HOUSE OF COMMONS
CHAMBRE DES COMMUNES
CANADA

Standing Committee on Government Operations and Estimates

OGGO • NUMBER 025 • 2nd SESSION • 41st PARLIAMENT

EVIDENCE

Tuesday, May 13, 2014

Chair

Mr. Pierre-Luc Dusseault

Standing Committee on Government Operations and Estimates

Tuesday, May 13, 2014

● (0845)

[*Translation*]

The Chair (Mr. Pierre-Luc Dusseault (Sherbrooke, NDP)): Good morning everyone. This is our 25th meeting. We are continuing our study on the government's open data practices, which is well on its way and indeed, nearly over.

We have several witnesses today, several experts: Ms. Francoli, Assistant Professor, Communication Studies, at Carleton University; Denis Deslauriers, Director of the Information Technology and Telecommunications Service, City of Quebec; as well as Alton?? Hollett, Assistant Deputy Minister, Economics and Statistics Branch, Department of Finance for the Government of Newfoundland and Labrador. We also have two witnesses joining us by videoconference from the City of Montreal a little later on: Mr. Chitillian, Vice-Chair of the executive committee responsible for administrative reform, youth, smart city initiatives and information technology; and Mr. Fortin, IT Strategy and Planning Advisor, Information Technology Service.

We will begin by hearing the witnesses present in the room, beginning with Ms. Francoli, who has 10 minutes for her presentation, just as all the others. Following that, once each witness will have spoken for 10 minutes at most, the committee members will ask the witnesses their questions.

I thank you for being here today.

Ms. Francoli, you have the floor for 10 minutes.

[*English*]

Professor Mary Francoli (Assistant Professor, Communication Studies, Carleton University, As an Individual): Thank you very much for inviting me to speak to your committee. I'm going to focus my opening remarks on the Open Government Partnership, or the OGP, in which Canada is a member.

The OGP secures commitments from governments to improve transparency, accountability, and citizen engagement; to fight corruption; and to harness technology to strengthen governance. A requirement of membership in the initiative is that each country agree to have an independent review of its national action plan and of its progress every two years. This is called the independent reporting mechanism, and it's part of a system of checks and balances built into the OGP. I'm the independent researcher for Canada. Our first progress report was published in February of this year. The foundation for that report and for my remarks today is based on stakeholder feedback that made up the bulk of the report.

Canada's national action plan to the Open Government Partnership focuses on more than just open data, but given the parameters of your study, I'm going to confine my comments to open data as much as possible.

There are a lot of different issues that I could talk about in relation to open data, but given the limited time, I thought I would focus this morning on some of the main areas of concern that users raised during my stakeholder interviews and meetings. So it's a bit of a critical analysis of our open data strategy. I'm not speaking as much to some of the positive things, but I'm certainly happy to speak to those later during the questions and answers.

I organized my comments today around seven main concerns or points that the majority of stakeholders that I spoke to during the course of my study raised as issues with regard to the open data strategy.

The first is the diversity of data sets. Currently the data.gc.ca portal is largely dominated by geospatial data. There are few to no data sets in many other areas, including employment insurance, health, and issues related to specific demographics such as seniors or aboriginal persons. A lot of the users that I spoke to during the course of my study found that quite limiting, just the nature of the data sets themselves.

The second point is the quality of data. A couple of the points relate to that. Quality of data was perhaps what the majority of stakeholders were most concerned about. There is a widespread belief that the quality of the data in the data portal will suffer, and will continue to suffer in the long term, as a result of steps that have been taken to cut data collection at its point of origin.

A prime example of this—and I have to say this was the example that was given by almost everybody I spoke to during the course of my study—is the cancellation of the mandatory long-form census. Those sorts of measures around data collection have led to concern about the availability of updated and comparable data sets at smaller units of geography in the future. We're already starting to see that be the case. Even in the last few days there have been a few news stories and reports about the loss of data from the last census exercise.

Those are the first two points.

The third point is the fragmented nature of the data sets that are found on the data.gc.ca portal. Data users noted that it isn't uncommon for data sets to be released in what they said were bits and pieces instead of in complete and wider-reaching data sets. Sometimes they said they're also separated from their methodology and their quality description. What data users were finding is that when they were trying to work with the data, they had to spend quite a bit of time and really did need quite a high level of expertise to be able to combine data sets and make them really useful.

I think it came out during the course of the conversations I had that this problem might be a function of a bit of a difference in the definition of "data set" amongst data scientists and data users and government. I think there's a bit more conversation needed around the definition of a "data set".

● (0850)

Another problem related to quality and the nature of the data is the format of the data sets on the portal.

In the past there have been some inconsistencies in the format of many of the data sets. I know that's an issue that the Treasury Board Secretariat has been working on. In the process of developing standards we really need to make sure that good metadata is included with the data sets. Missing and inconsistent metadata makes analysis really difficult; it makes it difficult for data users. The impression that I had from some of the users was that the standards for formatting are set a bit on the lower side, and that some of the metadata from certain data sets had potentially been removed in the name of standardization and consistency.

That brings me to my fifth point, which is the data portal itself. A lot of the people I spoke with had significant concerns about the data portal. I just came back from Open Government Partnership meetings in Dublin—they were European regional meetings—and heard many of the same concerns coming from civil society actors and assessors of action plans, coming from other countries that either have a data portal or are considering starting up a data portal. Data.gc.ca, as you know, is managed out of TBS, the Treasury Board Secretariat, which has the responsibility for the open government file. That centralization of the portal means that the data on the portal is effectively removed from its creators and its curators. It's removed from those, then, who have the highest degrees of specialization and understanding of the data itself. That puts TBS in the perhaps unenviable position of being a middleman, managing relationships and queries between those who are using the data and those who collected the data.

Some thinking needs to go into that issue, and perhaps the location of the data portal should be thought about. Some people I spoke with indicated that NRCan would perhaps seem a more logical home for the data portal, given that the majority of the data sets do belong to them and they have a high degree of expertise in data collection, presentation, and analysis.

Another issue with the portal is the search function. Users did quite widely indicate that it's not particularly user-friendly or well-designed, and they really thought that, at a minimum, with the portal, the functionality of the search function should be improved.

My second last point is that there is a growing data divide that's being created right now. Releasing data sets alone really doesn't have that much potential. It's not going to lead to any kind of significant change. You need people who can take the data and use the data. That requires expertise; it also requires resources. The raw format that the data sets are released in really does privilege data scientists, people who have high degrees of expertise in the use of raw data. Many others, non-governmental organizations, for example, would benefit greatly from the data sets and the information, but they're not able to use them because they lack the resources and they lack the expertise. If we're, in Canada, widely acknowledging that open data is important, then we need to think about potentially developing a mechanism for addressing that data divide and making sure that the data is accessible to a wider range of people than just people with a high degree of expertise—data scientists.

The final point that I'll make today is that open data is not open government. There has been a lot going on with open data, including the important study that you are undertaking with this particular committee. It's where a lot of other governments as well have placed their energy. We're certainly not alone in Canada in focusing on open data.

While there is certainly room for improvement, we have done some good things when it comes to open data. To be focused and careful of time, I didn't necessarily go over all of those good things. I'm happy to talk to them during the questions. My worry, after talking to a range of stakeholders, and conducting the Canadian evaluation of our open government progress, is that open data is becoming privileged at the expense of other areas of open government and some of the other commitments that we have made in our OGP action plan to the international community and to Canadians.

● (0855)

I'll close there. As I said, I'm happy to answer questions. I've provided the clerk with the link to a full copy of the report, and I can provide any other research that you might find useful.

Thank you very much.

[*Translation*]

The Chair: Thank you for that presentation, Ms. Francoli.

Without further ado, I give the floor to Mr. Deslauriers, from the City of Quebec, who has a maximum of 10 minutes for his presentation.

Mr. Deslauriers, thank you for being here. You have the floor.

Mr. Denis Deslauriers (Director of the Information Technology and Telecommunications Service, City of Quebec, As an Individual): Thank you, Mr. Chairman.

Good morning, everyone.

I am both pleased and honoured to be here today in order to share our experience with open data, in the hope of helping to improve the Canadian government operations with respect to data belonging to its citizens.

Like the Canadian government, we began in 2011 to release data, in response to requests from citizens.

For your information, the City of Quebec has a population of 575,000 people. As compared to Canada's population, which is about 35 million, that represents a ratio of 1 to 60. The City of Quebec currently has 65 data sets, as compared to 11,000 for the Canadian government, which is a ration of 1 to 180. The City of Quebec currently has 5,000 downloads per month on its website, where has the Canadian government has had approximately 15,000 per month over the last few months, including 5,000 during last February's EDOC, which is a ratio of 1 to 3. Since our launch in 2011, citizens have downloaded something 118,000 times on the City of Quebec's website. According to what I read this morning on your site, the Canadian government tallied 100,000. These numbers lead us to the question of whether volume necessarily equals usefulness.

