify if there are necessary factors ensuring the policy development process is more likely to be implemented along with the program. The initial thought is that if the program is initiated by Council as part of a formal directive, then the likelihood of policy being implemented is higher. If staff initiates the program in response to trends or peer pressure, then it is more likely that any policy is ad hoc as opposed to formal. It is also more likely that the program is properly resourced if Council or a corporate Strategic Plan initiated the program rather than staff initiation.

The Ontario Municipal Act requires all municipalities to adopt a policy on Accountability and Transparency. Section 270 of The *Municipal Act, 2001* (the Act) requires "that all municipalities adopt and maintain a policy with respect to the manner in which the municipality will try to ensure that it is accountable to the public for its actions, and the manner in which the municipality will try to ensure that its actions are transparent to the public." [58] Of the nine municipalities with published open data programs, only two updated their Accountability and Transparency policy to specifically address Open Government or the Open Data Program. The anomaly is Toronto as they are governed by the separate City of Toronto Act. This Act also refers to accountability and transparency but Toronto did not included open government in the specific policy. They published Open Data policies separately instead.

---

[58] The Ontario Municipal Act, 2001 Retrieved from http://www.e-laws.gov.on.ca/html/statutes/english/elaws_statutes_01m25_e.htm#BK325

One of the first steps in the development of a policy is the clear setting of the agenda or purpose.[59] This is a challenge for most municipalities as the agenda itself is still in a state of constant change. Open data became prevalent in the private sector long before its consideration by local government and yet due to pressure of available technology and social media, the public demands more access to government data and information. The survey research shows that eight of the nine current programs are created by the technology departments with input from others, either other staff departments or local community activist groups. One of these, the City of Mississauga, is actually created by technology staff within the Planning department as a means to improve response times to standard queries from the public.[60] Only one program, in Toronto, originated from Council. It should be noted that the City of Ottawa's program received Council's approval in response to a staff report. Typically an open data program is not identified as a problem to be solved rather as a response to a trend or as part of the broader open government program. The causality of the "prevailing ideas and ideologies" as well as the socio-economic environment proves to be factors in whether a policy is formally developed or not.[61] A formal policy, other than the federal government open data licence, is not established in those municipalities where the program started in the more ad hoc way or simply as a response to a technology trend. It is interesting to note that such a public-facing program can be initiated from staff without Council approval. This is likely due to the fact that no additional budget is part of the original program design and staff completed the work with existing resources.

Forty-four percent of the survey respondents noted they followed the guidelines and principles of the Sunlight Foundation. The Sunlight Foundation is a not for profit organization based out of Washington D.C. They state their purpose as "The Sunlight Foundation is a

---

[59] Howlett, M., Ramesh, M., Perl A. Studying Public Policy Policy Cycles and Policy Subsystems Third Edition, Oxford University Press, Don Mills, Ontario 2009, 92

[60] Steve Czajka (City of Mississauga) in discussion with author, July 7, 2014

[61] *Ibid.* 99

nonpartisan non-profit that advocates for open government globally and uses technology to make government more accountable to all.[62] The foundation incorporates three main guidelines for Open Data Policies: what data should be public, how to make data public and how to implement policy. Checklists are provided to help program administrators determine what data is to be published and what data should not be. The checklist is helpful but it could not be considered a policy and the municipalities who use it do not consider it one.

The roles of the actors are very different in each studied municipality. In the majority of municipalities, the program is developed by staff, whereas in just one case it initiated from Council. Where initiated by Council, a policy typically is developed as part of the program. One municipality selected the agenda using a group of staff, politicians and community members. This community involvement is considered a best practice according to the Sunlight Foundation criteria. Where staff initiated the program on their own, the only true policy is the licence to use policy. All but one municipality adopted the Government of Canada's Open Data licence which is widely promoted as a best practice and a good standard to be followed. The common theme is that this licence provides enough safeguards to protect against the inappropriate use of data and the municipalities did not have to be concerned with issues other than the posting of data. The variety of actors involved in the decision making process resulted in different policies being established and different degrees of complexity and coverage for the policy. The result is an ad hoc style of policies being set at this time, evolving as new issues arise and must be addressed.

