

TORONTO ELECTION 2010:  
Discussion Paper #6 — September 2010

# Open Data, Open City

MARTIN  
**Prosperity***Institute*

 Cities Centre  
UNIVERSITY OF TORONTO

# OPEN DATA, OPEN CITY

## WHAT IS OPEN DATA?

We live in a knowledge economy. We access, evaluate and compare many kinds of information every day, from checking bus schedules to sorting emails by priority. Cities deal with lots of information too, and most of it can be classified as “data,” collections of information generated and used by people and computers. The availability and use of data by and for the public is an emerging municipal issue as cities realize the many benefits of “opening up” their data.

Open data is both a philosophy and a practice. Municipalities collect extensive data about the city and its residents, and the practice of freely sharing this data online is gaining ground in cities across the country and around the world. Vancouver and Washington, D.C. both have official open data policies, and many other cities are launching their own open data initiatives. “Data” here means everything from statistics to address lists and recycling schedules, from information on election results by riding, to the locations of schools or streetlamps. Since such information is important for both the governance and understanding of a city, a clear policy on how that data is stored, used, and presented to the public is necessary.

The call for open data goes a step further than simply requesting access to the numbers; “open data” describes data that is digital and flexible, so that it can be shared and integrated into computer applications. It must be available in a machine-readable format<sup>1</sup>, so that it can be downloaded and manipulated by people who wish to use it. Open data advocates want access to the “raw” information so that it can be used and reused with other data sets and by any number of computer applications. For instance, the city’s water-use statistics<sup>2</sup> may be available as a list in a PDF document, which can be read by people, but which a computer program cannot interpret. If these statistics were in a format that a computer could read, for instance as a spreadsheet of populated columns, one could play with the data, looking for patterns, or comparing it with other sets. Homeowners worried about their water consumption could compare their use with neighbourhood averages, or even be automatically emailed their month-to-month statistics. Such services are but one example of the possibilities of open data for citizens and the city.

In order to realize the potential of a large, well-organized, and accessible collection of information, open data initiatives require planning and commitment. The immediate benefits and long term potential of open data include more efficient city operations, more interaction between the city and its citizens, a more data-literate citizenry, and the opportunity for entrepreneurs to use city data innovatively for public or proprietary use. Open data promotes democracy, fuels entrepreneurship and improves efficiencies by repurposing data for many uses.

## OPEN DATA AND CIVIC ENGAGEMENT

While open data may seem at first glance like something reserved for technology geeks, the intention of the open data movement is the opposite of exclusivity: it aims to foster an understanding of government information for the average citizen. While not everyone will make use of the data, it’s important for citizens to know that it’s accurate and available, and that accessible data is a right, not a privilege. It is also an increasingly important skill for participation in the information and knowledge economy. As Canadian entrepreneur and open government activist David Eaves wrote in a June 10th blog post, historically “we didn’t 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

---

<sup>1</sup> Being in “machine readable format” means, as the name suggests, that the data can be recognized and interpreted by a computer. Examples of machine-readable technology range from an old-fashion library punch card to a supermarket barcode to a cassette tape.

<sup>2</sup> Cities collect statistics on the volume and frequency of water consumption.

citizenry, we build them so that citizens may become literate in data, visualization, coding and public policy” (“Learning from Libraries: The Literacy Challenge of Open Data,” <http://eaves.ca/>). We all need to understand how data is manipulated to support arguments and ideas. A basic level of statistical awareness is necessary so that we understand the context of the data presented: we all need to know to ask, where did the data come from? Was the author paid to produce it? What are their biases? Can I check that the results are correct?

The process of understanding information is often characterized as an upward progression from data, to information, to knowledge, and finally to wisdom. Similarly, data becomes more valuable—more *useable*—through collection, structuring, and distribution. Cities conduct extensive data collection and structuring, but much of the data are not shared, often because its value outside of its original intent is not recognized. Some city data is distributed to us formally: garbage pickup schedules, smog warnings, and public library locations are all data sets collected by the city and presented to the public in some form, usually on a flyer or website. But this data is static. For data to be open, it must be flexible and manipulable.

Philosophically, the open data movement has grown out of a larger push for government openness. Advocates of open data believe that in a modern democracy, governments should share the data they collect, and that open data is a concrete way to teach and encourage government participation. Accessible data means that citizens can see and evaluate the impact of city policies. Citizens will know where to go to find out how tax dollars are spent, they will be able to calculate the numbers themselves, and compare them to other years, other municipalities, or any other metrics that interest them. Our ability to fully engage as citizens is enhanced when we have the ability to access the output of government work. In a country whose official motto includes the phrase “good government,” it seems obvious that we should be able to know when our government is being “good.” Richard Poynder’s UK-based blog, “Open and Shut” makes this argument very well. With regard to *The Guardian* newspaper’s launch of the “Free Our Data” campaign, he writes:

“... [*The Guardian*] wanted to draw attention to the fact that governments and government agencies have been using taxpayer’s money to create vast databases containing highly valuable information, and yet have made very little of this information publicly available. Where the data has been made available access to it has generally been charged for [... and] released under restrictive copyright licences prohibiting redistribution, and so preventing third parties from using it to create useful new services. The end result, *The Guardian* believes, is that the number and variety of organisations able to make use of the data has been severely curtailed and innovation stifled.”