Over the last few years, the City of Quebec has participated in several community events, events organized by the community rather than the city. For example, programming marathons that last between one and three days. We have also asked colleges and universities students to create, using our data, applications that truly benefit citizens, rather than creating bogus projects at the end of the school year that will be seen only by their professors.

We became aware that the one-to-three-day traditional programming marathons were much too short to allow the creation of quality applications, even if we provided funding.

We also realized that final projects submitted by students provided much better quality applications. However, at the end of the school year, students go off to work and forget about their applications. What is more, professors cleared the servers to prepare for the next session, and applications disappear, no matter good they are.

Among the problems we can raise, transparency and usefulness seem the most important to me. Take for example police data collected by each city in Canada, data which I have consulted myself. I cannot help but wonder how useful it is for a citizen to find out how many officers are working or the numbers of officers eligible for retirement. Although it is made accessible for reasons of transparency, does this data truly improve a citizen's daily life? I truly wonder.

As to statistics per province on break-ins, attempted murders, drug trafficking, gambling and so forth, they may well interest statisticians, but do they improve taxpayers' daily lives? I find all this a bit much.

Our mayor recently announced he would be making crime data available and break it down by neighbourhood. This will certainly have an impact on property values in neighbourhoods with higher crime rates. Furthermore, the residents' reaction will be to ask the city to double, triple or even quadruple patrols. In one hour, rather than having one pair of eyes watching, we will have four. Is that sufficient? Forgive me for my skepticism.

I would prefer the residents become more aware and stop asking themselves what government can do for them, thinking that government is responsible for surveillance and security. No, that should not be the case. As a citizen, one should be responsible for reporting events that seem suspicious. Thus, instead of having four pairs of eyes per hour, there would be 100, 1,000 or 10,000. If we raise citizens' awareness about the quality of life in their own neighbourhoods, we will improve safety in those neighbourhoods and bring property values back up. Generally speaking, citizens are unconcerned by their provincial or municipal crime rates. They are mainly interested by what is going on around them, in their own backyard.

I believe there are changes to be made, and I have other suggestions to that end.

What would you say if we provided real data on waiting times at border crossings, so that people could head to the right stations? What if we provided information about wait times in hospitals, medical clinics, and all the places where permits are issued? So that people can choose the least busy ones for themselves? That would certainly help mitigate problems.

What would you say if we provided information as simple as georeferenced data on our parks, fountains and public washrooms? These are things people regularly complain about, as they do not know where fountains and public washrooms are located.

Imagine the consequences if we provided the rate of success for kidney surgeries by hospital. Of the four hospitals in my own city, one has a success rate of 62%, and the others, 80%. What would people do? They would choose one of the hospitals with an 80% score. This data would allow us to improve as communities. We could try to find a solution to any given hospital problems, by checking whether the hospital received proper funding over the last few years, whether its staff is competent or has received the necessary training. It might be appropriate to reduce investments in hospitals with a success rate of 80%, to spend two years attempting to correct the situation in the one with the 60% score, and thus, rebalance supply. In my opinion citizens would be better off.

● (0900)

As for developers, I think we could make their lives simpler. We have already begun to do so this year by adopting the Creative Commons International licence for the cities of Gatineau, Montreal, Quebec and Sherbrooke, as well as for the Quebec provincial government. Developers no longer have to worry about which licence must be taken into account when using this data.

We have also begun standardizing data. For the first time, datasets for events and new ideas have been standardized among cities. We now host this dataset on our website and are uniting portals into a single stop for these four cities and the Quebec provincial government, in order to facilitate the use of the website for citizens.

We must not shy away from encouraging their use. We have kept citizens in the dark for a very long time. We felt we were in a better position to know what was in their interest. It is time to let them in. To that end, we have to show them how to go about it, as they have lost the habit. We must encourage use.

We recently held a municipal programming marathon. These four cities each contributed a $5,000 prize. Then, during one week, we asked the winning teams to polish their products in preparation for a provincial competition. We provided four organizations with $12,500 in funding for that purpose. It cost my city $8,000 for an application that is compatible with iPhones, Windows Phone and Android, as well as online in an adapted format. That application will be maintained for one year. That was one of the requirements to be able to win the contest.

We can also try to create wealth. Universities have begun to design products with us, for example, an application that helps you find a parking spot downtown during events. We encouraged them to partner with a not-for-profit organization, or NFPO. They continued to improve the application to make it possible, for example, to prolong the parking permit through a smartphone, without having to return to the parking meter, or, getting a discount at a nearby restaurant before attending a show. We are creating wealth at the same time.

We are supporting two CEGEP initiatives that will create a not-for-profit so that applications developed by students at the end of the year can be used permanently by future student cohorts. If these applications become too cumbersome to maintain, CEGEPs may ask the students to take them over and set up a business. That would create jobs.

We also need to think about facilitating data consultation by citizens. We have a great deal of data, but could we provide citizens with the required tools to view them easily and automatically generate graphs in columns and pie charts, rather than providing them with raw data? Ultimately, these measures are not focused on the developer, but rather the ordinary citizen. We could provide citizens with mapping that would allow them to view data on their area and benefit from it. Imagine the possibilities.

We could also invest in playing a role as liaison and catalyst for the community. Some software engineering students are quite good at developing applications, but not great at designing. In an another faculty, students who specialize in multimedia design are very good at designing, but have never developed applications. All these students are asking to work with us. We play a middleman role with citizens and try to maximize their contributions. According to the testimony we have heard, this is greatly appreciated. These students offer real benefits to citizens. They are happy to return the favour, because they know that their tuition fees are largely subsidized by citizens.

I will be pleased to answer your questions.

Thank you.

● (0905)

The Chair: Thank you for your presentation.

I will now give the floor to Mr. Hollett, who is the Assistant Deputy Minister for the Department of Finance for the Government of Newfoundland and Labrador.

Thank you for being here this morning. You have at most 10 minutes.

[*English*]

Mr. Alton Hollett (Assistant Deputy Minister, Economics and Statistics Branch, Department of Finance, Government of Newfoundland and Labrador): Thank you very much.

I'm very happy to be here this morning. I'm always interested and happy to talk to people about things that we're interested in as well.

I have brought along some slides, but unfortunately there wasn't time to have them translated into French, so you're not going to get the benefit of pictures that would probably make some of the things I'm going to say a little clearer. I'm going to have to take a slightly different approach to what I say because of that, but that's fine.

I looked at the questions of interest to the committee, and I thought the best way I could respond is to talk about where our government is going. We've recently announced an open government initiative, of which the open data is one component. I'll talk about that and ultimately tell you how we're dealing with that and where we're going. I will also go through some work that we've done in the past with data sharing, which will tell you why we're taking the approach we're taking. With all the things that I'm hearing, which Mary and others are saying, we've had the same experiences.

I will do that. I'll talk about our experience. I'm going to talk about the Newfoundland and Labrador community accounts data sharing initiative, from which we've learned an awful lot. It is the foundation for the way that we think about open data. I'm going to talk a bit about meeting user needs, and of course I'm going to end up with a little about where we're actually going.

We've had a long tradition of data sharing, and also supporting the users of our data. The Newfoundland and Labrador Statistics Agency has always been very interested in providing data to people in our 400 communities, scattered around 6,000 miles of coastline. We've given open access to a very wide range of information. We bought the data from Statistics Canada. We've developed it from internal sources, and we just put it out there. Statistics Canada has said that it's a peculiar thing to do; we're buying data and giving it away. However, we've always thought it was very important to do that.

To us, the open data initiative that we see across North America is essentially a focus on things we have always done. We're happy to see that focus. We're very engaged in the idea of open data, and very committed to it. However, it's not something that's absolutely new to us, by any stretch of the imagination.

Regarding the system of community accounts, on my slide I call it the “flagship” of Newfoundland data sharing, and it certainly is. We released it to the public in 2000. It has data for 400 communities, 200 neighbourhoods in our larger communities. It's actually a fully developed data set, in the sense that everything is documented; you can get back to the source. We have applications there, mapping, and so on.

The other thing we've done.... Dr. Doug May of Memorial University and I have partnered in this, and we've been at it for many, many years. The way we've packaged our data in the system of community accounts is that we use a well-being framework. My slides will be available, I think, and you'll be able to see it. I have a schematic there that shows an overview of that well-being framework. The reason we did that is that we wanted to make the data meaningful to people. The idea of the well-being framework is that we present data that gives statistics and measurements of factors that contribute to well-being in people's lives.

When you look at this at a community level, it's very powerful. People very quickly become experts because they know their communities. If you give them a number, all of a sudden it starts putting a quantitative dimension to basically knowing themselves. We have found that to be very effective.

In working with people at the OECD and the Australian Bureau of Statistics, and places like that, we found that we're probably 10 years ahead of things because that's where things seem to be going right now. That was very gratifying. In the beginning I was afraid we were going up the wrong alley, but we weren't.

