



HOUSE OF COMMONS
CHAMBRE DES COMMUNES
CANADA

Standing Committee on Government Operations and Estimates

OGGO • NUMBER 021 • 2nd SESSION • 41st PARLIAMENT

EVIDENCE

Tuesday, April 29, 2014

—
Chair

Mr. Pierre-Luc Dusseault

Standing Committee on Government Operations and Estimates

Tuesday, April 29, 2014

•(0845)

[Translation]

The Chair (Mr. Pierre-Luc Dusseault (Sherbrooke, NDP)): Order, please. Good morning, everyone. We are going to begin our 21st hearing.

We have several guests with us, and I thank them for being here.

Today we are continuing our study of the government's open data practices. We have the good fortune of having some representatives, who are mainly from Ontario. The first to make a presentation will be Mr. McKerlie, Deputy Minister, Open Government, Ministry of Government Services, Government of Ontario. He is accompanied by Mr. Foulon. Afterwards we will hear from a representative of the City of Ottawa, Mr. Giggey, as well as a representative from the City of Toronto, Mr. Low. We shall also be hearing from Mr. Lenihan, as an individual, who will also make a presentation. Afterwards, as is our usual practice, the members of the committee will have questions for the witnesses.

Thank you once again for being here.

We will begin immediately with Mr. McKerlie, from the Government of Ontario.

Thank you for having taken the trouble to travel here. I know that you have just gotten off the plane, so thank you again for being here with us.

Mr. Ron McKerlie (Deputy Minister, Open Government, Ministry of Government Services, Government of Ontario): Thank you very much.

[English]

Good morning. Thank you very much for the opportunity to be here today.

My name is Ron McKerlie. I'm the deputy minister of open government in the Ministry of Government Services in the Ontario public service, and with me today is Marc Foulon. Marc is the head of the open government project for the central agencies I and IT cluster.

Our open government office was created just 10 months ago, so we're fairly new to the business of open government and open data. It started in June 2013. At that time we started shaping the vision for open government in Ontario. We brought Marc on soon after to help lay the groundwork within our ministries and within the I and IT organization, particularly around the area of open data.

On October 21, 2013, Premier Kathleen Wynne publicly announced Ontario's commitment to open government. She penned a letter at that time to the public in which she wrote, "Our Open Government initiative will help create the transparent, accessible government the people of Ontario deserve."

It was at that time that she also committed to having each ministry craft an action plan on how they would pursue the open government agenda. She also introduced the open government engagement team chaired by Don Lenihan of the Public Policy Forum, and I'm delighted to see Don is here today as well.

The open government work that we're doing now falls into four broad streams. The first is open dialogue, which means increasing opportunities for the public to provide informed and meaningful input on legislation, on policies, and on programs that affect them. The second is open information, which means making government information normally only available through the freedom of information process available to the public on a proactive and ongoing basis. The third is open data, which means making government data available in a machine-readable format and covered by a common open licence so that people and businesses can access it, can utilize it, and can repurpose it to develop new ideas, new services, new applications, and hopefully jobs.

The fourth is accountability, which involves a number of legislative and non-legislative issues and items that would strengthen political accountability, enhance oversight, particularly with our agencies, boards, and commissions, and increase transparency across government and the broader public sector. The public sector and MPP accountability and transparency act, 2014, was brought before Ontario's legislature on March 24 and is currently in its second reading debate.

Of the four streams Ontario is pursuing, open data is the most mature. In November 2012, we launched our open data catalogue, which currently has over 175 data sets, such as energy consumption for the Ontario public service and the broader public sector, Ontario research funding details, greenhouse gases and pollution emissions, and Ontario library statistics, to name just a few. All are publicly accessible online.

One of the first goals of the open government office was to create an inventory of the data sets held by the Ontario public service, and working together with ministries, we have identified over 1,000 of these data sets. We have now posted them online and they're available with a voting tool for the people of Ontario. We've since shared our inventory collection process and guidelines with the provinces of B.C. and Alberta, and with numerous Ontario municipalities and regions. We don't have the resources to make all of our online data available in an open format immediately, but we also wanted to ensure that we pursue a quality over quantity approach, which is why we considered the voting tool and how we would release it to the public.

In the first 24 hours, over 10,000 votes were registered on different data sets, and we have more than doubled that in the first week that this voting tool has been available to the public. Ontario is the first government in Canada to encourage and empower the public in this way and there's been a great deal of interest in telling us which data sets are most valuable to them.

Another key piece of our work on open data has been to introduce a common licence on our data sets. Data released under a common licence becomes even more valuable because people can more easily use and combine data from any source that uses this licence. Ontario worked with the federal government, with Alberta, and with British Columbia to create the licence and we have adopted it for all published data sets from June 2013 onwards. Since then numerous municipalities across Canada have adopted this licence, including the City of Toronto, and we are continuing to encourage other Ontario municipalities and regions to use this licence as well.

● (0850)

Ontario has also been committed to building relationships with the developer and the research communities. We've hosted a number of events to get developers thinking about how they can use our data and how we can support them in doing that. Industry data owners have also attended these events in order to help developers understand the nuances of the data. Developers are one of the key audiences for open data. They're the ones who build new applications using the data, and we value their enthusiasm and support.

In late March, we received the open government engagement team report. The team spent the early winter travelling around the province, engaging the public, elected officials, public servants, and journalists on how they want to see government done differently. The report, entitled "Open by Default: A new way forward for Ontario", includes 17 recommendations, such as improving the freedom of information framework, launching a one-stop open government platform, and other things that would help consolidate information for all public engagement initiatives across the province. We have been reviewing the report, and we're preparing a response that will go forward shortly to government.

One of the recommendations they made was to implement an "open by default" data policy in the OPS, and the development of this open data policy is now under way.

The open by default concept has several tenets to it. The first is that we would publish all government data in commonly accepted open standards unless there are privacy, security, or legal reasons not

to do so. It would also mean publishing data in a timely manner, with the highest level of detail possible and in a machine-readable format. It would mean making data available free of charge in non-proprietary or commonly adopted formats. It would mean ensuring that no data is destroyed. It would mean waiving intellectual property for data that the government collects or creates and ensuring we don't transfer intellectual property of data to a third party. It would mean extending these open by default principles to agencies and to our broader public sector organizations when renewing existing governance agreements.

We're in the process of drafting an open data directive that would support this open by default approach in the public service and in our classified agencies. The directive will define key principles and requirements on publishing open data. It will define data that is exempt from disclosure due to privacy and confidentiality or security reasons, and it will promote a culture of openness and collaboration, both within the public service and externally, to the people of Ontario.

The key principles that Ontario will commit to under this directive are comprehensive. The first and primary principle of the directive is that government data that is not exempted by this directive will be made open by default, as understood by the tenets I previously spoke to.

The second is that government data must be accessible to the public via the Internet in a machine-readable format, at no cost and under an open licence with few restrictions. It will also direct data stewards to prioritize the release of high-valued data as defined by public demand and government priorities. We will commit, to the extent possible, to releasing data in its original unmodified form and at a level of granularity that will not compromise privacy and privilege, confidentiality, and security of government data.

Finally, we will apply detailed data quality guidelines to ensure that any government data released is complete, accessible, fully described, and timely.

This directive will provide a strong foundation for the Province of Ontario to build a policy process aimed at implementing a 21st century open government organization, and will be one of the most comprehensive data directives in place anywhere today.

Ontario is focused on strengthening partnerships with other levels of government. We are very motivated to participate in the delivery of a national data set search capability, for instance. We believe this will be of great value to public servants, the research community, and developers, as well as to the general public. A federated search functionally for Canadian governments, including federal, provincial, and municipal, would help the adoption and use of open data by improving access.

I also believe that we should be pursuing common standards and principles around open data. Interoperability in data standards and how we release it between the platforms we use is more than a technical exercise. It's about building value for users and ensuring that all levels of government work together to provide citizens with a seamless experience of government.