While it can be argued that hoarding data can curtail innovation, it can also be said that sharing too much information can cause problems, too. In political environments where every decision is criticized or ridiculed, civil servants will be understandably reluctant to share the information that led to their decisions. The challenge for cities creating open data initiatives is to release as much data as possible while ensuring legitimate privacy and security concerns are met.

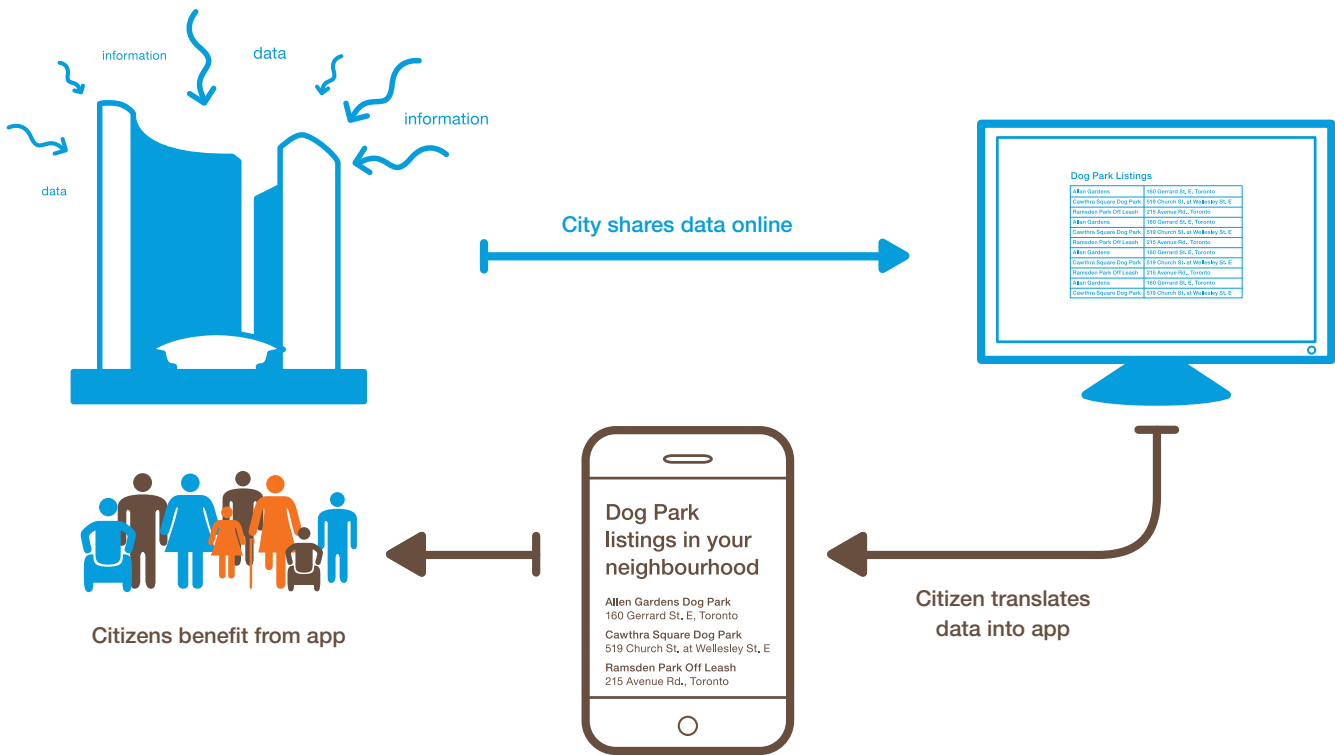
## OPEN DATA, ENTREPRENEURSHIP AND INNOVATION

A solid commitment to open data would not only benefit Torontonians on principle. Opening up data sets leads not only to increased accountability and trust but, for the entrepreneurial among us, innovation through the development of applications, most often in the form of “mashups,” digital products created from pre-existing elements. Programmers and developers create mashups by combining two or more technologies or data sources to create a new product or service; some of these services are created simply to benefit our fellow citizen, while others include an intention of profit. A simple version of a mashup would be a store’s “Is It In Stock?” feature, which combines location data and stock data to tell a customer if the product they want is available nearby. Toronto software developer Mark Headd used data sets on location and space availability of child care centres to create a daycare search by postal code using texting or Twitter. Mark describes the application as “quite simple, and is still rough in many ways, but it was completed within several hours and demonstrates how governments that release interesting and valuable information empower developers to build useful things” (“Toronto Opens Government Data,” [voiceingov.org/](http://voiceingov.org/)). This year Jeff Aramini, a former