The assignment of resources appears to also be linked to who initiated the program. Only two municipalities use full time resources that are dedicated to the open data program and both of the municipalities started with some form of Council direction and approval. All the other programs are operated as an additional task for existing staff and within existing budgets. This helps to explain the lack of variety of datasets being published as most are property based sourced from the geographic information systems (GIS), which typically are published on

---

[62] The Sunlight Foundation retrieved from http://sunlightfoundation.com/

websites via online maps. Susan Chin Snelgrove of York Region's Open Data program noted that spatial data is less risky for publishing. It is already catalogued, already includes metadata and is already available online.[63] Other data types published tend to be more quantitative data such as basic organizational information, transit, transportation, budgets, expense, call statistics and documentation. Many seem to believe there is little to no cost to publishing data since most of this data is already created and much of it is already posted to the website in some form. Most of the datasets need to be converted from the proprietary systems into multiple open formats that are required by most open data programs however. Keith McDonald of the City of Toronto's Open Data program estimates that maybe just 10 to 20 percent of all possible data is actually made available to be published.[64] The program is not yet viewed as an important part of the operation; instead it is viewed as a one-time technology project or simply as extra work for as long as it is without permanent funding and resources.

One very poignant response is from The City of Windsor whose economy is struggling in response to recent market changes. Harry Turnbull, the CIO for the city, noted the IT department started the program in response to a Council directive for more open government. The open data program is seen by staff as a quick response to the current trend. There is very little feedback or use of the program and the interviewee noted it is likely because the community demographic is not of the creative or knowledge based businesses.[65] This example supports the hypothesis that the new creative community and new urbanism is a driving force for more open government and in particular the open data program.

The interviews provided for a more in-depth understanding of the program and of any concerns or barriers identified by the municipalities. Everyone commented that the most significant barrier is the lack of resources as well as the lack of corporate understanding of the project. When asked about the challenges they needed to overcome, the lack of data standards

---

[63] Susan Chin Snelgrove (York Region) in discussion with author, July 18, 2014
[64] Keith McDonald (City of Toronto) in discussion with author, July 11, 2014
[65] Harry Turnbull (City of Windsor) in discussion with author, July 11, 2014

not only within the organization but also across municipalities is mentioned by several. To become a truly open society, data should be accessible for every municipality so it can be compared and analyzed against others. McDonald noted all municipalities collect essentially the same data and yet they are all storing and using it differently.[66] The lack of standards makes this comparison very difficult. It is also difficult within the organization since data must be reworked and delivered into standard, open formats before it can be published. This adds to the workload and the delivery schedule. One of the barriers noted at the City of Windsor is a lack of demand from the community. Staff is creating new datasets and yet very few people are accessing them. It is difficult to maintain the program against so many other competing priorities for limited resources.[67] Chin Snelgrove of York Region noted that one of the barriers to their program is the lack of interest and input from the local municipalities. Chin Snelgrove estimates the Region holds just 20 percent of the total data for the area with the rest being created and maintained by the lower tier municipalities. In order to be able to provide a full spectrum of data, the data from the local communities is required.[68]

Two of the organizations mentioned they follow the Sunlight Foundation checklist when asked about the process for identifying what data should be published and how do they ensure no compromise of security. Czajka, the lead for the Open Data program in Mississauga, uses this checklist as part of their process. There is an ad hoc team who research different data sources or a potential topic through a peer review. The results are sent to him who reviews it against the checklist and then he publishes the data once he receives final approval from the operational department Manager.[69] Two other municipalities formed an ad hoc group of staff who are interested in the program. Turnbull identified there is such a group with representatives from different departments who volunteered their time to identify new datasets for the

---

[66] Keith McDonald (City of Toronto) in discussion with author, July 11, 2014
[67] Harry Turnbull (City of Windsor) in discussion with author, July 11, 2014
[68] Susan Chin Snelgrove (York Region) in discussion with author, July 18, 2014
[69] Steve Czajka (City of Mississauga) in discussion with author, July 7, 2014

program.[70] The groups identify what data should be published based on questions they receive in their department. This list of data is then sent to IT to produce which is then sent to the Records Manager for scrutiny before it is returned to IT to be published. This is a very effective process and yet it is unfortunately unrecognized and could be disbanded at any time. There is likely no true implications for improperly posting data with such an ad hoc process either. McDonald identified a "SWAT team" approach that is used in New York City as being a good solution. They established a team consisting of programmers, developers, mappers and data experts that work with different departments to identify potential data that might be published and then they make it ready for the website. This takes the onus off the individual operating departments and gives the responsibility to a team dedicated to the program. This led to a highly successful open data program for New York. [71] Chin Snelgrove suggested it might be helpful to invite application developers in to show how they might use the data in addition to what data types they would like, to further develop the program.[72]