● (0855)

We also need to describe our data with comparable metadata standards so that policy-makers, researchers, and citizens can tell if a particular data set is relevant to their questions or if they are comparing apples to apples. We need to release our data in compatible file formats, under a compatible licence, and with a relatively similar user experience. Otherwise, only a few very technical individuals will have the time, skills, and resources to extract, convert, and even find the data they need to start a project or test an idea.

Common standards and interoperability are not a “nice to have” but a strategic requirement needed to do public policy and to tackle with citizens, based on evidence, the complex and very real policy problems we're facing.

This is why we're happy to be participating in the new open data Canadian leadership forum, and the open data and information working group. Working together on these common goals will open up more possibilities for both our governments and our citizens.

We're obviously very proud of the work that Ontario has accomplished in a very short time. We came to the game a little bit late but we've been persistent and dedicated in our efforts since then and I believe our achievements speak for themselves.

I also believe that there are still many very exciting potential opportunities for Canada to emerge as a leader in the open data space. By working together and alongside other levels of government, I believe we can create something that is greater than the sum of its parts. By working together, by developing an agreed-to set of standards and protocols, by making search easier, and by supporting the developer community, I believe we can become global leaders in open government and open data.

Thank you very much for your time and attention. Marc and I look forward to answering your questions.

[Translation]

The Chair: Thank you, Mr. McKerlie.

[English]

Thank you very much for your presentation

[Translation]

We are now going to continue with Mr. Giggey, for 10 minutes.

Thank you for your presence before the committee today. You have the floor.

[English]

Mr. Robert Giggey (Open Data Lead, City of Ottawa): Thank you very much.

Thank you for inviting the City of Ottawa and for inviting me to speak today. I've been running the open data program at the city for the last three to four years. To give you a bit of history, I've been working in the municipal IT sector for the last five or six years. I've served on several boards that have been working with open data, and we've been partnering together with other cities. Four of the larger cities, Toronto, Edmonton, Vancouver, and Ottawa, have been collaborating on open data for the last three years. As well, there is now a provincial open data group that's working together that I've been participating in. The municipal IT associations have created special interest groups around open data that I'm participating in as well.

Just to give some background, and the city plays host to you so you are aware, we are the fourth-largest city in Canada but largest by land mass for the big cities. We have 17,000-plus employees and a \$2.5-billion operating budget. That gives you a sense of comparison, I guess, with the other cities. We're trying to be representative of the cities and how we use open data, and maybe can reflect those.

For the city itself, we initiated open data through a review. We were requested to review our data dissemination policy. Before “open data” took hold as a common term, we were simply looking at mimicking what some of the private sector was doing. Could the city leverage the release of data to tap into primarily the local technology sector to see if it would support them in building solutions for the city, but also to encourage innovation based on what we were seeing from local developers?

We launched in 2010. Our launch was fairly quick and simple. We simply started to put up easy-to-release data sets. For a city, that's pools, park locations, and facilities, stuff that's very common sense to put up. We immediately launched an application development contest to make use of the data, to create awareness, to have people involved, and to start gaining some of the benefits of having the data available.

For the first few years, our focus was on community engagement. That's ongoing. We've continued to do that. It's a critical part. I think any jurisdiction that has gone forward with open data has seen the tremendous value of reaching out to those who use the data. Because the topic itself and the skills required to use the data aren't necessarily in the greater population, you want to reach out to those who can create stuff with it and help deliver the value to the residents or citizens.

Along those lines, we've hosted many events in the city—hackfest may be a term you've heard—where we bring members of the community together. We bring staff along, and we try to encourage the development of applications or services or research based on not only the needs of the city but also what the community groups are interested in. We're on hand to be the experts around the data itself, with the intent that we can focus on being good data providers, of running that and of having good-quality clean data and understanding the topics, but they can help deliver what's of value to the community based on their own needs, or reaching out whether from a profit motive or for community engagement or community development.

We have run several contests along the way that have been successful with Apps4Ottawa. We'll probably be looking to doing a third one in the not-too-distant future. These are useful activities in terms of not only gaining the attention of the public, who do benefit from the outcomes of it, but also reaching out to a broad set of users of the data, going beyond application developers into academics and researchers. It's getting students involved, students in universities and high schools, so that from an early age they can see that open data exists and there's value in it. It can help lead them to greater community engagement, get them involved with their cities and with government, and get them interested in doing their part to contribute to society.

● (0900)

We have had many successes along the way. One of the great things, as we've seen, is that when we launched the open data program we spoke of the benefits that we were hoping to see but we've actually now attained them so it's no longer just theoretical. We have proven cases where open data has shown its value manyfold over. For the cities, we are lucky in a way because we have so many front-line services to deliver and it's very amenable to, say, front-line applications that the public can use. We've seen tremendous benefit whether that's around recreation, transit, traffic, garbage, or recycling. The benefits we see are cost reduction from us not having to develop solutions but also driving new revenue from the city to getting more people to using our recreation programs, our cultural programs, and doing it without further investment from the city.

Finally, one of the most recent developments is a new routine disclosure and proactive dissemination policy that the city has adopted. This is in line with the direction that the Province of Ontario has moved in and aligns with the Privacy Commissioner of Ontario with their access by design. This allows the city, moving beyond open data—it can include open data but what information should the city be proactively disclosing for the same purposes?—to have the information available to the public, to the community

groups, to associations and researchers, so they can create the value from it as well as promote transparency and accountability.

I would like to use this time to suggest several recommendations from the point of view of a municipality working around open data of where we would like to see the federal government move or opportunities for them to adopt.

The first is around a common data and it's for the provinces as well. Looking at many of the topics, whether transportation, health, environment, education, governance, or spending, for people to have the complete picture they need the data from all three levels of government and I think the federal government is in a good position to help drive that forward and promote the release from the three levels of government.

The second is around federated data. There would be seemingly tremendous value in having a single place of access for data sets. So if we were all releasing transportation, environmental, or health data, if there were one place to access that data, it would make it more accessible to more users and the federal government's portal could serve as that place. So not only would cities release their data on their sites but they could also be releasing them through the federal portal to simply make it easier to access.

I think the federal government is in a good position to help promote standards for interoperability and policy. I think they've done that through the release of the common licence that they've been promoting which is, in itself, of tremendous value in enabling many people to participate and use open data. I think there's more that can be done in this area specifically around data standards but also in promoting formats that will support interoperability. This is so that data from all the many jurisdictions can be more easily used by those who choose to use it.

The final point is around close engagement with the municipalities. There is, obviously, a separation between the federal government, the provinces, and the cities. There are formal processes in place to help with that collaboration, such as the Public Sector CIO Council; however, it can be limiting in terms of getting that real collaboration between the three levels. I think, whether it's through the federal portal, through other means, if there's a way to connect, say, the bureaucrats at the federal level to the municipalities directly with the provinces involved, we'll have a better outcome for citizens and users of the data.

Thank you for your time. I look forward to your questions and helping you. Again, I appreciate your inviting us to speak and that you are looking at this topic specifically. Thank you very much.

● (0905)

[Translation]

The Chair: Thank you for your presentation and your recommendations.

Without further ado, we are going to hear from Mr. Low, who is the Manager of the Social Research Unit for the City of Toronto.

[English]

You have 10 minutes.

Mr. Harvey Low (Manager, Social Research Unit, Toronto Social Development, Finance and Administration Division, City of Toronto): Good morning, everyone, and thank you for having me.

If I could just ask you to talk to Environment Canada to change the weather out there, I'd appreciate that. That's my first ask.

I'd like to break down my comments into two main sections. First is a little bit about the history of open data and open government at the City of Toronto and some lessons learned that might help you. Second, I'd like to address the key questions that you've asked of me in terms of our response and input into the federal data portal.

Just briefly, my name is Harvey Low. I'm the research manager of the social research unit at the City of Toronto. In 2012, I was the OGP representative down in Brazil on behalf of Canadian civil society and Canadian municipalities. I am currently the chair of the city's research committee, and the human services cluster representative on our open government committee. The city has an open government committee, actually.