We found it very useful. When you use that framework, you look at income, employment, unemployment, demographics, and so on. It helps to give you a sense of what data you should put in your system. It also helps with your prioritizing. You get people coming and they're asking, when are we going to have this, and when are we going to have that? We found it to be very effective, and we found that our communities and our neighbourhood people really liked that approach. A lot of people just don't know what the possibilities are.

In terms of lessons learned, which is a driving force in terms of what we're doing with open data, we found that people came to us and said, “You're a statistics agency and we'd like some data”. We asked what they wanted and they asked what we had. That's a hard question for a statistics agency to answer. It's very hard, as you can imagine. You can think of it as a warehouse that's full of all kinds of wonderful data. Most people don't know what they want, and a lot of people don't know the possibilities. This is the power of the conceptual framework that we've put in there with well-defined objectives, and so on and so forth.

● (0910)

Our experience with open data versus more developed data sets is that the majority of users really are not coming to us looking for the open data type of data. We find that most people, as Mary said, who use these data, who are looking for these data, are academics or seasoned data users, and quite often it requires a lot of work to actually use them. Of course, we've always provided those kinds of data when people ask for them, and I wouldn't want to give the impression that we don't think that because people are not asking for the data right now that good open data initiatives that are well delivered and well structured can't develop an appetite and develop a lot of interest and a lot of usage of those types of data. But I think we have to be realistic about where we are today. The market for raw open data, if you want to think of it that way, at this point in time, is certainly not very well developed, and if it is, it's clustered in specific places.

We view the open data approach as really most simplistic—and I don't mean that in a negative way. It's pretty elemental, how the concept defines open data, and then, of course, there's the value-added data, which is the community accounts type of data, in terms of my example. We look at data in terms of a sort of spectrum. There's data, information, and knowledge. From our perspective, the open data, the raw data, would be sort of just the data end of it, but when you do things with that data to make it more useful, you turn data into information, and when more fully developed, you begin to turn it into knowledge.

We've always put a lot of emphasis on trying to provide information and knowledge data. I do believe that in the future, when I get into a visioning mood, I really think we'll put a lot of effort into open data. We'll learn a lot about it, and eventually end up coming back to data that are better supported, better defined, and not simply dumped out of administrative data sets because they were never designed for those kinds of reasons. It will go full cycle. The market for open raw data will be there, and probably bigger than it is today, but I think most of the demand will not be there over the medium to long term.

In my slides, which you won't see.... I had a couple of slides there that I refer to as repairing data usage. The behind-the-scenes challenge is a messy business, and it really is. I encourage you to take a look at the slides when Marc puts them up, because I'm not going to get into it now. It is by no means simple or straightforward to take a set of data that people would consider raw data, and even to do marginal work to turn it into something that's going to be useful to pretty much any user. It doesn't matter how technically strong and numerically literate an academic person or any seasoned data user is. Administrative data files are nightmares to deal with, and that's where a lot of the open data interest actually lies.

These data sets, to be useful, require a lot of support. This is one of the main reasons why our government has had us into the data side of this, as a professional and well-developed statistics agency, to make sure. We want to be a leader in our province in providing a good data product. We don't want to get out there and just churn it out and have all our staff on the phone all the time trying to answer questions as to what this is.

We want to make sure that.... The value-added will vary across the spectrum, but the value needs to be there if this is going to be successful. I would argue, based on experience, that if we don't put effort into making the data sets clean, even the rawest form, if we don't make them clean and well-defined so they can be used properly and efficiently, we are creating a resource nightmare for our organizations in trying to deal with people who are going to be coming looking for help, looking for how to interpret, how to use—where do they come from, what do they mean, what can you do with them? Ultimately, I think this could be the foundation for the failure of open data initiatives, which I think are a very good way for governments to go.

● (0915)

In terms of what data can be shared, what we find at this particular point in time is that it's really a challenge to know which way people are going and which way people actually want to go. For many of the sites we look at, there is no obvious organizational framework. You see that the offerings are all over the place when you look across the different sites that are out there. The word I have on my slide is "spurious", and in many cases the quality is questionable.

But as for the way we think of it, we think of data as answers to questions, so where we start.... We've been going with the open data, and we've been encouraging our stakeholders across government to do so, the people who are into open information but don't really understand the open data as well as we do because we spend our lives at it.

First, we have to decide what questions we actually want to answer. Once we know what kinds of questions we want to answer, that begins to give us some idea of what the objectives of the initiatives are going to be. Who is the target audience? Are they highly skilled? Are they less skilled? Do we know what they want? Then, based on all of that, what's the best way to provide it across the spectrum? That's from raw data to knowledge, if you want to think of it that way.

As for what we've done in the approach we've taken, our government is fully committed to open government, to open data. There's absolutely no question about that. What we've done is establish a preliminary website. It's almost a demonstration website, but it's not something that will be withdrawn. It's something that will be made bigger. There, our Office of Public Engagement is beginning to consult.

I'm finishing now, Mr. Chair, because I'm sure I must be close to 10 minutes.

They're doing a consultation to see if we can engage with people and see where their interests might lie. A big thing we're doing that's going to be very useful for a variety of reasons is that we're actually building an inventory of all data sets across government. That is not simple. It's a big job, but we do have it under way. Of course, we're making sure as we go that privacy, confidentiality, and all that sort of thing is appropriately dealt with.

Based on our consultations, and also on our judgment, because I sort of feel that we're not going to get an awful lot of feedback from our consultations based on experience.... As I said earlier, you ask them what they want, and they don't really know for sure—

[*Translation*]

The Chair: I will have to ask you to please conclude.

[*English*]

Mr. Alton Hollett: Yes, absolutely.

That's the essence of what I have to say or what I can say. I'm happy to answer questions.

The Chair: Thank you very much for your presentation.

[*Translation*]

It's now the turn of the City of Montreal officials, who should be in a great mood this morning. We will hear from Mr. Chitilian, the Vice-Chair of the Executive Committee, and also from Mr. Fortin.

On behalf of the 10 members of the Standing Committee on Government Operations and Estimates, I would like to thank you for being at our meeting this morning.

You have the floor. You have 10 minutes.

● (0920)

Mr. Harout Chitilian (Vice-chair of the executive committee responsible for administrative reform, youth, smart city initiatives and information technology, City of Montreal): Thank you very much.

Committee members, Mr. Chairman, it is a great honour for us to present the point of view of the City of Montreal on open data.

With me is a senior official from our public service, who will answer questions later on.

[*English*]

Let me start by giving a quick history of the open data policy and open data initiatives at the City of Montreal. We started in 2011 and put in place, first and foremost, an open data policy, and then we went ahead to develop a website to free up the data to the public.

We started gauging the interest of the public in the data we started liberating and freeing up. We noticed that in the early goings, the initiative of the City of Montreal was perceived as a compliancy issue. Everybody was doing it, so we also had to do it, especially because there was a lot of pressure from the bottom up. Now, almost three years after the development of this policy, we recognize that it's not a compliancy issue anymore; it's a management-transparency issue. In other words, the city has to be transparent towards the citizens; therefore, we will and we do free up data that demonstrates how we use public funds.

Also, we will use the open data policy in order to develop solutions around three key issues we face on a daily basis in the urban environment. The first and foremost issue is transport. All of the future transport systems we are developing have built-in open data mechanisms and they will also have mechanisms for crowdsourced data, equally. It's one thing to have the public administration data, but where you get the synergy is if you join this data with that which comes from the citizens themselves.

The second issue we will concentrate on is sustainable development. There again, we have a lot of data, but we need to share that data—to help university students and help companies add value to that data; to develop policies and solutions to tackle most of the issues we will face in the future.

Last but not least is emergency services. Again, we have some data, but we need data provided to us by the citizens in order to build lasting solutions for this issue. How do we achieve that? On our end, we will tackle the three Ps that I always repeat to our people, to our citizens. First, we are always adapting our policy. Therefore, just like the City of Quebec and the Government of Quebec, we went with the Creative Commons 4.0 licensing. Second, we are reviewing all the processes of the city; therefore, existing systems and new systems have to have built-in open data mechanisms. They have to produce open data.

Finally, the greatest challenge of any public administration is to change the culture of its people. The data we have in our organization belongs to the people who are the different public servants in different services. Therefore, the greatest challenge we have is bringing all these people up to par with our policy decisions because a lot of the services are still very hesitant to free up their data.

On this, I will let Mr. Fortin, follow up with four specific areas where he's going to develop this idea.

Thank you.

[*Translation*]

Mr. Jean-Pierre Fortin (IT Strategy and Planning Advisor, Information Technology Service, City of Montreal): Thank you, Mr. Chitilian.

I will continue by speaking to the points of interest which were submitted to us when we were invited to appear. In the invitation, it was suggested that we talk about the needs of users and about socio-economic benefits, if there were any, as well as the ways different governments could collaborate. It was also suggested that we talk about best practices, and of the ones we felt were better than others.