It appears that the most significant barrier to the open data program is the lack of understanding throughout the organization. If data is made ready for publishing at the initial capture, and each dataset design adheres to specific standards, then it becomes much easier to post data in a timely and safe manner. Departments also need to understand that it is easier to respond to requests from the public if the information is already available. Each municipality interviewed noted the number of requests through MFIPPA reduced since launching their program. One municipality noted that there is an issue with one department not willing to publish their data because they currently sell the data. This revenue source is approximately $300,000 per year and is significant enough that they are not willing to give it up.[73] The

---

[70] Harry Turnbull (City of Windsor) in discussion with author, July 11, 2014
[71] Keith McDonald (City of Toronto) in discussion with author, July 11, 2014
[72] Susan Chin Snelgrove (York Region) in discussion with author, July 18, 2014
[73] Steve Czajka (City of Mississauga) in discussion with author, July 7, 2014

interviewee hoped that at some point the decision is made to make all the data available at no cost through the open data program but that this likely requires a council decision.

McDonald pointed out the risk of the correlation of data that was not intended when asked about any concerns they might have about publishing data. He felt this is a potential problem that may become more of an issue as the technology becomes more significant. He also pointed out that the impression of public service is 'not the best right now' and an open data program may be intended to improve transparency but there is also the risk of exposing information that may be misinterpreted or misunderstood by the public or lead to a negative impression. He identified the need for context with some datasets but it is often not possible to provide this additional information as part of the program.[74] Chin Snelgrove identified the need for improved communications about the program internal to the organization. This would create a greater understanding and buy-in for the program and encourage people to be less resistant to sharing their data.[75] Czajka identified that his concern is the suppression of data. He noted that data from Statistics Canada is actually suppressed if it is sourced from too small an area that makes it easily identifiable. He believes that it is better to identify the data as is than it is to suppress it from publication. He is also concerned with the risk of developers overlaying multiple sets of data from different sources to come up with an answer or outcome that may or may not be correct.[76]

The large cities that do not have published open data programs were contacted to investigate if there is a specific reason why not. All respondents noted that an open data program is not a priority for Council. It is clear that staff is interested in implementing the program but there are so many conflicting priorities with limited resources that a project not supported by senior management or endorsed by Council could not be considered a priority. They hope that best practices will have emerged by the time they are ready to implement the

---

[74] Keith McDonald (City of Toronto) in discussion with author, July 11, 2014
[75] Susan Chin Snelgrove (York Region) in discussion with author, July 18, 2014
[76] Steve Czajka (City of Mississauga) in discussion with author, July 7, 2014

program. In the meantime, they are researching the track record of others, beginning to discuss the program internally and hope to include the program as part of their upcoming new strategic plans for information technology and communications for their organization. All the respondents displayed the same issues surrounding big data programs. There are many disparate systems in place and in order to do true data mining or analytics, they need to design the datasets into a data warehouse type structure and there simply are no available resources to move forward.

Three basic levels of policy development around the open data program have emerged. The first level is those that publish data in a machine readable format, but that information is already available on the website so it is seen as being low risk. These organizations tend to not have a specific policy or process around the open data program. There are also those who believe that all data should be open therefore a specific policy is not required. The second level is some form of policy developed but typically it is the licence to use the data and maybe the existing Records Management policy for internal paper and digital documents. They see no rationale for separating open data from any other data so a formal policy is not required. The third level applies to those that have a more mature process. These organizations have a defined data identification process that involves multiple stakeholders, both internal and external. There are clearly defined policies that identify what data can be published, how data is excluded from the program and how data is protected before it can be published. Resources are permanently assigned to the program and policy development is part of their responsibilities. The program will include the consultation and approval from council and senior administrators and readily involve the community as key stakeholders in the program.

It appears that the open data program is still too new and immature to have policy fully implemented or even evaluated. Research suggests that the policies will be further developed in response to problems or as issues arise. Tindal and Tindal identify that policy making is a very complex and "a far from tidy process." They note that it is difficult to develop policy when there is constant change and new problems and service demands. They identify that "rather

than undertaking thorough research and analysis, the more likely reasons is to grope for a plausible remedy and hope it works better than previous responses."[77] This seems to be the response for policy development around the open data program. An incremental approach is clearly being undertaken and the policy will develop as the program continues to grow and change. A rational approach to policy development is not entirely possible with the evolving program. Tindal explains that a complication of municipal policy development is that so many policy issues are interconnected and interdependent which is also evident with this program.[78] Policies regarding openness and accountability may conflict with privacy protection and data security and yet they are all connected and impact each program. These complexities and relationships make it difficult to develop an effective policy that can react to the changing technologies and expectations.