We have learned some lessons at the city. We've been in the open data game since about 2010. As Robert mentioned, I'm one of the original four. We have an open data site, and now the city has actually moved to multiple portals for access to free data. Essentially we have not just the open data site, which has about 110 data sets up there, if my count is correct, but also an application called Wellbeing Toronto. It's an online mapping application that provides an additional 200 to 300 variables, many of which are from federal and provincial data sets that have been long in coming. So I agree with Robert that if there is a way the federal government can help us stickhandle and manage access to both provincial and federal data, that would be wonderful.

I am also on the steering committee of the Canadian Council on Social Development's community data program. For those of you who are unaware of that program, that was formed about eight years ago, with 22 Canadian municipalities, to get easier and low-cost access to federal data. That's been going on, as I said, for eight to ten years.

One thing we've learned at the city is that it's not always wise to simply release data sets without context. I'll speak to that in a little bit more detail.

We want to make sure that a website that is open to the public also has a context explaining what all those data sets are about. Typically users out there are of a common profile for open data sites. They are usually more technically savvy, younger in age, and very competent in the area of social media.

The first question you asked us was how we compare to other municipalities and how we meet the needs of Canadians. You asked about how we become accountable and transparent while maintaining privacy. I think another question you asked was how we use the data for very specific issues such as job stimulation.

I'll try to address those now through a couple of key theme areas. The first one is to engage the unengaged. We need to decrease the digital divide. As I mentioned before, typically the users out there are people who are hackers and developers—and that is wonderful—and

they're very good at data manipulation. Ron mentioned that a lot of them know how to manoeuvre around data. A lot of them know how to analyze data, and they know how to change formats in data. The common public don't have those skill sets.

I think, in line with the provincial and local municipalities, the federal government also needs to begin to recognize that there are other users of your data out there. You need to engage the unengaged.

Who are those? Well, outside of the developers and hackers, there are cities and municipalities. We need data for very strong, place-based evidence planning. Everything that we do in municipalities is done by neighbourhoods, by place. I'll speak to the issue of geography in a moment. You need to reach out to non-government agencies, community agencies, and those that also use data and analyze data on behalf of the public and the clients they serve. The other area of engagement is of non-traditional users of open data sites: professional organizations. They are sitting on a landmine of information from professional geographers and urban planners. Those associations would recognize the value of consolidated federal data. Of course the last group is the people we serve. These are the at-risk population groups—seniors, new immigrants, youth. Those are the groups that also need the data.

So you need to begin thinking about it not being just about hackathons. While I think hackathons are great, you asked whether we are serving the needs of Canadians. What we've heard in the city is that they need an intermediary. They need an intermediary to take that data, such as that from municipalities or the province, and analyze it on their behalf.

The second thing I'd like to talk about is releasing the relevant. What do I mean by that? You asked how we compare to other jurisdictions. The real question is how we can be better than other jurisdictions.

● (0910)

I did a quick scan of the international federal data sites from the Netherlands to Australia. There is one thing in common with us here in Canada. I went in there pretending to be a common user from the public, and I wanted to type in an issue area. We talked about job stimulus. I typed in "unemployment" and I came up with—and I hope I'm not putting out anyone from Australia—zero on the Australian site, zero on the Netherlands' site, zero on the U.K. site. People search by issue areas. The search engines that we have out there...and I'm glad, Ron, that you mentioned we need a national search engine, because that's exactly what we need. We need a better way to tag data. Right now, you need to know what that data set is called before you search on it, which is a bit of a problem. The developers will figure it out, but the general public won't.

We operate right now in a local municipality in what we call a shared-service delivery model. We as municipalities work with communities, non-government agencies, in interpreting data on behalf of the public. Why do we need the data? We need data on vital socio-economic data right now about people's economic circumstances, education levels, those types of things. Those types of data sets, while they may be available at federal ministries like CIC and StatsCan, are not connected right now with your data portal. Raw data's in one place, and that's fine. That serves a specific audience, but we need to pay attention to that other audience.

The other thing we need to do is to link and leverage. You asked the question about how data can stimulate job growth, as an example. One way is to begin linking your data portal with the service provider areas in the various ministries across the federal government, and down through the provinces and municipalities. So if somebody goes in there looking for the economic situation of their municipality, their neighbourhood, or Canada, it would be good to have the data in your data portal linked with, say, economic development or the ministries, even if it's a link to their website. Then you're connecting and networking that raw data to a very specific mandate of a particular ministry. If I understand your standing committee, one of your mandates is cross-departmental connection of services for the public. If you just provide data on a raw data site, that will satisfy the hackers and developers, but it may not truly get at providing the services to the public you want as a federal government. That's what I think we're all here for, for all three levels of government.

The other thing we want to do is to make sure we tag the data by thematic areas. To get back to my question and issue about not having sufficient tags when I type in something as simple as "seniors" and "youth", I'll get summary papers, which is great. We need to provide better tagging and search capabilities on all our websites.

We need to begin thinking about geography. We need to give the gift of geography. We do everything at municipal levels by place. It's wonderful to have statistics at the city level, what StatsCan calls the CSD level, but that level of data we need to be more granular. We need it down at the neighbourhood level. There should be search engines out there that allow users to search by geography, as well as theme. When we look at the data we need, we're going to need unemployment rates, we're going to need health care statistics. Those have to be at a level of granularity that makes it useful for cities in their place-based evidence planning.

The question on privacy can also be solved through geography. We do not have to look at individual data, we aggregate data to a level of geography that allows us to do that aggregation. Right now, you could say we have the Stats Canada site, but as I said before, begin to link the data portal at the federal site with other ministries and their service mandates so when people look at a table they know where to go if they have questions about that data, and if there are interconnections.

One of the ideas is that you look at service delivery from a multiple factors analysis. We call it in social planning, social determinants of health model. That means if a person has an issue with unemployment, and it has to do with health, housing, transportation, all of these things need to be linked. I think there's

an opportunity to link your open data site with service providers in the federal ministries, down through the provinces and the City of Toronto.

● (0915)

Robert mentioned that—not the horizontal connection, which is one issue, but the vertical connections between all levels of government. This is because people understand geography. They don't understand jurisdictions. They don't care where the service comes from. They don't care where the data comes from.

Where do we go from here, as a conclusion? We need to be thinking about intergovernmental data sets, not just data out in Victoria or Halifax, or one ministry over in Alberta or a federal ministry. We need to begin thinking about what the public wants and creating data sets that represent intergovernmental service delivery to particular groups.

Thank you for the time, and I look forward to your questions.

● (0920)

[Translation]

The Chair: Thank you for your presentation.

Our last witness is Mr. Lenihan who is the Senior Associate of the Public Policy Forum.

[English]

Thank you for being here. You have 10 minutes.

Dr. Don Lenihan (Senior Associate, Public Policy Forum, As an Individual): Great. Thank you very much. *Merçi beaucoup.*

First of all, let me thank you for the opportunity to come to speak to you today about some of the work I have been involved in. I guess the reason I'm here, as Ron McKerlie has already said, is that I was the chair of the Ontario government's open government engagement team. We had nine members. We were all from outside government and were asked by the government to go out, travel around the province, and think about the future of open government.

I want to talk a tiny bit about our mandate. In fact, I want to stand back and try to say something from a much more big picture point of view because that's the kind of discussion that we had. We had lots of fine-grained discussions as well. I think Ron has already done an excellent job of saying some of the more important things we had to say about data, so I'm going to stand back and just think about the question of open government and what it actually means for us.

Let me start by saying something about our mandate, and I take it yours, or certainly the Government of Canada's, because there are those three streams as we like to say: open data, open information, and open dialogue, and the Ontario government, when it gave us our mandate, identified those three streams and we thought a lot about this.

First of all, I would say this. My background is that certainly I have lots of experience in e-government and open data, and so on, in the past, but my real interest is in dialogue, collaboration, and engagement. So what am I doing here? Well, it turns out that open dialogue is a really important part of this.