epidemiologist at the Public Health Agency of Canada, developed HealthAndSafetyWatch.com, a website that tracks product and food recalls using aggregated federal and provincial open data in combination with media monitoring software. This information is then overlaid on a map, alerting users to recalls in their area. Several iPhone applications have been developed—some free, some fee-based—to deliver TTC schedule information on demand and in a mobile format. Examples of useful mashups abound and span the spectrum of generic public service tools to highly specialized proprietary applications. What these programs have in common is their reliance on flexible, open data.

Open data uses City money wisely, promotes entrepreneurship, and enhances civic engagement

Exhibit 1



Source: Martin Prosperity Institute, 2010  
Design & Illustration by: Michelle Hopgood

## THE IMPORTANCE OF STANDARDIZED OPEN DATA

Standardization is an essential, albeit expensive, component of this openness. Standardization means that the data sets can be more easily shared, but it also means that the city will have to invest dollars to standardize existing and legacy data systems in a more flexible format. While the start-up costs of standardizing data collection are high, both financially and in terms of human resources, over time the benefits to the city are manifold. David Eaves, who advises the City of Vancouver on the development of their open data policy, argues that data is a strategic asset that cities should be using to help themselves, that the centralization of data will help the city manage itself more efficiently. An excellent example of using standardization to improve both efficiency and user experience is Toronto's 311 service, which integrates city service responses into a single system, allowing for ease of access and faster response time.

The cost of developing systems to standardize, share, and archive data could be shared between cities, who could come together to negotiate with potential vendors of the appropriate tools. An open data policy is a mechanism for cities to “help themselves,” and, in addition, it is a tool for cities to learn from each other while encouraging important policy learning opportunities.

## OPEN DATA PORTALS AT WORK

Toronto is one of several Canadian cities aiming to publish its data in an open format catalogue, along with Vancouver, Nanaimo, Edmonton, and Ottawa. Most recently, London (Ontario) launched Open Data London and Montreal has announced Montréal Ouvert. The City of Toronto launched *toronto.ca/open* in November 2009. The catalogue offers about thirty datasets, most of them static geographic data, including ward, park and neighbourhood boundaries, locations of places of worship, traffic signals and food banks, and some TTC schedules. The sets are available in flexible digital formats and some data are automatically updated or available in real-time. The portal’s contents are available under an open license, which means that anyone, anywhere in the world, has the city’s permission to a “world-wide, royalty-free, non-exclusive license to use, modify and distribute the datasets in all current and future media formats for any lawful purpose”(City of Toronto Open Data project). Created by local data enthusiasts to support Toronto’s open data project, *dataTO.org* is a collaboration site inviting the open data community to request new data, suggest enhancements, and share in the development. Participation has been low, perhaps due to the limited number of data sets offered in Toronto’s Open Data catalogue. Additionally, despite the ease with which datasets can be released in many formats, the catalogue has typically chosen only one format per data set. The catalogue also uses internal city terminology (“cuts” for excavation work or potholes, “Toronto Centreline” for a set on road markings), terms not always familiar to citizens.

Operating since September 2009, Vancouver’s more advanced *data.vancouver.ca* currently offers over 100 data sets. But it’s not just the number of sets available that is impressive: the city has turned the portal itself into an interactive feedback and discussion space, and has built RSS feeds that aggregate city information such as job postings and bike route detours.

Across the pond, *data.london.gov.uk* acts as a hub, pointing visitors to raw data sets available on a number of agency websites. Data available includes life expectancy at birth by ward, household waste disposal by borough, and the locations of every polling station for the 2010 elections. The portal calls itself a first step, and promises that they are “committed to influencing and cajoling other public sector organisations into releasing their data here too.” (*data.london.gov.uk*). London’s datastore should be a source of inspiration for Toronto’s future efforts. At the federal level, Canada lags behind both the United States and the United Kingdom. In the US *data.gov*, founded in 2009, has over 270,000 data sets available. With the help of Tim Berners-Lee, credited with inventing the World Wide Web, the British government has launched *data.gov.uk*, “a key part of the Government’s Transparency programme.” The Canadian government has thus far shown less enthusiasm; there is no national open data strategy, and therefore no central hub for national data sets.<sup>3</sup> As the largest city in the country, Toronto has the resources to be a leader in government transparency, and to encourage the use of its data for myriad purposes, championing the value of a digitally literate citizenry.