The success of the program seems to be based on support from senior management from across the organizations, Council and the community. The internal advocacy seems to impact not only the success and depth of the program development, but also policy development as staff are able to impact policy decisions without Council direction. This seems to be the case with staff-led programs that change very quickly such as open data, big data and open government. A more formal process appears to take place when there are stakeholders from the community involved in the program development as there is a greater likelihood that council will then be actively involved. As the program matures and becomes more widespread, it is anticipated there will be greater involvement of all the stakeholders and thus, a more robust policy process. Open data concepts will be part of the everyday operation and will not be perceived as a 'technology project.'

---

[77] Tindal and Nobes Tindal, *Local Government in Canada*, 365
[78] *Ibid.* 367

Conclusion

There is increasing demands for transparency of government operations and accountability for the decisions made by politicians and public administrators. Residents want to be active participants in the community's development and expect assurances that public funds are being well-managed, notwithstanding the legislative changes at the various levels of government that focus on accountability and transparency. The Canadian government launched their Open Data program with the development of an Open Data Action Plan and the continual release of new datasets.[79] The Ontario Provincial Government recently embarked on an Open Data program, offering provincially acquired data online along with a portal for sharing applications.[80] The primary goals of the decision makers that lead to this Open Data program are to improve accountability and transparency and most importantly, to further engage the community and improve relationships among community partners. The Open Data program can help to establish greater trust in the local government and public administration and possibly raise the levels of compliance, as the community better understands the factors that impact the many decision making processes.

There are many issues for policy makers in the open data forum. The need to provide open data programs and to make all information accessible and transparent and yet the need to balance this with the protection of personal information and a person's right to privacy. Technology can make this accessibility very easy but the challenge is for the policy makers to ensure appropriate access and privacy protection. It is important to identify that dealing with different generational expectations present challenges for future consideration. As is the changing socio-economic pressures for the development of creative cities and knowledge communities causing changes in program delivery. Government distrust is another factor for

---

[79] Open Data Canada Action Plan http://data.gc.ca/eng/canadas-action-plan-open-government (accessed June 2014)
[80] Open Data Ontario http://www.ontario.ca/government/government-ontario-open-data (accessed June 2014)

the policy makers to consider as they balance the need for transparency and accountability with the need to provide effective context with the data being presented.

Once policies are developed, the changing technology, use of technology and expectations of the users, will require the constant evolution of the policy itself. A solution is the implementation of effective processes and practices that help to protect the information as it is being readied for publication on an open data site. McDonald from the City of Toronto pointed out the need to keep open data principles in mind right at the beginning when the data is being created. This will help to develop the program in a more accessible and efficient way than the current process which involves redesigning datasets and redeveloping formats. [81]

It is important for government to be accessible, transparent, and accountable when launching programs such as open data and big data, which provide significant benefits to all. Government program administrators need to address privacy and security concerns so they can assure their citizens of protection of personal information and that all ethical concerns have been addressed. Research shows there are different approaches being taken by administrators but common themes emerged:

- Open government is important but it needs to be cognizant of privacy and security.
- An informed and engaged public will help boost the creative local economy and improve interactive knowledge sharing.
- Transparency and accountability must be balanced with trust and credibility.
- The use of public funds is a trusted task but data should be recognized as a public asset and delivered securely and with privacy protected.
- Changing societal values means that the program and policies will be continuously evolving and emerging as requirements and expectations change.

There are strong advocates for open data programs and they put a great deal of pressure on government to publish datasets, yet they are downplaying the risks. It will prove interesting to determine if these risks are actually strategic ploys to gain access to the data in order to provoke change in government decision making. A central question for future research arises: Should taxpayer funds be supporting this program considering there are still has many

---

[81] Keith McDonald (City of Toronto) in discussion with author, July 18, 2014

unanswered questions, current constraints on government spending and the policy is still evolving? In comparison, are the principles so compelling that the risk should be considered minimal with the understanding that the policy will develop over time? The approach and development of government policy are never prepared for the onslaught of technology and the internet. Government actors need to be fully aware of the risks and implications of publishing data sets before making the decision to move forward with the program. A methodical approach to the policy design should be undertaken and should include all the different stakeholders in the design process. With the trends in technology evolving and the adoption by more and more people, government needs to be agile in their response and yet still they have the onus of ensuring they can protect the personal information of people before they lose this trust.