If you look at the Open Government Partnership and I'm sure you all have had lots of discussions around this, like your committee, it's interesting. Ask yourself where most of the focus is around those three streams. Not surprisingly—and this is a good thing, not a bad thing—it turns out that much of it is focused on data. Why is that? Why aren't we talking about information today? Maybe it's your committee's mandate. Why aren't we talking about dialogue today or talking about it only indirectly, although I see a lot of it surfacing in the conversation? I guess what I'd want to say to you—no offence to anybody in the room or outside the room—is that for governments, the easiest thing to do is open data. Guess why? It doesn't really compromise a lot of the traditional forms of governance we have. You don't have to say a lot. You don't have to give away any power. You don't have to open yourself up. There are some risks, and we heard about some of them, but by and large open data is the easiest place to start, and that's a good thing. I'm not opposed to it, folks. It's a great thing.

But what do we need to do after that, or where do we start to go after that? I want to talk just a little about that. It was interesting, our mandate for the open government engagement team was to focus on open government, and in fact, what the government said to us was to spend 60% of our time on open dialogue, 20% on data, and 20% on information. I thought that was really interesting, and the reason they did that was the recognition that it's time we moved the yardsticks on the dialogue piece and began to ask ourselves how do these three things actually fit together and what's the connection between them.

Let me come at this a little bit through the lens of data, and again in a very high level kind of way. First of all, if someone were to ask me, or if I were to ask myself why we would want open data, there are probably at least two very big things that we've heard a fair bit about today. It won't surprise you. We hear about the commercial benefits, right, that if we unlock these natural resources of the future of the information age we will be able to create new products and services, and that's good for the economy and that's good for all of us, and I'm hugely in favour of that.

Another reason, which we probably talk a little less about but we surely have heard about, is evidence-based decision-making. If you want to make good decisions, you need good information, and the availability of data makes that promising and important.

Let me say just a little bit about each of those and the role that dialogue plays in helping us to realize the full potential of open data.

Harvey talked about geography. I want to say a little tiny bit about it as well. You may or may not have heard of the Canadian geomatics round table. It is hosted by Natural Resources Canada, but the geomatics round table is actually quite a remarkable group, a new one. It's developing. It has about 25 or 30 members, and the round table is focused on geomatics information. It's a multi-sectoral partnership. It is a formal partnership or a round table that involves provincial governments, a number of federal departments, NGOs, universities, and a variety of other stakeholders all around the same table.

The basic reason all those people are there is that over the last four or five years it has become increasingly clear that spatial information is hugely important. I think Harvey has done a fabulous job of beginning to point out the complexities of thinking about how we

not only need spatial information but also how we're going to use it and what we're going to focus on when we organize and make this available in the future. It's not like there's just something out there called spatial information. It's how we put the stuff together, how we use it, and we can make choices about how we invest our resources and about what's important, whether it's for commercial purposes or other purposes.

● (0930)

[*Translation*]

The Chair: Thank you for your presentation.

I thank all of you for your very interesting statements.

We will now be moving to a somewhat more interactive part of our meeting, that is to say the period reserved for questions from the members of the committee.

We will begin with Mr. Ravignat.

You have five minutes.

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

Thank you to the witnesses for being here today; we appreciate your presence very much. The presentations were very stimulating.

[*English*]

At the beginning of the federal government's initiative with regard to open government, they were talking about open government writ large. What's interesting about this study right now is that it's a great representation of how they've narrowcasted what an open government is now.

This study we're doing is not about open government. It's about open data, and it's particularly about a portal. In a way it's about congratulating themselves that they went ahead with something concrete. There's nothing wrong with that, but the results are rather small. There are all sorts of reasons for that, which we could get into, but the reasons don't really interest me right now.

I have a very specific question for Mr. Giggey and Mr. Low. The largest municipalities in Canada are represented here, so just a really quick question to the two municipalities. How useful is the data that is now on the open portal of the federal government?

Mr. Robert Giggey: In a way, I think it is a question for the consumers of the data. Harvey will speak to it as a consumer and his role with the city of how useful that data is to the city itself. I think it goes a long way in identifying the gaps in the data. From my point of view, data has been placed up there across different industries or topic areas of the government. Now we can identify the gaps of what's missing, but that first step needed to be taken.

Mr. Mathieu Ravignat: But it's not necessarily the data on the portal that is useful but the absence of data that wasn't put on the portal that is useful.

Mr. Robert Giggey: I think it does serve a purpose around that, whether it's data standards or formats or the topic itself that's accessible.

I'll let Harvey speak to it.

Mr. Harvey Low: First of all, I want to congratulate the government. I agree with Don. It's small steps. It's great to have a data portal, but you asked me whether it's relevant, useful. So here goes. I'll give you a candid answer. When I was looking through CivicAccess, an online user group of users of data, I happened on a Twitter trail that said that one of the most popular data sets was caribou trails or migration patterns. I've never seen a caribou in Toronto. I don't know if Robert has seen one in Ottawa.

So there's the whole issue of making sure there are data relevant to social issues. I come from the social realm. Someone on the infrastructure end might talk about transportation data. There are data out there that you've picked, which is great, it's low-hanging fruit, but begin to think about data in terms of what people need, what service agencies need as well. There's that angle.

The other piece, and Don talked about it as well, is geography. The data to us in municipalities is not as useful if it is only at the city or even provincial level as a geographic layer of data. We need data at smaller levels of geography. StatsCan does this already. I'll be technical for a moment. Census tracks, dissemination areas, the geography is there. You have a federal department in StatsCan that has a file that can link and aggregate data to many different levels of geography. Use it. Use StatsCan to bridge that link across federal data sets and provide some of that data at different levels of geography. It'd be meaningful to us and users of geomatics across the board.

Mr. Mathieu Ravnagat: That's great, thank you. I have to stop you there because I have other questions, but that was really helpful.

Maybe more big sky, it was very interesting to hear about the open dialogue piece and the three pillars of open government. I'm looking for best practices. The federal government has done very little on the open dialogue front, not zero, but very little, which in a way has made these data inoperable for a majority. Had there been the open dialogue piece, there would have been an understanding of perhaps what was needed to be put on that portal, in what format it was needed, and how it could be searchable and so forth. There seem to be best practices of the provincial government.

I'd be interested in hearing what the federal government could do more of with regard to open dialogue.

• (0935)

Mr. Ron McKerlie: One of the things the provinces moved forward with is a consultation directory, which we're about to launch. What it does is provide, in one place, information on all of the opportunities the public had to weigh in on. Policy or discussion or dialogue are issues that the government is interested in hearing the public's opinion on.

It will give them social media ways to ask questions or to offer ideas. It requires a best practice that after we go out and hold these sessions, we document all the learning from that—what we learned from the public—and archive it permanently on this site, so we'll be able to go back every five years when we do a five-year poverty plan

to look at what we learned five years ago when we went out and talked to the public.

That consultation directory is developed and ready to go, and it will be evergreen all the time. Every time we think about going out to talk about transit funding, or minimum wage rates, or whatever we're talking to the public about, that will be the first place you will be able to go. You will be able to see how long it's open, where the physical sessions are, and what the social media ways of interacting are so that we can listen to what the public has to say.

The other thing we're doing with Don's help is experimenting with some other ways to engage the public rather than the standard: bring everybody into a room, frame the issue, hear what they have to say, thank them, send them away, and then create our own set of policy options and implement them ourselves, and so on.

What we have been trying for over a year now, with Don's leadership, and it has shown some good success to date, is to bring all the vested stakeholders around an issue together and have them own the process, where the government simply sits at the table as a voice. It's challenging because it requires us to give up power. It's challenging because we lose control of the timing of the process, and it's challenging because you have to get groups that might not normally speak together and work together to speak together and work together. But what it has proven is that you get much better solutions, you get much richer dialogue, you get stakeholders who are willing to bend and compromise, and you get much better outcomes.

[*Translation*]

The Chair: Thank you. I must stop you here, since Mr. Ravnagat's speaking time has expired.

Mr. O'Connor, you have five minutes.

[*English*]

Hon. Gordon O'Connor (Carleton—Mississippi Mills, CPC): Good morning, gentlemen.

I've been listening to your briefings, and the basic message I'm getting is that it's not so much the public that is involved in this, it's more like information nerds, or whatever you want to call them, people who are specialized in these areas and can manipulate the data.