But before doing so, I think it is important to specify certain statistics.

Up until now, the City of Montreal has released 107 datasets. They deal with subjects of interest to citizens, which Mr. Chitilian mentioned, including transportation, administrative data and services close to the people, such as sports, recreation, culture and so on.

As far as the needs of users and citizens are concerned, I believe that others before us have already said this, but we really have to insist on the fact that beyond the accessibility of data, citizens are asking for information and structured data so they can improve the way they use the services the city has to offer and the way they can access these services. Structuring the data of course depends on the availability of platforms. In this situation, we are not talking about platforms which would only receive data, but really about systems which make the way these data can be used understandable. Taken to its logical conclusion, we could even draw a parallel with environments involving business intelligence. Otherwise, these data would really be of interest to no one.

I also think that citizens need to feel that their city is transparent, and therefore their government as well. In return, we can hope that the public's cynicism towards its institutions would go down. So what we are talking about is creating and maintaining a relationship of trust.

As for the socio-economic benefits, I have just mentioned the first one, namely the feeling of belonging and of pride people have when they contribute to a more open society, one which is more dynamic and which makes sense. Another effect is that this creates bottom up work and initiatives, that is, initiatives which create value based on this data. For example, citizens could take the initiative and create applications for their fellow citizens.

If you do your job well and if you like your community, you can expect that a virtuous circle will develop. It would be a kind of ecosystem which includes a city that has data and makes it available, that includes supporters or creators of solutions who use these data, as well as informed and engaged citizens. So you would find yourself in what could be called a virtuous circle.

Regarding best practices, at the City of Montreal, we have always found inspiration in Europe's best practices. This includes both top down initiatives, where governments strongly participate, and, to the contrary, initiatives which strongly call on community involvement.

We also are clear on the fact that the British government contributes not only because it publishes wide ranges of datasets, but also because of the open quality of this data. We were recently consulting an index on open data of various governments, and Great Britain ranked first precisely because of the fact that its data is so open.

● (0925)

Which leads me to talk about the choice of licence.

As with our friends from Quebec City, whom I would like to recognize, and officials from the Government of Quebec and those of the other Quebec towns, we believe that this is an extremely open licence which has very few restrictions. It's the Creative Commons 4.0 licence, whose only requirement has to do with attribution.

In our opinion, it is essential that governments which want to work together agree on a licence which is as open as possible. Otherwise, even though there might be common standards, if the licence does not allow for combining data in a very general way, the work will be in vain. This is why we are working very hard to get all of Canada's public organizations to adopt a licence which is as open as possible and which, of course, comes with the fewest restrictions.

In addition, as far as collaboration is concerned, there are licences and standards, but in this case, as in other countries, the process is moving forward by trial and error. Everybody wants to do their own thing. On the other hand, we are witnessing a form of industrialization of all our processes, and because there are so many platforms, it will probably not be necessary for everyone to develop and maintain their own. Perhaps we can think about sharing these platforms, which would be defined based on common criteria and interests. At the end of the day, we might have super platforms, within which all public organizations could deposit their data. The level of interpretation of these data would largely exceed the level of interpretation of each order of government.

For example, it might be interesting for a Canadian citizen to not only know the extent to which people engage in recreational activities or use public transit, but also, generally speaking, to have an idea about the way in which Canadians engage in recreational activities in their hometowns. For that type of information to be available, the data would obviously have to be combined and integrated into common platforms.

I will stop here. I am ready to take your questions.

● (0930)

The Chair: Thank you for sharing your expertise in this field with us.

We will now move on to questions from committee members. Since we are also talking with witnesses via videoconference, I would ask you to please indicate whom you are putting your questions to.

Mr. Ravignat, you have the floor for five minutes.

Mr. Mathieu Ravignat (Pontiac, NDP): Thank you, Mr. Chairman.

[*English*]

My first question would be for Madam Francoli.

You began your presentation by saying that open data is only one part of open government. This is a rather narrowcasted study but I would still like to ask you about it. Regarding the relationship with other segments of open government and open data, it's hard to view open data and only open data without talking more broadly about open government.

Do you have any thoughts with regard to other principles of open government and the relationship to open data?

Prof. Mary Francoli: That's a good question.

We've actually defined or structured our action plan around different aspects of open government in Canada in a really rational way. In Canada we talk about open government in terms of open information, open data, and open dialogue. The open information bit of things relates more to access to information and it relates more to what we often hear referred to as unstructured data. That is, files that public servants might have on their computers and information that's generated more in a documentary form and not just in a raw data set. The open information bit is around that. The open data, obviously, we know; we've talked about that. The open dialogue bit is around engaging citizens in an ongoing and meaningful way. That's really necessary to a good open data strategy as well. All of those things really do work hand in hand.

Mr. Mathieu Ravignat: Is there an open dialogue piece going on right now with regards to data in the government's initiative, or have they abandoned that dialogue piece?

Prof. Mary Francoli: No. I mean engagement in dialogue around the open data bit has actually been not too bad compared to some of the other areas of open government. There have been various consultations that have been held since we joined the Open Government Partnership. As part of our membership in that body, we had to have a consultation around the development of our first action plan. We had to have the consultation around the post-year-one self-assessment. We didn't do a good job at those consultations, and the government has acknowledged in its own self-assessment report that we didn't do a good job in citizen engagement around those.

On the open data side of things, it was actually a little bit better. The consultation process involved an online consultation, but it also involved a series of round table discussions where people were engaged on a face-to-face basis. So really, in terms of the various consultations we've done around different aspects of open government, open data is actually probably the strongest.

● (0935)

Mr. Mathieu Ravignat: Now at the same time you kind of have to find a way to build in continuous improvement with regard to the data sets and with regard to the what's available and the format they're in. At TBS I guess there are practices in place or structures in place that would allow that feedback mechanism to ensure that. Because the portal is rather young, and we'll give some benefit of the doubt to the government, which I'm not one to do often, but it's a rather young thing and it needs to improve, obviously, given your presentation and other witnesses we've heard. But in order to do that, you have to build in a kind of continuous improvement capacity. Is that currently going on?

Prof. Mary Francoli: I think that's something that could definitely be strengthened. I think you raise a really good point to say that the data portal is really young, and our commitments in general to various aspects of open government are pretty young. So really we're in the second year of our commitments under our national action plan. We're trying to develop a second one. It's been a learning curve, and there's work to do for sure.

I think part of that work is developing a more ongoing dialogue between a range of different users. I heard academics mentioned a few times this morning. I'm really happy to be part of that discussion. It's always nice to get out of your office, certainly, to make your work meaningful, and not just say it's on the syllabus 100 times a day. But it's not just academics; it's people maybe you wouldn't think of at first blush. A lot of non-governmental organizations really want to be engaged in this dialogue around open data and tell the government how they think the data portal can be improved. Data scientists want to be involved in the conversation on a more ongoing way to say, “These are the problems we're seeing realistically with being able to use the data sets”, so some of the problems that I spoke to directly during my opening remarks.

Even though we've done what I think is a fairly good job with citizen consultation and engagement in relation to our particular open data commitments, that kind of mechanism for an ongoing and sustained dialogue amongst different actors is something that certainly needs to be developed in a much better way, and I think the government has acknowledged that. So right now, if you go to the data portal, you'll see that they're running a consultation on how to do consultations. So they're kind of asking people, “How do you think we can do this? How can we develop a mechanism for engaging people, and who needs to be part of that discussion?”

We're not the only government struggling with that. I mean I certainly heard that from other national governments at the OGP meetings in Dublin last week.

[*Translation*]

The Chair: Thank you for your answers.

Mr. Trottier, you also have five minutes.

Mr. Bernard Trottier (Etobicoke—Lakeshore, CPC): Thank you, Mr. Chairman.

I would like to thank the witnesses for taking part in this morning's meeting.

Mr. Deslauriers

[*English*]

and also Mr. Hollett,

[*Translation*]

You spoke about the importance of data in specific communities. People want to have data and information about what is happening in their neighbourhood. That's very important. I know that it was said that a great deal of data in the Government of Canada's portal were geospatial data. I think that's very important.

As you mentioned, Mr. Deslauriers, people want to have information on the crime rate in their area; they are not interested in macrodata.

Are there any problems with the way geospatial data are provided in the Government of Canada's portal and in all of the other portals? Is there anything we can do to help people get information on what is happening in their own neighbourhood, be it about the crime rate, the environment, transportation or the challenges of daily life?

I would like Mr. Deslauriers to answer first and then Mr. Hollett.

● (0940)

Mr. Denis Deslauriers: I believe so. The closer you get to people's daily lives, the more they are interested. For example, during the month of August you could ask people who live in heat islands whether they would be interested in knowing how the heat spreads throughout their neighbourhood and what they could do to play down its effects. If you want to be an open government, it means that you have to be open to suggestions from citizens and to their involvement in the development and maintenance of any solutions.