Work Cited

Bates, J. Manchester Metropolitan University, The Journal of Community Informatics Volume 8, No 2 (2012) Retrieved from http://ci-journal.net/index.php/ciej/article/view/845/916#foot1

Bradshaw, S., Harris, K., Zeifman, H. *Big Data, Big responsibilities: Recommendations to the Office of the Privacy Commissioner on Canadian Privacy Rights in a Digital Age.* Waterloo, Ontario: Centre for International Governance Innovation http://site.ebrary.com/lib/uwo/Doc?id=10746713&ppg=1 (accessed July 2014)

Canadian Broadcasting Corporation (CBC) CBC news reports accessed from http://www.cbc.ca/news/politics/telecoms-refuse-to-release-information-on-private-data-given-to-feds-1.2626286

Cate, F.H., Cullen, P., Mayer-Schonberger, V. "*Data Protection Principles for the 21st Century – Revising the 1980 OECD Guidelines*" March 2014 http://www.oii.ox.ac.uk/publications/Data Protection Principles for the 21st Century.pdf

Chin Snelgrove, Susan (York Region) in discussion with author, July 18, 2014

City of Brampton, http://www.brampton.ca

City of Hamilton, http://www.hamilton.ca/ProjectsInitiatives/OpenData/

City of Kitchener, http://app.kitchener.ca/opendata/

City of London, https://www.london.ca/city-hall/open-data/Pages/default.aspx

City of Markham, markham.ca

City of Mississauga http://www.mississauga.ca/portal/residents/publicationsopendatacatalogue

City of Oakville, http://www.oakville.ca/data/index.html

City of Ottawa, http://ottawa.ca/en/mobile-apps-and-open-data/open-data-ottawa

City of Toronto, http://www1.toronto.ca/wps/portal/

City of Vaughan, vaughan.ca

City of Windsor, http://www.citywindsor.ca/opendata/pages/open-data-catalogue.aspx

Civic Access, http://civicaccess.ca/ (accessed June 2014)

Croll, A. "*Big data is our generation's civil rights issue and we don't know it*" August 2, 2012 http://radar.oreilly.com/2012/08/big-data-is-our-generations-civil-rights-issue-and-we-dont-know-it.html

Croll, A. *Private Data in Public Ways* . http://radar.oreilly.com/2012/08/big-data-is-our-generations-civil-rights-issue-and-we-dont-know-it.html Posted August 2, 2012

Czajka, Steve (City of Mississauga) in discussion with author, July 7, 2014

Denham, Elizabeth. *Evaluating the Government of British Columbia's Open Government Initiative,* Victoria, BC, CAN Office of the Information and Privacy Commissioner for British Columbia, 2013.

Eaves, D. Learning from Libraries: The Literacy Challenge of Open Data accessed from http://eaves.ca/2010/06/10/learning-from-libraries-the-literacy-challenge-of-open-data/

Florida, R. *Startup City: The Urban Shift in Venture Capital and High Technology* report Martin Prosperity Institute, Rotman School of Management, University of Toronto, March 2014

Florida, R., Martin, R. "*Ontario in the Creative Age,*" Martin Prosperity Institute, February 2009 http://martinprosperity.org/2011/01/01/ontario-in-the-creative-age/ (accessed June 2014)

Gartner Research, 2013 Accessed from http://www.gartner.com/it-glossary/big-data/

Government of Canada - Open Data Plan, Accessed from http://data.gc.ca/eng

Government of Canada, Standing Committee on Access to Information, Privacy and Ethics. Privacy and Social Media in the Age of Big Data. Ottawa, ON, CAN: Canada. Parliament. House of Commons, May 8 2012

Government of Canada Freedom of Information and Protection of Privacy Act, RSO 1990, s. 41(1)(a)

Government of Ontario Municipal Freedom of Information and Protection of Privacy Act R.S.O. 1990, CHAPTER M.56 Section 31

Government of Ontario, The Open Government Engagement Team, *Open by Default*, June 2014 https://www.ontario.ca/government/open-default-new-way-forward-ontario (accessed June 2014)

The Government of the United States of America, The White House, Open Government Initiative. http://www.whitehouse.gov/open (accessed March 2014)