In one of our first briefings, members of the government bureaucracy came here. I asked them how they are radiating the information of what they are doing. Essentially the answer was, they are not. They just put it out, and they don't do anything with it.

I'd like each of you, please, to comment on how you would get the information out to the real public who can use this information.

I'll start with Mr. McKerlie.

Mr. Ron McKerlie: Thank you.

First of all, our experience has shown that the number one users of government data are the public servants in that jurisdiction. Number two are public servants elsewhere. Numbers three and four are developers and researchers.

What we're trying to do is to go out to those communities, specifically talk to those communities—public servants, developers, and researchers—and make them aware. We are connecting them through Twitter and through other online media. We are making them aware when we release data sets so they know exactly what's available and how they can use it. We used the voting tool so we could hear from them in terms of what data they wanted access to, and we are now working together to release the top 25 data sets they had voted on.

• (0940)

Hon. Gordon O'Connor: Mr. Giggey.

Mr. Robert Giggey: Thank you.

I really think open data itself is kind of a nerdy thing, to use your term, so it can only be expected that the biggest users are those who can manipulate it. I think what governments can do is to leverage that.

The point of this is that we typically had released reports, released analysis, showed data aggregated, and provided reports with that. But that's not all of it. People wanted the data so they could do their own analysis and they could come to their own conclusions. In a way, we're going to end up going in a little bit of a circle because we have now released some of the data, at least, to help them along that way. Now the question is being asked: should we not be taking that data and making it into a format that the everyday user can consume?

I think the focus now should be to continue to work on getting data out that's usable and useful across a broad sector. Then depending on what comes out of that and what the users are doing with it, we'll see if there's a gap to make it so that the everyday member of the public can consume that. But we should be reaching out to those who can do it and who can create the value for the end citizen and user.

Hon. Gordon O'Connor: Mr. Low.

Mr. Harvey Low: Open data is not open government. These are terms that should not be used interchangeably.

Maybe the best way to answer your question is to give an example. When we provide a data set through the City of Toronto's open data site.... Let's use a federal source. We just put some data on immigrant landings from the CIC and some other data from income tax files about people's income circumstances. When that data goes on the open data site, that same data is also released through that application I mentioned earlier, Wellbeing Toronto. It's an application that is a tool set. It allows users to look at that data through a different lens. It's a geography lens at neighbourhoods so it's the same data.

This is one place that's raw data and the other place is data in a business context. It's an economy of scale. Don't put it in one place. Put that same data set elsewhere in areas of the city or areas of the federal department where you think that the data would support and give awareness to the public about that service area's mandates. Those are just a couple of examples.

Dr. Don Lenihan: I think it's a great question.

I might come at it a little differently. First of all, let me just say this. At least in the short-to-medium term I'm probably maybe a bit

like you—I hope I'm not reading this into you—I'm a little skeptical about the likelihood that ordinary Canadians are going to use a whole lot of data sets. I'm not.

There are two things. I think if we want to engage them on this and make it meaningful to them there are probably two other ways we have to come at it. One of course is benefits. If there are really important benefits, it's making the public aware, whether it's economic benefits or better outcomes from government or whatever it may be, so that they see that. I would go even further. I would go back to the engagement question and the dialogue question. I would give them some ownership.

What I mean by that is the more we move down this road and the more we solve the short-term, immediate problems of data management and start to raise the bigger questions of what we want to do with this new resource and what we want to do with coal or oil or something else, people have a view on that. If they start to realize that oil is worth a lot of money and it's going to make them rich or poor or it's going to change the world in which they live, they start caring about how it gets used even if they don't know what oil is made of.

We really might want to start thinking about how to manage and make data more valuable and useful and the benefits to flow. Why wouldn't we ask people what they think we should do with it, because what they say will make a difference? They may not be experts on how to integrate it but they are the experts on what we should get out of it at the end. Is it going to make us better off and how?

[Translation]

The Chair: Thank you.

Mrs. Day now has the floor for five minutes.

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

I thank the witnesses very much for being here.

You know that according to the Charter and the G8, data must be free, universal and accessible. That is part of the basic tenets.

Setting up such a structure means there are initial costs. Can you tell us how much the City of Toronto and the Ontario government invested?

• (0945)

[English]

Mr. Ron McKerlie: I can start if you like.

On the question regarding how data has to be free and what the costs are of opening up data, our costs depend on the size of the data set and what has to be done in the conversion. They're relatively modest. They're being borne by the ministries and they're being absorbed as part of their normal operations. We're looking now at changing systems. New systems are being installed so that the data element can be stripped out, cells can be totally compressed if they need to be or they can be eliminated altogether if the sample sizes are too small. They can be auto-published to a portal. That's the vision in the interim costs. Is there any sense of magnitude of cost?

Mr. Marc Foulon (Head, Open Government, Ministry of Government Services, Government of Ontario): As Ron McKerlie said, right now there are a lot of manual costs. That's why we did the voting process, to try to limit those costs and find out what people care about the most. Internal staff are manually releasing data that people find of high value. There is some cost associated with that, obviously, time and resources. In the longer term it's to build our net new IT applications and solutions where they're going to be automatically generating open data in a very automated way where it will be very low cost for us in the future.

[Translation]

Mrs. Anne-Marie Day: Is there some complementarity, some harmonization between the open data portal of a large city and that of the province or the federal level so as to maximize its use?

[English]

Mr. Ron McKerlie: Yes.

I think the biggest opportunity is around a common search feature. I have a slightly different view from what was expressed earlier. I don't think we should replicate the data all over the place. If we do that we're going to have huge storage costs, plus the complexity of trying to keep track of the original source of that data.

I think we should jointly develop a common search engine, so wherever that data resides we can search federal government data, provincial data, municipal data. There's no wrong front door into finding that data. I think that would give us some huge economies of scale.

I think there are some other savings we can create, if we can get to standards. It will make it easier for the research community and the developers if we have standard formats, and if we agree on meta tags, for example.

I think those are areas where we could save, perhaps not money for us immediately but money for the users down the road.

[Translation]

Mrs. Anne-Marie Day: Do you have any statistics on the use made of open data? For instance, are the province's or city's data accessed on a regular basis? What type of data is being asked for? Are they mainly local data such as data concerning the weather, the subway, schedules, or are they data related to geomatics, for instance? Do you receive a lot of requests from businesses?

[English]

Mr. Ron McKerlie: The top voted data sets right now are transportation data, finance data, and health and education data. Those would be the top ones.

Am I missing anything?

Mr. Marc Foulon: In addition to those would probably be general government types of services, so procurement and HR information. A couple of other higher ones are freedom of information statistics, as well as general information about the open government directory of staff. That is in the top 10.

Mr. Ron McKerlie: In terms of actually using the data, we've had a number of applications created with our data. One is called iamsick.ca. The developers took information on hospitals, emergency clinics, pharmacy locations, hours and language of service,

and a host of other things. Some of it came from Stats Canada and some from the province, and some came from the City of Toronto. They created an application. If you're moving into a neighbourhood and need medical care or attention, you can find a pharmacy or an after-hours clinic or a doctor or an emergency room.

There are a number of other applications that have been created from our data.

I would say that the usage is still modest, though. It's not huge volumes; it's much smaller volumes. A lot of it is being used to answer questions.

One of the applications took water sampling data and source water data, Google mapping, and Stats Canada population data. Now you can click on it and find out the source of your drinking water anywhere in Ontario, what percentage of the population uses that water, and the recent test results. This is post-Walkerton and the problems we had with water quality in Ontario.

It's smaller volumes.

Thank you.

• (0950)

The Chair: Thank you.

I will have to stop you.

[Translation]

Perhaps Mrs. Day can discuss cities a little later during the meeting.

We will now hear from Mr. Aspin for five minutes.

[English]

Mr. Jay Aspin (Nipissing—Timiskaming, CPC): Thank you, Mr. Chair.

Welcome, gentlemen, and thank you for assisting us with our study.