So yes, I think that there is a way to do that. We already provide some information, but it's not complete. For instance, there is also data on heat which comes from the provincial or federal governments. It would be interesting to get this information at the same time in a simple manner.

When I looked at the Canadian site, I saw a TIFF map, but it was completely useless for me, since I am an ordinary citizen. Unless I am an expert, that is useless to me, because I cannot use the data. I don't have the ability to do that, even though I work in IT. You have to be able to easily use this data, to draw conclusions and to take measures at the community level with your neighbours and the people in your neighbourhood.

We need to find ways to help citizens easily consult this data and then act accordingly. The point is that people should be able to do something without asking government to fix the problem. Ideally, people should be able to do things on their own initiative which we would simply support. This would be much less onerous and much more sustainable in the long term.

[*English*]

Mr. Bernard Trottier: Mr. Hollett, maybe you could comment on that. You mentioned some of the things you're doing by community. I'm not sure how Newfoundland and Labrador defines a community. Is it latitude and longitude? There are certain challenges at the very micro-level to define what is the relative data for a specific location. I know all governments around the world must deal with this challenge. What's the approach that Newfoundland and Labrador has taken?

Mr. Alton Hollett: It's not so hard for us, because we have 600 communities spread around 6,000 miles of coastline, and almost every one of them is separated from every other community. So basically our approach to defining a community is: it's Fogo, it's Harbour Grace, it's whatever. That hasn't been a problem for us at all.

That's one of the reasons we did the neighbourhoods. It's easy to do the communities themselves, but then people are interested in the larger areas, such as, if you know Newfoundland—some people here certainly do—the larger communities of Corner Brook, St. John's, and Clarenville. So we broke those into neighbourhoods of 1,000 population, and we did that with residents of those neighbourhoods as well, by the way. We had consultations with them and discussions. They essentially designed the neighbourhoods, but then we superimposed Statistics Canada geography over those so we could standardize the approaches, and so on, that we made.

I would like to just respond a little bit to some of the points that you made.

The geospatial aspect of it is critical and at our statistics agency we have very strong geospatial capacity there. One of the reasons it's so important is that I find in Newfoundland and Labrador—and I would predict that there's not much difference elsewhere—that what you said is absolutely right. Most people don't care about the macrodata. Obviously governments care, because we have to do our best to manage our economies and we have to have those indicators. But what really means a lot to people in communities and neighbourhoods is to take a look at data about themselves. We found that to be very....

The way we've always looked at it is that when you think about economic or social development, putting data into that equation is a very powerful new thing to add there. But if you deliver the data at the right level and in the right forums—we call it accessible, which means you can understand it and it's easy to get, easy to manipulate, that sort of thing—you actually turn people, who may not have much in the way of quantitative background at all, into experts. If I tell somebody in Arnold's Cove what the unemployment rate is in Arnold's Cove, they have a number that begins to put a dimension around their community. But they're experts on that, because they know what's happening to the guy next door. They know who's going to Alberta, they know who's working in the fish plant, and so on and so forth.

We've taken that kind of an approach for getting data into people's hands at the micro-level. I really believe that the buzzwords that we hear about evidence-based decision-making, and all that sort of thing. I think that the secret is moving in there at the micro-level and making data available to people in a way that it means something to them, so that they can actually start thinking that way and understand an awful lot more, and for that matter understand what governments are dealing with, because we have to deal with the realities. Lots of times people don't have enough information to know what the realities are and the dimensions of that in the same way that we do.

● (0945)

The Chair: Thank you, Mr. Hollett. I'll stop you.

[*Translation*]

Mr. Trottier, your time is up.

Ms. Day, you have five minutes.

Mrs. Anne-Marie Day (Charlesbourg—Haute-Saint-Charles, NDP): Thank you, Mr. Chairman.

I would like to thank all of the witnesses for their participation. We always appreciate their testimony.

My first questions are for Ms. Francoli.

On May 10th last, I was reading an article in *La Presse*. Amongst other things, the piece was about the poor quality of the search engine on the donnees.gc.ca website, and about the lack of communication following feedback from users. The piece also criticized the lack of transparency which came as a result of the elimination of the long-form census, in addition to the fact that the short-form census had cost taxpayers another $22 million. All of this showed that there were serious problems.

You remember that the Government of Canada participated in an international initiative to get governments to give more power to their citizens, to promote this fact, to improve transparency, to fight against corruption and to take advantage of new technologies in order to strengthen governance. This is the government's responsibility, and not that of citizens.

In your view, how does the Government of Canada fare compared to other G8 members in the race towards transparent data and as far as its commitment is concerned?

[*English*]

Prof. Mary Francoli: I think our commitments aren't totally out of line with commitments made from other countries, if we look at the context of our national action plan. With the G-8, we were involved in the charter on open data, so we adopted an open data charter with the other G-8 leaders. The charter commits Canada and the other G-8 member countries to a set of norms and standards for the proactive release of more high-quality user-friendly data that's unrestricted in the way people can use it and reuse it. In that sense, it's talking about norms. It's talking about standards and potential for use and reuse.

In terms of my study, I found that across the board various people I interviewed were very worried about the types of data being released. I'm not sure if this is quite what you were getting at, but they were really worried that what we're doing at the federal level is making commitments internationally to transparency and to improving accountability, and those things are good commitments to make. There's potential for better public policy. There's potential for strengthened democracy. But what people are seeing or the way they perceive what's happening right now—and I'll try to say this as neutrally as possible so as to not say that they're right or wrong—you kind of have the rhetoric about transparency and accountability going on, while at the same time we're cutting the origin of the data, so that goes back again to the long-form census. They see the government as saying that they're going to release things, that they're going to try for greater transparency and more accountability, but in doing that, they're very selective about what's being released, because they're cutting the collection of certain types of data, a move perceived by some stakeholders as a bit of information control over what's going on or what's feeding into the portal.

People are very worried about the long-term impact of that. They're worried about the impact of that for transparency and for accountability. There seem to be two things fundamentally fighting with one another there, and they're worried about the impact of that for good policy in the future and for what we'll know about Canadian communities on the smaller scale.

[*Translation*]

The Chair: You have 20 seconds left.

Mrs. Anne-Marie Day: Do you have the impression that there have been improvements?

[*English*]

Prof. Mary Francoli: A lot of the data in the data portal is data that was already released to the public. We can think of this as potentially an improvement, although there's work to be done in terms of the actual portal itself. We're seeing the data we have, what we own in Canada, and how it can be structured in a way that makes it more accessible. There's room for improvement, but it's good that those conversations are happening.

● (0950)

The Chair: Thank you very much. I have to stop you.

[*Translation*]

Ms. Day, your time is up.

I will now give the floor to Mr. O'Connor for five minutes.

[*English*]

Hon. Gordon O'Connor (Carleton—Mississippi Mills, CPC): Ms. Francoli, I have just a quick question.

The Treasury Board claims they have more than 190,000 data sets published, and then I see city and province representatives come here with different numbers. Is there any sense between these numbers? Do they mean the same thing? Are they talking about the same quantities of information or not?

Prof. Mary Francoli: I can't speak for the cities. My own research has really focused on the federal level, so I'm not quite sure what the municipalities are defining as their data sets.

It would be logical that the federal government is the owner of quite larger amounts of data than the municipalities. I don't think that would be unusual. There have been some bizarre exercises in counting, if you want to call it that, where I think even at the federal level we've been trying to figure out how many data sets we have.

In my report—and I had provided a link to Marc-Olivier—if we look at various points of time, we can see Treasury Board Secretariat releasing different sets of numbers around the quantity of the data sets that we hold. I can't remember them off the top of my head, but it's up and then it's down, and then it's a little bit up and then it's down again. So I think there is a lot of work to be done around what a data set is. Part of that is figuring out the standards as well. I know that TBS has said that in trying to develop a set of standards, they've worked towards combining some of the data sets and that accounts for some of the fluctuation in the numbers.

Hon. Gordon O'Connor: Thanks.

Mr. Fortin, we've been briefed from a number of people, and basically the same message keeps coming out that experts are who you're dealing with. Experts go to all the databases. They withdraw the information, manipulate it, etc.

But I believe that our purpose should be the people at large, and I suspect that the people at large don't know about these portals. Certainly they may not know about the federal; maybe they know about yours. They may know, in your city, about transportation or things like that. But I think we have a real problem getting to the people, rather than to the experts.

I just wonder if I could have your opinion.

[*Translation*]

Mr. Jean-Pierre Fortin: From the outset, I would say that it is fairly predictable that federal government data are not necessarily directly useful. After all, people are closer to services provided locally, that is, by municipalities or cities. Recreational services, garbage collection and public transit, for example, are things for which cities are responsible. It is normal that, in their daily lives, citizens would rather get information on these types of things.

Further, I agree that, for data to be useful, they must be relevant and easily accessible to citizens. I say this while being aware that these data are extremely useful to administrators. It is important that the government have statistics, so that it can understand its own processes, and that it can understand how its various programs perform. This type of data is not of direct interest to citizens, but it is for the government.