Graham, K and Daley, M *"Big Data and Local Government: Turning potential into Practice" Municipal Monitor* Spring 2014

Howlett, M., Ramesh, M., Perl A. Studying Public Policy: Policy Cycles and Policy Subsystems Third Edition, Oxford University Press, Don Mills, Ontario 2009

Information and Privacy Commissioner, Ontario, Canada, "*A Primer on Metadata: Separating Fact from Fiction*" July 2013 by Cavoukian, A. http://www.ipc.on.ca/English/Resources/Discussion-Papers/Discussion-Papers-Summary/?id=1316

Information and Privacy Commissioner, Ontario Canada "*The Unintended Consequences of Privacy Paternalism,*" March 5 2014 by Cavoukian, A.

http://www.ipc.on.ca/english/Resources/Discussion-Papers/Discussion-Papers-Summary/?id=1381

Information and Privacy Commissioner, Ontario, Canada *. Using Privacy by Design to Achieve Big Data Innovation Without Compromising Privacy, June 10, 2014 by* Cavoukian, A. Stewart, D., Dewitt, B., http://www.ipc.on.ca/images/Resources/pbd-big-data-innovation_Deloitte.pdf

The Institute for Citizen-Centred Services, *Citizens First 6 Report* http://www.iccs-isac.org/research/citizens-first/citizens-first-6/?lang=en

Kundra, Vivek, January 2012, *Digital Fuel of the 21st Century: Innovation through Open Data and the Network Effect.* Joan Shorenstein Centre on the Press, Politics and Public Policy, Harvard University, John F. Kennedy School of Government.

McDonald, Keith (City of Toronto) in discussion with author, July 11, 2014

Newsom, G., Dickey, L. "*Citizenville: How to take the Town Square Digital and Reinvent Government.*" (New York: Penguin Group, 2013)

The Regional Municipality of York, York.ca

Statistics Canada, Population and dwelling counts, for Canada, provinces and territories, and census subdivisions (municipalities), 2011 http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=302&SR=1&S=51&O=A&RPP=9999&PR=35&CMA=0

The Open Government Partnership, http://www.opengovpartnership.org/ (accessed July 2014)

The Sunlight Foundation, http://sunlightfoundation.com/ (accessed June 2014)

Talen, Emily. "New Urbanism." *The Encyclopedia of Housing.* Ed. Andrew T. Carswell. 2nd ed. Thousand Oaks, CA: SAGE Publications, Inc., 2012. 513-18. *SAGE knowledge.*

Tapscott, D. The Toronto Star, *Big Ideas* in partnership with Martin Prosperity Institute. *"Rob Ford? John Tory? Olivia Chow? Here's what matters"* March 14, 2014

Tapscott, D. Williams, A.D. 2012 Edition *Macrowikinomics*: Canada, Penguin Group.

Tapscott, D. 2009. *Grown up Digital*, New York, U.S.A.: McGraw-Hill

Tindal, R.C., Nobes Tindal S. 2009. *Local Government in Canada*,7th Edition Canada: Nelson Education Ltd.

Turnbull, Harry (City of Windsor) in discussion with author, July 11, 2014

Wigan, M.R. *"Owning Identify – one or many – do we have a choice?"* IEEE Technology and Society, vol. 29. No. 2, 2010 page 33-38

Wigan, M., Clarke, R. "*Big Data's Big Unintended Consequences"* Published by the IEEE Computer Society June 2013

Appendix A

Online Survey

1. Who led the Open Data Project in your municipality?

    a. Staff? IT or other

    b. Council

    c. Strategic plan/objective

2. How was the Open Data Program created?

3. Why was the Open Data program created?

4. Do you have policies in place for the Open Data Program?

5. If yes, what policies do you currently have?

6. Who manages and maintains the program now?

7. Does your organization perform any analytics or data mining on your City-owned data?

8. If yes, who has access to the outputs and/or reports?

9. Do you have any comments or thoughts on the future of open data and big data programs in local government?

10. Would you be available to discuss the Open Data program in more detail?

Appendix B

Interview Questions

1. Did you have a decision process to determine which datasets are published?

2. Have you received any feedback from the public – is the data being accessed as much as you thought?

3. Do you have any concerns about security, risk exposure, etc. with the published data?

4. Have there been any issues to overcome?

5. You publish primarily property or spatial data and 311 call stats. Is this deliberate?

6. Are there any barriers to moving your program further?

7. Has there been any political involvement?

8. Is the program formally resourced on a full time basis?