According to the information we've been given, the Province of Ontario and the City of Toronto, and I guess most of the larger municipalities, are working with the open data of municipalities in a group called the Public Sector Open Data.

I wonder, are the formats that are being developed by the province and participating municipalities in this forum consistent with those being offered by the federal government? I'd like each of you to comment on the level of consistency or where improvements need to be made, starting with Mr. McKerlie.

Mr. Ron McKerlie: I'm going to delegate that to Marc. I think he has that level of detail.

Thanks.

Mr. Marc Foulon: Thank you.

Yes, as you mentioned, we sit on the PSOD, in the Province of Ontario, as well as some other committees with the federal government, other provinces and municipalities, and have some of those conversations. I'd say there's not a set standard or metadata that's out there right now that is being used across all the different levels of government. That's something that we do need to improve on and come together on to put something in place. Even within the Ontario government, with our various ministries, sometimes it is difficult to have some common standards in place.

Our catalogue is an example. We have five or six standard mandatory metadata categories, and a few others that are optional, or depending on the data set, can use certain other characteristics. We have tried to put that in place for most of our data over the last few years, but definitely different levels of government, not only within Canada, but even, say, in North America. That's something we should look at, so researchers, either internal to government or external who are using the data know they're using apples to apples data, and can compare them, match them up, use them as part of their evidence-based policy work.

Mr. Robert Giggey: If I can add, it's difficult to get common data formats for particular topics across the levels of government simply because in many cases you're working with different types of data. There'll be a few circumstances where it's the same. It could be around transportation. If you're reporting on incidents or traffic on federal or provincial highways or municipal roads, you could do that. But in a lot of cases we're looking at topics separately, so this is difficult to do.

One of the biggest gaps is around all the cities themselves. For those using the data, one example is governance, and those that are looking at how governments are run from looking at minutes and agendas and voting records. In the federal government you have one place to look at data formats. In the provinces, you have a few, and you can work on getting that data out so it's usable, but once you get down to the municipal level, if you're trying to show what's happening at all three levels of government, you now have thousands of cities to work with in trying to get a common format.

The cities themselves have a lot that we need to do to help with using common formats. One thing is to look at what the rest of the world is doing, because everybody's tackling this and trying to solve it. In terms of interoperability some global standards have developed, and we have to look toward that.

It's still early. I think municipalities have a lot of work to do to get common data, as well as work with the other three levels to see. Even though there are different types of data on the same topic, how can we get them to work together as cleanly as possible?

Mr. Harvey Low: In my opinion, there are two types of standards. There's a technical standard, and then there's a policy standard. Technical standards deal with different formats of data, whether they're mapping-format data sets or Excel or that type of thing. I think those are probably less of an issue.

Of greater importance are the policy standards. For example, when we work with low-income groups, there are many different definitions and many different measures of poverty. When you have different levels of government, even ministries within governments, releasing data sets called poverty, there needs to be a consistent

metadata set that defines what all those indicators mean. I would say that's probably more of the challenge.

Finally, the other thing that we haven't really considered, haven't spoken about, is information technology. When I mentioned earlier that we release data sets on numerous different portals at the city, we've used a centralized data warehouse and we've used centralized geographic spatial software to make sure that consistency and that vintage is the same throughout the city, so the other standard there is geography, and that one is probably the easier one.

• (0955)

[Translation]

The Chair: Thank you. Mr. Aspin's speaking time has expired.

Mr. Dubourg now has the floor for five minutes.

Mr. Emmanuel Dubourg (Bourassa, Lib.): Thank you, Mr. Chair.

It is now my turn to greet you and to thank you for the statements you made. They were very interesting. We have talked a lot about transparency and accessibility. Accessibility has been discussed because we know that data is for everyone, be they youths, children, experts or specialists.

I agree entirely with Mr. Low when he says that people do not necessarily pay attention to the various jurisdictions of government. They need information, and they look for that.

I would like to know if you have had any concerns regarding the accuracy of these data. There are many users, be they students, experts, researchers or historians. Did you have any concerns of that order when you worked on setting up your platform?

[English]

The Chair: Fine, in the same order as usual.

Mr. Ron McKerlie: Thanks very much for the question. It's a great question.

What we know for sure is that all data has errors in it. Some of them were created, some of them occurred when the data was collected, and some of them were created and built-in as the data was put together. What we've found though is that the more sets of eyes on the data the higher the quality becomes. So it's actually improved as we've started to open it up because as people, particularly public servants, start to look at it they question missing data, they question anomalies that don't seem to make sense. So it's improved the quality of the data. Yes, we do have concerns, absolutely, that the data isn't perfect, but we understand that a lot of the people working with it are giving us some grace in terms of understanding that it won't be perfect and the quality is improving as more sets of eyes look at it.

[Translation]

Mr. Emmanuel Dubourg: I have another question.

The Chair: Just a minute, I think that Mr. Low would also like to reply to your first question.

Mr. Emmanuel Dubourg: Go ahead Mr. Low, please.

[English]

Mr. Harvey Low: Yes, basically the issue that we have in quality of data.... Ron is completely correct. There is no perfect data. I think that the solution here is to provide equal resources and opportunities to support federal departments in explaining their data, providing a proper metadata base or definition and one that's written in English. Do not have statisticians write it. Having somebody who is aware of communications so that they can explain information in a user-friendly manner certainly goes a long way in increasing the public's trust of data.

The final comment is to begin to leverage and use those departments that understand the data, Statistics Canada being one of them. If people know that it's coming from the national census, we find that there's very little question on the reliability of the federal data. The national household survey is a different matter, but the census, certainly, is a reliable source of data and we've heard that loud and clear in the community.

[Translation]

Mr. Emmanuel Dubourg: I would also like to put a question especially to M. Giggey, since he works for the City of Ottawa.

Did the two official languages pose a problem when you worked on open data? Some information may be available in English, but partially available for francophones, for instance. Was that a problem in your case?

• (1000)

[English]

Mr. Robert Giggey: Thank you.

We are I think in a unique position. There aren't actually many bilingual cities or jurisdictions releasing data, so I've actually connected with staff in the Treasury Board on this topic to help try to solve it.

One of the key problems we have is the operating language at the city, the language that most of our base systems are in, is English. So it can make it difficult to translate data. The position we have at the City of Ottawa is that we've made all the metadata—all the information about the data itself to help with discovery and access—all bilingual because that's somewhat easy to deal with. The data itself comes in whatever language the base systems are in. Many of our front-line systems like recreation, culture, those are all being translated by the operations anyway because they are being used for public information. But there is quite a bit of data that currently isn't translated. It's all available to be translated upon request. But in the interest of moving forward and trying to get the benefits out of open data, we've chosen this model. So far I believe it's worked well. For users so far no issues have been identified and they are able to use technology to translate the information around the data itself.

[Translation]

The Chair: Thank you for your questions and answers.

Ms. Ablonczy now has the floor for five minutes.

[English]

Hon. Diane Ablonczy (Calgary—Nose Hill, CPC): Thank you, Mr. Chair. Thank you, gentlemen.

I think it's fair to say that this whole concept of open data is in its infancy to a big degree so we're now working together to try to do this as best as possible. Bearing in mind the great plea that we've heard about consultation and dialogue, I'd like to ask the smartest people in the room—who happen to be our witnesses—what databases or data sets from the point of view of the province or the cities that you represent are of most value to you or do you reckon will be the most valued to you so that we can start to do a reality check on our own priorities?

Mr. Ron McKerlie: That's a great question. I'd have to give that some more thought. We use a lot of the federal government data now, particularly a lot out of Stats Canada. We use a lot of the place data, the geospatial data, that we've talked about already.

Most of the data we use is very contextual. If we're trying to solve a transit issue, we're looking for anything you know about funding or building transit, for any data you would have on that. If we're looking to roll out electronic health records, we're trying to figure out what Canada Health Infoway and others would have in terms of health information.

So the answer is that it really depends on what problem we're trying to solve at the time. Then we're usually looking for information to help us solve that problem. I'm not sure I could give you much more detail than that, because each ministry has a very different set of issues they're trying to deal with. What would be useful, though, would be to see what data you possess, such as an inventory, so that we could get a sense of what data might be available to us as we're trying to solve problems.