If data is to be accessible and useful to citizens, it must be processed and it must be turned into information. This data must also be collected, used and applied within the framework of applications. These applications could be on a platform provided by government. This would allow for all kinds of questions and analyses, if need be, but also for applications in everyday life, as those mentioned a little earlier by Mr. Deslauriers. For example, these applications could relate to parking, that is, where parking spots are available in a neighbourhood. This is really something that's tangible. This information could be made accessible in real time.

So by using existing systems, we could extract data in real time and transfer them to applications which provide information to people who need it. For example, it could be information about parking, as in this situation, but it could also be information about registration for recreational activities. Any data contained in our systems could be used in applications provided by either a public organization, or by third parties supported by government.

With regard to the infamous programming marathons, Mr. Deslauriers did not talk about their futility, but the fact that their usefulness seems questionable. Indeed, they result in the development of applications which have a limited shelf life. But also, there is the issue of quality, since sometimes these applications are developed too quickly, and also the fact that once these applications have been created, people start to lose interest. So the data will have to be used in applications which will be of lasting usefulness. Of course, I am talking about those applications which will be deemed to be useful and effective.

● (0955)

The Chair: Thank you, Mr. Fortin.

Mr. Simms, you have five minutes.

[*English*]

Mr. Scott Simms (Bonavista—Gander—Grand Falls—Windsor, Lib.): Thank you, sir.

Mr. Hollett, I certainly appreciate where you're coming from in talking about communities and everything else, but even the smallest communities are trying to use this data for several purposes—obviously, for attracting industries, for abandoned plants, that sort of thing. But you said something that caught my attention earlier, which was that they were surprised you would buy this from Statistics Canada and then release it to the public for free.

How does that work? What exactly are you buying? What kind of contract do you have?

Mr. Alton Hollett: I was thinking specifically of the data that we have in the community accounts, and of course, there's a wide range of data there because there are a large number of domains—income, employment, demographics, and on and on. We buy any data that we don't have readily available. We have historically bought the data from Statistics Canada in a generally pretty raw form, and then we've prepared it for the community accounts. So it's basically any data that we needed for the community accounts that were required to respond to the framework that we've established.

Mr. Scott Simms: You mentioned the term “well-being framework”. Obviously this relates to what is user-friendly regarding open data for people to use. In many cases, I agree that there are tools set up, but people don't know they are there and therefore don't utilize them enough.

I think about certain industries and companies that want to set up in certain areas, accessing a concept that we talk about, which is a skill set inventory. There has always been the traditional circumstance in which people try to find work by going to a particular website to find data for open jobs. But how can commercial enterprises find out what skills are available in a certain area in terms of people working in that region? Much of this stuff just exists as raw data. Nobody really packages it as something accessible.

I guess I'm asking all three of you this question. What are some of the best practices by which municipalities or the province or the federal government are packaging this data in such a way as to allow people to use it to better their communities?

Mr. Alton Hollett: That's an interesting question. This was one thing that led us, when we were building the community accounts, to put in an organizational framework for the data. When you talk about people who are interested in labour market-related data, for instance, that was one area we specifically gave a lot of thought to.

It goes back to what I said earlier. If you ask people what they want, quite often they don't know. Economists have ways of thinking about this. We look at our labour market and our labour market participants in a certain way. A lot of this is quite simplistic, as you'd be aware: males, females, age, wage rates, and so on. There are many other categories.

We began to set the data up so that those kinds of data were readily available. Instead of leaving people saying that they want to know what the unemployment rate is but would like to know more about the labour market and skills and what type of people are in the community and so on, we set it up along the lines of how an economist would think of it, which is basically as a description of the way the world works. That was very helpful to people for understanding what kind of occupations people are in now in our communities.

There are a couple of other things that we're doing as a government. One of them is obviously.... With our tight labour markets in recent years, it's a big issue. What we've been doing is developing occupational profiles for people in different occupations. We're also now building occupational projection models, whereby we take the forecasts we do at our branch, look at labour demand, and then look at labour supply and at where the mismatches are and provide occupational profiles—for individuals, to say where the jobs might be; and of course for companies, to give them a better sense of where the labour markets might be tight.

On that particular subject, it's not difficult to provide general information so that people can understand what is happening in communities from a labour market perspective. But when you start moving forward and asking about companies, or when you dig into the industries and that sort of thing, it requires a variety of measures, really.

● (1000)

The Chair: Thank you. The time is up. I'm sorry. Maybe you will have a follow-up question later today.

Mr. Aspin, you have five minutes.

Mr. Jay Aspin (Nipissing—Timiskaming, CPC): Thanks, Chair, and welcome to our guests. Thank you for helping us with our study.

I'm going to begin with a question to Ms. Francoli. I'd be interested in getting your assessment of how Canada compares with other countries in terms of open data.

Prof. Mary Francoli: That's a bit of a tricky question, given my position as the independent researcher for the reporting mechanism that is part of the OGP. The independent reporting mechanism tries not to rank countries in terms of who is and who is not doing the best.

You can see some reports out there. There is one released in February of 2013 by Capgemini that you might want to look at. I can provide a link to the committee, if you're interested. It did that kind of comparative ranking. It ranked Canada as a trendsetter in the area of open data, but a trendsetter behind the U.K. and behind the U.S.

Their metrics, which I haven't had time to critically evaluate, put us behind those two countries because of such things as the lack of diversity or breadth in the data sets that are out there—that goes back to the point I made earlier about the dominance of geospatial data—and also the lack of a good forum for engagement and lack of a good mechanism for ongoing, sustained engagements. Those two things together set us back a little bit.

Mr. Jay Aspin: Thank you.

To our other three guests with the province and the municipalities, I'd like a comment from each one of you. Are the provincial and municipal governments collaborating to link the data available on their respective open data sites?

Perhaps we could start with you, Mr. Hollett.

Mr. Alton Hollett: I think the answer is that's not very well advanced in our province right now. I know that communities use our data an awful lot. I get contacts from mayors throughout our province all the time, quite often thanking us and saying that they wouldn't have too much if we weren't there with our data. But I think there's really a lot to be done in terms of collaboration between the municipalities and our group—for sure—which will be a data leader in Newfoundland and Labrador.

● (1005)

Mr. Jay Aspin: Thank you.

Mr. Deslauriers.

[*Translation*]

Mr. Denis Deslauriers: We are working closely with the provincial government. This initiative brings together many municipalities. In Quebec, the directors of IT services in cities of 100,000 people or more meet every six weeks for different projects. We have united our efforts to meet with the provincial government and to ask for its collaboration.

We therefore already have a common licence for the entire province as far as open data is concerned, as well as a first standardized dataset, which is about events and ideas for things to do. Further, we are in the process of building a common portal, which is managed by an NPO, in which we will deposit our respective data.

[*English*]

Mr. Jay Aspin: Thank you.

Our friend from Montreal....

[*Translation*]

Mr. Jean-Pierre Fortin: I would like to add something to what Mr. Deslauriers from Quebec City said.

There is also an initiative where Canadian cities collaborate. Our group of Quebec cities is also a member of the Municipal Information Systems Association of Canada. This association brings together the people responsible for computer services in Canadian cities, especially the big cities. Within the association, there is a specific interest group on open data. We are in the process of getting organized to at least establish basic standards and criteria with regard to data so that they can be used by everyone.

Over the last few weeks, one person in this group was chosen to work in close collaboration with the Chief Information Officer Branch of the Treasury Board Secretariat of Canada to help Canadian municipalities work together in the interest of creating a common space or, at the very least, of creating conditions conducive to everyone being able to use these data.

The Chair: Thank you, Mr. Aspin.

Mr. Ravignat, you have the floor for five minutes.

[*English*]

Mr. Mathieu Ravignat: I'd like to dig a little deeper on the lack of diversity in the data sets that are available on the portal, particularly the lack of useful social data, which no doubt—and I fully agree with the stakeholders and with you—would have been better if the long-form census hadn't been eliminated.

You said in your presentation, Madam Francoli, that there was a lack of data on aboriginal people and seniors. I'd like to know why you think that is.

Prof. Mary Francoli: I think the geospatial data was already out there. There was already a lot of it. What happened when the data.gc.ca portal went up is that the existing data sets got amalgamated and pushed into that common portal.

A lot of the data on some of the areas that you mentioned and that I referred to earlier are data sets that might be seen as a little more sensitive. They are data sets that often have personally identifying information. I think that issues of privacy and respecting the Privacy Act are some of the main concerns about making that data public. That data needs to be really cleaned to make sure that personally identifying information is not part of the data sets.

Mr. Mathieu Ravignat: Do you know of plans to clean that data and make it available?

Prof. Mary Francoli: That I don't know. Treasury Board did say, during the course of the interviews, that there are efforts to expand the data on the portal but I don't know exactly what kind of data.