---

<sup>3</sup> Not content to wait for our administrations to take the lead, a few Canadian innovators have attempted to address the gap. Their efforts include *openparliament.ca*, run by Michael Mulley, and *datadotgc.ca*, run by David Eaves, which he describes as his effort at “showing our government, and Canada, what a federal open data portal could and should look like.” *CivicAccess.ca* and *datailbre.ca* are two other websites involved in championing the necessity of open, free data in the name of democracy and transparency in the information age.

## POLICY IDEAS

A successful open data initiative needs more than just accessible data to be effective. It requires formal commitments from the city on a number of fronts, the development of a robust centralized portal, and enthusiasm from the general population. It also, particularly in the early years, needs a dedicated team with a variety of skill sets, including technophiles, marketers, and those who can best promote citizen engagement.

Cities who make data available and actively market what they are sharing demonstrate their belief in the benefits and potential of open data. In order to make open, accessible data sets part of standard operating procedures, cities need to have policies in place that will help standardize creation and facilitate pro-active distribution. All city data should be created in shareable formats, regardless of whether it is deemed allowable or worth sharing at that moment. Policy should include a written, official commitment to updating data sets, and to communicating these updates through a variety of channels. Certain sets, such as public transportation and election information, should be recognized as being priority sets, with more regular and formal updates required.

Toronto's open data catalogue has a lot of room to grow, and should be looking to the plethora of information on the city and its governance already available, and known to be of interest, through <http://toronto.ca/>. All of this data should be housed in the centralized online portal, allowing anyone with an idle curiosity to browse, download, and play with what is already public information on a standard platform. This portal must be designed to protect sensitive data, and to ensure that privacy concerns are addressed, investigated, and dealt with promptly. This will help build trust between the specific department overseeing the portal, City of Toronto employees, and the general public, and encourage the enhancement of the portal's development.

As important as it is to open up data, it's equally important for citizens to make use of it. Cities are unlikely to be convinced of the value of open data unless there's a community of enthusiastic users who can demonstrate its value. Every municipality is low on cash and resources, resulting in a focus on efficiency in delivering city services. Open data is a way for cities to create more value for every citizen; citizens benefit from the development of innovative and useful applications, and the city is viewed as having played a key role in offering value-added services and fuelling entrepreneurship. When cities open up their data, they are initiating a conversation with the people using the data, and the initiative will succeed as long as both sides are listening, participating and working towards a common goal.

## Conclusions

Open data initiatives work when the host city understands the benefits to both the citizens and to itself, and is willing to reap those benefits by committing to success through investment. When open, accessible data is harnessed by data-savvy citizens to create valuable tools that serve all kinds of needs, there are many wins: the city has encouraged and satisfied entrepreneurship; the entrepreneur has found satisfaction in creating a valuable tool, and the citizen benefits from the tool itself. Open data can enhance the relationship between the city and its citizens by sharing what it collects, and in return the citizens can offer value to the city by taking the time to create tools the city hasn't the time or budget to build. Bringing a sense of cooperation that is necessary for a healthy and progressive future, open data initiatives make cities and their citizens partners in growth and innovation.

## **SOME KEY QUESTIONS ABOUT OPEN DATA:**

**How do you see data literacy fitting into the education of citizens?**

**Where should open data collection, production, and monitoring sit within the City's structure?**

**How do you envision open data policies changing in the next five years?**

**Many civil servants fear open data initiatives because they believe that access to data will result in people (and the press) constantly undermining and second guessing all of their decisions. What is your opinion about this issue? How would you address City Hall, your new employees and staff, about this issue? How would you convince the various city departments of the importance of open data?**

**What would you do to encourage data sharing across the many different departments of the city?**

## ACKNOWLEDGMENTS

### OUR TEAM

#### Authors

Kimberly Silk  
Jacqueline Whyte Appleby

#### WEBSITES

[martinprosperity.org](http://martinprosperity.org)  
[www.citiescentre.utoronto.ca](http://www.citiescentre.utoronto.ca)

Thanks to **David Eaves** ([eaves.ca](http://eaves.ca)) and **Pete Forde** ([unspace.ca](http://unspace.ca)) for their input during the research phase of this paper.

### CONTACT US

Phone: (416) 673-8580  
Fax: (416) 673-8599  
Email: [info@martinprosperity.org](mailto:info@martinprosperity.org)

Martin Prosperity Institute  
Joseph L. Rotman School of Management  
University of Toronto  
MaRS Centre, Heritage Building  
101 College Street, Suite 420  
Toronto, Ontario, Canada M5G 1L7

Richard Florida, Director  
Kevin Stolarick, Research Director