Mr. Robert Giggey: I can say that at the city level, for sure our most popular right now, although some of this is based on what we have available, are recreation, transit, transportation, garbage collection, cultural events—all of those front-line, immediate services that people are looking for. They are the most popular.

To pick up on a previous question, though, sometimes you can't tell the value of a data set simply by the number of times it's been downloaded. You can have cases where somebody has downloaded the data once, but viewing that data can be thousands of users.

Hon. Diane Ablonczy: But what does the City of Ottawa want to see from the feds?

Mr. Robert Giggey: To go along with that, I think transportation is a key one, as is environment, health, and spending itself. We've had requests to see the spending as it goes through the three levels of government.

Certainly around transportation it's an easy one. Take the conditions of the highways and the roads. There's the City of Ottawa's jurisdiction, for example, and the NCC's. If we're going to look at helping people get around the city, we need the city roads, we need the provincial highways, and we need the NCC, because they have the parkway on the edge of the city. For recreation, it's the same thing. They offer events, they have parks, and they have the Rideau Canal skateway. This is the kind of day-to-day stuff that people want to do.

• (1005)

Hon. Diane Ablonczy: Does the federal government keep that data? Or would that be provincial?

Mr. Robert Giggey: In the case of the NCC, I'm not sure, but I think federal highways would be one example in terms of that indication of how you get around.

Mr. Harvey Low: Given that our major role is to identify in a better way the clients we serve, we need that profile data, that socio-economic data.

If you're asking about what data we need from the federal departments, I could give you a list right now, as follows. We are in very short supply of health data. The federal government is sitting on a gold mine of health data through CIHI, the Canadian Institute for Health Information. That would be a great start. We get very little data from them.

There's also the CMHC housing data. That one is a little different. We would like to get that data at a specific level of geography. There are little nuances to the data.

The big one, of course, is the national census. I realize this is a political hot potato, but I'll say what's on my mind: bring back the long form. We've heard it loud and clear in our community. That is the only substantive source of socio-economic data that's reliable.

Finally, sometimes it's not just about the type of data, it's about the way the data is provided. To give you an example, employment is huge in Toronto. Job creation is huge in Toronto. The labour force survey and all of that data that comes from the feds is only done every five years. We do not get that level of granularity in neighbourhoods, so you need to begin thinking about expanding your delivery and dissemination of data on a more timely basis rather than every five or seven...or 10 years on religion data.

[Translation]

The Chair: Thank you.

Mr. Lenihan, did you want to add something?

[English]

Dr. Don Lenihan: As somebody who's not part of a government, I would just to defer to my colleagues here. They would have a much more direct view on this than I would.

[Translation]

The Chair: Thank you.

Mr. Ravnat, you have five minutes.

[English]

Mr. Mathieu Ravnat: Just as a quick comment, I hear you, Mr. Low, about the long-form census data. I come from a research background. It's basically a national tragedy that we've lost so much useful data. It's really unfortunate. We've called on this government to reinstate it, but there seems to be some resistance there, for unknown reasons.

I want to come back to two things. One of them is access to information and the relationship between open government and access to information. Unfortunately, there's a lot of frustration with regard to access to information out there, particularly the practices of this government—how the act is being applied, delays, etc. There's a relationship, I think, between reducing the amount of access to information and the accessibility of data. I think if you make useful data open by default, then there's a potential to reduce access to

information requests. It could wind up saving taxpayers money, really, and a lot of frustration with regard to journalists, etc., trying to find information.

Just your comments on that relationship between access to information practices and open government would be greatly appreciated.

Mr. Ron McKerlie: Maybe I can start. Thanks for the question.

All the research we've done suggests that the easier it is for access to information the more demand for information you'll get. We honestly believe that even though we are moving down a path that will both automate the front end of the FOI process to make it easier for people who ask for the information, as well as proactively releasing information before they ask.... We know it's information that people always ask for. Our goal is to try to make information more readily available, to point to it.

The other thing we're trying to do is to use data visualization and information visualization to make it easier to understand once you find it. A lot of the financial information is available through three volumes of public accounts. The problem is nobody can read or understand it. Taking that financial information and visualizing it is the next step for us to try to make it simpler to understand.

Mr. Mathieu Ravnat: I'm rather impressed with all the efforts that have been made by the provincial government on this concept of open government. There are many things that the federal government could be doing, and there are clear examples at the provincial level where we could probably learn from your experience.

With regard to your understanding of “open by default”, a lot of research exists in the federal government departments that is being done in a transitional way, that's not necessarily complete, not necessarily fully baked, yet, useful. I wondered if you could comment on the default portion and the timeliness of that portion with regard to what the ideal is in making that data available as quickly as possible.

• (1010)

Mr. Ron McKerlie: We haven't got it approved yet but we're going for 90 days, so the information gets created and it's out within 90 days. That's the goal. We'll see whether that survives cabinet or not but that's what we're offering.

Mr. Mathieu Ravnat: It's amazing—90 days.

Do I still have some time?

[Translation]

The Chair: Yes, but Mr. Lenihan seems to have something to add on that matter.

[English]

Dr. Don Lenihan: I'd like to make a couple of brief comments on “open by default”, which of course is the name of our report, and we had a lot of discussion around that concept or principle.

First of all, to go back to your question about whether more open data would reduce the request for open information, I doubt that, frankly. I think it's about culture and expectations and about how we understand our relationship to government. If it turns out that we start to think that this is public property and should be available and start thinking about how we're going to govern, my guess is that's not going to restrict the number of requests for open information. We're going to expect access to that too.

I want to go one step further and just say this. This is a much longer discussion but we spend a lot of time talking about this with each other in the group and also with the participants in the various sessions. The way we've governed for a long time has been around an assumption that governments retreat. They make decisions and they need to be in private when they make these decisions. Then they announce and they communicate and defend their policies.

We kept hearing from people of all sorts that the more complex it gets out there, the more information technology and communications and other things are out there, the reality is that model simply doesn't work for all kinds of reasons. It doesn't mean it's bad; it served us very well. If the model of policy-making requires a high level of privacy or secrecy that you can no longer control, it puts it at odds with ourselves. I think what we heard a lot of people saying is that the real challenge for politics over the next 10 years is recognizing that you can't make policy that way. I'm not saying any policies but just having a model based on secrecy in the traditional way increasingly will not work, and how you get ahead of it where the principle is not open when we say it's open, it's open by default, may well be the challenge that lies ahead for all of us.

[*Translation*]

The Chair: Thank you, Mr. Ravnat. Your speaking time has expired.

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

[*English*]

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

Thank you to all witnesses for being here today.

Mr. McKerlie, we spoke earlier about intergovernmental data coordination. Is this not like negotiating a trade agreement in a multilateral sense? If you're talking with different governments and when you begin to group data, you have to come up with a common set of definitions of what each word means. Is that an issue as you see it right now or a potential issue going forward?

Mr. Ron McKerlie: Sure. Whenever you're going to spend time trying to agree on language issues for metadata and standards in terms of formats, it's tedious work and it's long and involved. The benefits though, I think, are huge, and they're huge for a very long period of time. I think it would make Canada unique in the world if we had, among our various levels of government, a common set of metadata terms so that people, if they wanted environment data from any level of government, could actually find the environment data they were looking for from any level of government. If we had a single search engine to tie that all together, we would be unique in the world.

Mr. Mark Adler: I agree.

Who, in your mind, would create all of that?

Mr. Ron McKerlie: I think we already have a number of forums that work well.

I used to be the corporate CIO for the Province of Ontario. We have a forum at which the CIOs from Treasury Board as well as from each of the provinces, and so on, get together regularly anyway. Maybe that would be something we could task to that group. We need the municipalities there as well, obviously. Maybe it's an expanded version of that.

But it's technical work.

•(1015)

Mr. Mark Adler: It's very technical.

Mr. Ron McKerlie: It's grind-it-out difficult work.