Mr. Mathieu Ravignat: I want to come back to what I think was Mr. Aspin's question, a kind of comparative, on how the Obama administration talked about "open by default" government. This government hasn't gone there yet. In the case of open data, do you have a sense of what that might look like if we were to go to a kind of open by default policy?

● (1010)

Prof. Mary Francoli: I'm not sure the open by default policy necessarily will have the greatest effect on open data. The open data initiative is really pushing data sets out there so it would be like, okay, if you have new data sets that are developed, then that information will automatically be made public. But there are parameters around that as well, so again, relating to things like privacy and national security and making sure that you're not releasing data that has potential negative consequences.

I think perhaps a bigger impact of that open by default policy will be on more the unstructured information, so information that Canadians would typically get via the access to information system, access to information request. I think that's where you would probably see the primary difference. We do have, in our national action plan, a commitment to developing an open government directive, which, as I understand it, is leading us down a path of that open by default framework that's been adopted already in the United States. This was something that we had committed in our national action plan to having in place earlier this year, but it's one of the commitments we're actually a little bit behind on. We are ahead in some commitment areas and behind in some. That was one that we're behind in.

Mr. Mathieu Ravignat: With regard to the results of your research and recommendations going forward, has there been, between yourself and the government, an open exchange of ideas? Have your recommendations been well received and do you expect them to act on some of your recommendations?

Prof. Mary Francoli: I hope that they pay careful attention to the recommendations. They're not, specifically, my own feelings about open government, so they were largely based on what I heard from the majority of stakeholders and things that were supported by documentary analysis as well. I certainly hope they pay attention to that.

Mr. Mathieu Ravignat: Have you met with Treasury Board officials?

Prof. Mary Francoli: I have not met with anybody from Treasury Board.

Mr. Mathieu Ravignat: Have you requested to meet with them?

Prof. Mary Francoli: I sent a letter. The way the OGP IRM process works is that the national governments get to see a copy of your report prior to it going public. They don't get to veto anything in there but they do get to say, "Hey, there's a factual error or a problem here". So that process happened and I did get feedback from the Treasury Board Secretariat in response to which I wrote them a letter addressing each of the points that they had made and offering to meet with them to tell them a little bit more about what I had learned and what's going on. I haven't met with them since the report was published. They did schedule a meeting between me and two other people, which they cancelled the morning of and then I never heard anything again.

[*Translation*]

The Chair: Thank you, Ms. Francoli and Mr. Ravignat.

Mr. Adler, it's your turn now for five minutes.

[*English*]

Mr. Mark Adler (York Centre, CPC): Thank you, Chair.

Thank you, witnesses, for being here today.

I do want to just begin with Ms. Francoli. We're pretty much in the infancy of open data. Is there anybody, any jurisdiction out there, in your estimation, that so far is getting it just right and is sort of the gold standard at this point?

Prof. Mary Francoli: That's really difficult. Different jurisdictions have approached it in different ways and I've been looking primarily at the national levels. I think really interesting things are going on at the municipal and provincial levels but most of my research is focused on the national level. If I look around at what's been going on in other countries, I've certainly seen some other initiatives that were really interesting. The U.K., as one of the committee members noted, led the charge towards the open licence, so we're looking to other countries like the U.K. and the U.S. and modelling some of our initiatives on them. So in terms of best practices we can say that's happening. There's the open licence as an example, and the open by default from the U.S. where we're modelling our open government directive on that.

One thing the U.K. has also done that was really interesting is a big data audit of their data holdings, and in addition to publishing a range of data sets they've also published the list of unpublished data sets. They've been very transparent with society to say, look, this is what you have, this is what you don't have. It gives them a mechanism to refer to and to go through and say maybe we can work on releasing this and this in the future. So there's a very clear understanding of what's there.

● (1015)

Mr. Mark Adler: Okay.

Since you've been studying and comparing national jurisdictions, you're probably in a good position to answer this question. In terms of a lexicon, or what standard definitions we may have or what certain sets of data define themselves as, is it uniform what you've been finding across jurisdictions, or does it vary from jurisdiction to jurisdiction? Could that create a problem, if it does vary, and should we come up with a standard definition of what specific words mean and have that universally applied?

Prof. Mary Francoli: I have seen glossaries of what the different terminology means when it comes to open data. On the data.gc.ca website there is a list of things, if you want to look up the background of open data.

I think, though, there's a divide that's happened. I think the divide exists between government as well as those involved in public administration and those with expertise in the area of data use. Data scientists are probably the best example of that. Governments, for example, seize data sets in a bit of a smaller way than the data scientists who want to use the data. I think there's room for more ongoing consultation and more ongoing dialogue between those two groups in particular to say, okay, well, what is it you need; what makes a complete data set?

A complete data set isn't maybe just a range of data sets with the same information for different provinces; maybe it's one big data set with all the provincial data in it. That data set can be easily comparable, or combined with other data sets, so that it can be used well. I think that conversation needs to happen a little bit more.

But as you say, we're still in the infancy of this. There's certainly room for that conversation to continue.

Mr. Mark Adler: Yes.

You can't just take data, dump it out there, and call it open. Do you find that government has undertaken enough of a public awareness campaign that, first of all, this data is available, and second, what this data may mean and how it could be helpful to you, the citizenry —

Prof. Mary Francoli: No, I don't think—

Mr. Mark Adler: —and one that cuts across all jurisdictions?

Prof. Mary Francoli: Again, I can't speak for the provinces or the municipalities—

Mr. Mark Adler: It isn't what you've studied.

Prof. Mary Francoli: —but at the federal level, certainly we haven't had that great of a campaign.

The one instance that really pops into my mind where that might have happened in an interesting and good way is the CODE appathon that the Treasury Board Secretariat sponsored to really say to people, here are that data sets we have on our portal, take them, do something interesting with them, and we'll publicize that. That was an interesting initiative.

But there's no communication, really, around the fact that we're a member of the Open Government Partnership, let alone the various commitments we have.

Mr. Mark Adler: You feel it would be very helpful if that were something that we undertook.

Prof. Mary Francoli: I think so. People need to know, in the context of open government in general, and open data as well, as part of open government, that we're doing interesting things as a country. There's room for improvement, but that improvement won't happen unless there's dialogue between different actors and unless citizens are engaged in a meaningful and ongoing way. Certainly that requires some sort of information or public relations campaign so that people know about it in the first place.

Mr. Mark Adler: Great. Thank you.

[*Translation*]

The Chair: Thank you, Mr. Adler.

Ms. Day, you have five minutes.

Mrs. Anne-Marie Day: Thank you, Mr. Chairman.

My question is for Mr. Deslauriers.

I will use the expression "open data users". A data user can be a child, an adult citizen, an NGO, a government, an expert, a researcher or any other person on the site. Users use the site depending on their own research needs or individual needs. For example, if I was going on holiday, I would look for weather data. This is an area for which the federal government is responsible, but this information should also be on the portal.

In the case of Quebec City, public money is being used to pay for the development of these data sources. How does our city—I am also from Quebec City—evaluate the data site?

● (1020)

Mr. Denis Deslauriers: Could you be more specific? You want to know how much the site cost or how the site is perceived by citizens?

Mrs. Anne-Marie Day: No. I would like to know how you find out from citizens what they think of the value of the site.

Mr. Denis Deslauriers: As with other levels of government, we haven't really promoted the fact that people can use the data on the site. As it now stands, we are still dealing mostly with developers, so they can create applications which will improve the collective well-being.

A few moments ago, I talked about parking, traffic, recreational activities and how to find municipal equipment. These are things which are useful for citizens to know about as they go about their daily lives. However, they cannot transform the existing data in information which is useful to them. For that, we need developers. We are still at that stage.

Mrs. Anne-Marie Day: You referenced Quebec City.

Mr. Denis Deslauriers: We have to go further and we need to have tools which will allow everyday people, people who are not computer experts, to interpret the data, to draw conclusions and to share information with others. You know that everybody is plugged into a social network. This is what we saw in Quebec last year, during the student protests. People communicate with each other. How can we provide them with the right information so that they can talk to each other about real things, and not share information which might not be reflective of reality? That is the second part which we will work on with regard to the portals, in other words, to provide accurate information which is useful to ordinary citizens.

Mrs. Anne-Marie Day: My next question is for all the witnesses.

One of the objectives of our study is to give people access to open data which is useful and usable, and which would help to stimulate economic growth. Do the various portals reach that objective?

Mr. Deslauriers can answer, or maybe one of the representatives from the City of Montreal, or the representative for Newfoundland and Labrador. In fact, all three witnesses can respond.

Mr. Denis Deslauriers: Generating wealth is not easy. You have to encourage people to do so.

As a comparison, if I had put 25 smartphones in a room even before they had been put on the market, you would not have known what they were. Perhaps you would have looked at them, but you would not have known how to get them to work, nor what they might be useful for. You would not have used them.

It's more or less the same thing with open data. People don't really know how to use it. You have to get them to do so. We had to convince one CEGEP after another to develop applications. Today, they understand the usefulness of these applications and they appreciate them.

Citizens who use these applications are also happy. The applications do not yet officially generate wealth, but one way to make sure they are sustainable is to get young people to create businesses for that very purpose. These applications will allow people, for instance, to find a parking spot without risking a $45 parking ticket. People will gladly pay $1, $1.50 or $2 for an application from an application store. If there was a way to help people easily find out what is available in the area so they can have an even nicer evening without paying much more, by accessing an application store, that would be a way of creating wealth. We have to find ways to do this.

There are other data, for instance, on contaminated land on which there has been drilling activity. These lands are located in municipalities and provinces; the federal government also is responsible for contaminated sites. It is possible to transform this information into open data so that people who have to work on a given piece of land, regardless of whether the land belongs to a level of government or a business, can find out what others have already paid for the drilling.

Mrs. Anne-Marie Day: Let me stop you here.

It seems to me that in G8 member countries, data must be accessible, universal and free. In fact, these data have already been paid for by taxpayers. Did you say that you are making people pay for open data?

Mr. Denis Deslauriers: No. The data are free, but people will have to buy the application created by using the open data. The application, for instance, might be to find a parking spot without risking a parking ticket. This type of application could be sold in an application store which would have been developed by others, who would have worked on that. Those people would not be reselling data; they are selling a service. If you want to get to a certain place, an application could tell you where parking is available and where you could park on the street. The application could tell you, in real time, where available parking meters are located and whether there are parking spots left in major parking lots. So with the help of this application, you will be able to find out where you can park your car. People might be willing to pay $1 or $2 for this type of application. The city would not have developed the application, nor paid for it; this application would have been designed by the private sector.

The Chair: Thank you, Ms. Day. Your time is up.

Mr. Hillyer, it's your turn for five minutes.

[*English*]

Mr. Jim Hillyer (Lethbridge, CPC): Thank you.

Mr. Hollett, you said that Newfoundland has been working with open data since 2000.

Mr. Alton Hollett: I said we've been very active in sharing data since 2000. If I stretched the definition of open data, I could say that, but I'm not going to do that.

● (1025)

Mr. Jim Hillyer: At what point would you say the notion of open data even became part of the conversation in Newfoundland?

Mr. Alton Hollett: Last year.

Mr. Jim Hillyer: If you don't want to stretch the definition of open data, I don't expect you to call it something...but what were you doing that's changed now that open data is part of the conversation?

Mr. Alton Hollett: That's an interesting question.

In the past we'd have people come to us and say, “I've looked at the community accounts, and you have your table, so you must have an awful lot more data to have created those” and our answer was yes. If somebody wanted that and they asked for it, we would deal with the privacy and confidentiality issues and we would provide it to them on request.

What's changed is that with open data there is essentially a focus now on the more elemental forms of the data that before, for example, we might have put in a table form, or for that matter, the data could be in a lot of government administrative databases and that sort of thing. From our perspective, the big change now with open data is that governments have embraced the idea of providing that additional elemental form. In the past, as I mentioned earlier, a lot of what we provided had value added to make it easy to understand, easy to use, and easy to access.

With regard to business development and that sort of thing in particular—because I think that's probably where a lot of the action is going to be in terms of open data in our province, and I suspect, in lots of others—what we plan to do, as we consult, is to see what businesses think they need and what would be useful. That, in combination with our own thinking in terms of the data sets we may have in government that we can bring forward and that will be useful to people for that particular purpose, presents an opportunity now, with the open data commitment by government, to actually begin moving some of that data out there.

One point I haven't heard anybody mention here this morning, which I would like to just toss out there, is that if you really think of what a best practice would be in terms of a good, solid organization providing open data, one of the things would be rejigging how that organization does its business and builds its databases and sets up its databases, and so on and so forth, so that it gets the data out there efficiently and effectively.

For example, you would build your administrative databases in that manner, change your organizational process so that you could actually get access to those data efficiently, and then do at least the minimum that is adequate to make sure that raw data, when it's out there, even though it's raw, is still good quality. That's one thing I haven't heard anybody mention. I haven't heard that mentioned pretty much anywhere.

I don't know if you have, sir, but it really is something that needs to be thought of, because the resources required to provide open data are huge. Right now I would argue that the resources required to do it today are probably 50 times what they should be, just because you have the administrative databases that are just not designed to provide that kind of data.

Mr. Jim Hillyer: Mary, would you like to comment on his suggestion and his recent comment?

Prof. Mary Francoli: I agree. It takes time, effort, and resources to develop quality data that can be put out there. It's not just about taking data and putting it up on a website. There are a lot of other components to it. It needs to be good quality data that people can use.

Ideally, there should be not only data specialists but also, as I said during my comments, people who might have lower levels of expertise with raw data. That's a really difficult thing to address. There are different types of software programs that allow for data visualization, which people can plug raw data sets into and they might get something out of them in a much more meaningful way. It might even be just a matter of identifying those sorts of technologies and providing them on a data portal or linking to them on a data portal for people to use.

I'm kind of departing a little bit from what you said, but I think you made really very good points.

[*Translation*]

The Chair: Thank you.

Thank you, Mr. Hillyer.

To conclude, I will give the floor to Ms. Day for five minutes.

Mrs. Anne-Marie Day: Thank you, Mr. Chairman.

My questions will be mainly for Mr. Fortin, from the City of Montreal.

Mr. Fortin, data is collected, stored and transferred, but one often gets the impression, in the case of collecting and storing, that the data accumulates into something like a library with a million books. One needs the right tools to be able to read them. However, an average citizen is not usually capable of reading datasets.

I believe the City of Montreal has 106 datasets. How often are the new datasets made available? Which datasets are downloaded the most frequently? Does the City of Montreal expect to realize savings or make its cultural and creative services more popular by having an open data website?

● (1030)

Mr. Jean-Pierre Fortin: Your question contains many parts.

It has maybe been difficult to make data available because of the approach that has been used up until now, which has been more or less successful.

Currently, data is released on a somewhat discretionary basis. There has maybe also been a lack of understanding about ownership of the data. Those in charge of the various municipal services are automatically deemed the owners of this data and it is at their discretion that data is released or not. Therefore there is, generally speaking, resistance. There is also a lack of consistency from one service to the other in terms of which data is released.

Of course we are thinking about an approach that would involve a model for others to follow. Take, for example, the American model which, besides being open by default, also includes an action plan. Under the American model, each department, each agency must, within a specific period of time, release a specific number of datasets from the accumulated legacy data. Furthermore, the expression "by default" implies that data within future systems will have to be what is called open data.

In terms of the quality of the data, there have been many discussions on what is open data and how it should be defined. I think we should be referring back to the charter signed by the G8 countries, of which Canada is a member. There are certain characteristics that define open data. Open data must be complete, primary or raw, timely, accessible, machine-readable, non-discriminatory, commonly owned, free of licensing restrictions, permanent, and available at little or no cost. When those characteristics are met you have true open data. Those qualities should also be true of any government initiative on open data programs. It is not enough to provide raw data; that data has to meet the criteria that I have just listed for them to be truly open.

You asked if the city will realize savings. That is not the current purpose. The main concern or goal of the current administration is that of transparency, for all kinds of reasons, including recent events. The city must be very open and agree to releasing all data as well as information on its own administration.

The data that is most frequently consulted deals with daily life. I am referring here to city websites on sports, recreation, parks and cultural events. We recently organized a programming marathon with the cities of Quebec and Sherbrooke on the events occurring in our respective areas. That is the kind of data that is of particular interest to citizens and that is downloaded.

Other kinds of data that are used deal with transportation, traffic lights, bus schedules and other similar kinds of information.

Could you please remind me what your other questions were?

● (1035)

Mrs. Anne-Marie Day: I think my time is up. Thank you very much.

The Chair: Yes, Ms. Day, your time is up.

On that note, I would like to thank you all for coming to this morning's meeting. Ms. Francoli, Mr. Deslauriers, Mr. Hollett, and Mr. Fortin, live from Montreal, who was accompanied at the beginning of the meeting by Mr. Chitilian, thank you all for your expertise. Your testimony will certainly help us with our study, which is reaching its end.

I would like to remind committee members that we will meet again on Thursday at the same time. It will be our last meeting for hearing testimony on open data. After that, we will come back on Tuesday, May 27 and Thursday, May 29 for the purposes of other study topics.

Before adjourning, I would just like to confirm that the Clerk of the Privy Council wanted to appear on his public service report, but he was not available on May 27. He has asked if he can appear on June 5. If the committee agrees, we could modify our schedule. There appears to be agreement.

We will set aside some time at the end of our Thursday meeting in order to discuss witnesses for our next study. I know that the government party has already submitted a few names but if the other parties also want to submit names, they will have an opportunity to do that next Thursday.

The meeting is adjourned.