Mr. Mark Adler: Is that possibility on the horizon right now or is it being discussed at least?

Mr. Ron McKerlie: I'm not aware that it has been discussed. I think everybody understands the need and I think everybody is waiting patiently for a champion.

Mr. Mark Adler: Okay.

Mr. Lenihan, you mentioned earlier that we'll be making better public policy decisions because we'll have more information at our disposal.

I would tend to agree with that. However—and Mr. Low may also want to come into this—do we not run the risk, first of all, of policy paralysis when we have too much data at our disposal, because the information is just coming out of a fire hose and we don't know when it's going to end and when enough is enough? Also, secondly, I foresee a bigger problem. Because we're relying only on empirical data, do we not lose the fact that for the big policy decisions in Canadian history, like those on national medicare, building a railroad, or these big vision things, no data was available? Those were based on somebody's vision of where they wanted to take our country. Do we not lose that if we rely too much on the information and the empirical evidence?

Could you address those two things?

Then, maybe, Mr. Low, you could chime in too.

Dr. Don Lenihan: First of all let me just say this. I don't think I said we will be making better decisions; I think I said I hope we will.

I actually agree with what you said. I think we can become overwhelmed by information and data, and that would be a bad thing. Part of what I wanted to emphasize with regard to the importance of open dialogue is that if we don't talk these things through about how we organize the data, how we use it, how we understand it, how we interpret it, then we're at risk of having so much of it with no real coherence to it that we don't know what counts and what doesn't. So part of my argument would be that it's one of the reasons we need dialogue, so we can understand and agree, at least on some levels, as to what it means to say this is about poverty or about financial success or some other thing.

The last thing I want to say is that I absolutely agree with the last thing you said about vision and so on. I don't ever want to live in a world that is run by nothing but scientific policy. I've railed against that all my life.

Here's what I would say. Big policy issues are a complex mix of information, knowledge, and choices. Choices are about values and priorities and lifestyle and all that. I don't want to lose that for a moment.

I think what we don't want to do, on the other hand, is decide that everything is just about priorities and values. It actually isn't. We do know some things about the world, and if we knew things about the environment, about the social environment, about business development, and about a whole range of other issues, those would inform our policy-making.

What I want to argue is that we have a chance to advance policy-making beyond where it ever was. That's not to say it's just about science. It will never be that way.

So I don't think we disagree at all.

Mr. Mark Adler: I don't either.

Thank you.

Mr. Harvey Low: That's a great question.

I totally agree that too much data is confusing.

The real solution to that is not counting, and that's why it's not good as a metric to say your measure of outcome is x number of data sets released. That's not the right metric of outcomes for an open data site.

You can get around this and you can have lots of data out there, but what you need—and Ron touched on this—is a good search engine.

I spoke a bit about needing a better way to tag data sets in terms of issue areas. If you can organize your data that way, it doesn't matter. Look at StatsCan. They have tens of thousands of data sets, but they do it well so you can easily find what you're looking for by either subject area or term or search. I think those are what the government needs to work on—that search engine and that taxonomy—and those will go a long way. Then you can have 10 million data sets out there and it doesn't really matter.

[Translation]

The Chair: Thank you Mr. Adler. Your speaking time has expired.

Mrs. Day, you have the floor for five minutes.

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

Sometimes our governments use data that has been redacted or blackened so that they are not accessible, claiming that they are confidential and that cabinet has decided that they would not be available.

Mr. Low, how can that type of situation be avoided when it comes to open data?

• (1020)

[English]

Mr. Harvey Low: It's a very good question.

The solution to that is already in the works by several ministries. Again, I go back to Stats Canada. I think geography is one of your solutions. When we as a municipality want to better understand the clients who we serve, one way of doing that is to get that aggregated level of socio-economic data. We're not interested in individual profiles of people and how much did they make out there. What we're really interested in is a higher level of geography that makes sense and allows us to do that place-based planning in neighbourhoods. That would be one way around the privacy issue.

[Translation]

Mrs. Anne-Marie Day: Mr. Lenihan, my next questions are for you.

We are heading for a communications era where people will have mini computers on their wrists, where they will have everything they need to surf the Internet and access open data at any time. People will know the subway schedules, the weather, etc. That is first-level information for a lot of people. However, if we see something flying by in the sky, we may wonder what is going on and what that is. So now we are talking about space or aerospace. So then that may involve the planets and so on. All of this is unfolding very quickly.

The ordinary citizen may access highly specialized data as well as specific municipal data. For instance, he or she may want to know what day the recycling truck will come by; that is important for people. This will lead them very quickly to provincial data on health and federal data on filing income tax, or geomatics. People will have access to all of this data easily and immediately.

Earlier you talked about forums you had set up that allow people to communicate interactively. How do you assess the data, and the effectiveness of the systems that have been put in place? How is that dialogue with the population carried out?

[English]

Dr. Don Lenihan: That's a great question.

I want to return to this. That is, when we think about the role that dialogue plays in this, and I guess that's what I keep returning to, we do have some emerging models. The example that I gave of the geomatics round table is a really good one because we are going to have to make some fundamental choices that will resonate with citizens, and this goes back to their ownership of it, with citizens and popular use. Ultimately, there are only so many resources and so much effort and opportunity to make things available. As I think you've just rightly indicated, there is already and there's going to be more and more data. So how do we decide and how do we make available the stuff that really matters?

I would turn to models like the round table and say that if we don't learn how to have these discussions we won't be able to make those choices. Those really are policy choices at the end of the day.

I could go into greater detail about how these sorts of discussions will take place. The one thing that I would want to say, and I hope I'm not missing your question, is that there's justified fear among governments, especially at the political level, that in opening up dialogue—even the word is scary—we're sort of giving away all this power to somebody else to make decisions. I don't think that's necessary at all.

If I could make one point about open dialogue, it doesn't mean that government has to give away any authority at all. I think it does mean that government has to change the way that it makes decisions. So if you're part of a collaborative table having a discussion on setting priorities for how we will use our resources together in various ways, if you were there representing the federal government my advice to you if you were the minister would be, you don't ever let anybody tell you what to do. You're not giving your authority over to somebody else to make that decision. What you're doing is sitting there to work with others differently, where we're all trying to reach a collaborative answer that essentially serves the interests around the table of everyone in the best possible way.

My view about open dialogue is that it's not about giving away authority; it's about exercising it differently. That is, openly in a way that we recognize the need that we have to work with others to make choices so we can solve the problems you're talking about. My argument would be that we will never solve those problems in a vacuum because increasingly they involve more and more players.

I hope I haven't missed your question.

[Translation]

The Chair: Thank you.

Mrs. Anne-Marie Day: That is fine, except for the assessment of the data; I would like to know how that is done.

However, I think my time has expired.

Is that the case, Mr. Chair?

The Chair: Yes, but perhaps you can come back to that later.

I will now yield the floor to Mr. Hillyer for five minutes.

[English]

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

As much as I've appreciated a lot of the vision and the recommendations on where the federal government can better provide data and make it more accessible to the general public, I think the purpose of the open data project was to put it to the developers and anyone else out there to get in on it so that it's not driven just by the federal government. So while I think a lot of the recommendations are valuable, and we need to consider... Well, let me put it another way. We shouldn't be abdicating our own responsibility to provide better data just because we have open data. Just because we are inviting developers and other people in on it doesn't mean that we're done having to provide context, as has been put.

But one of the challenges in the context of the open data project is that we don't want the government itself to put too much context in the data for the open data project. Maybe we want to put context in for other things, but with the open data project itself, by putting

context on it we're eliminating a whole bunch of ideas that may have come up without providing the context.

That brings me to a question out of Mr. McKerlie's presentation. You said that you wanted to make sure you're pursuing a "quality over quantity" approach. Is that true just for the kind of data that the government wants to provide, or is it true in the open data context? If it is, I wonder why that would be. Why not have quantity now, which can be followed up as we go with quality, so that we don't have to wait for the quality to get people's hands on this stuff? Maybe they can help us provide the quality.

• (1025)

Mr. Ron McKerlie: We looked around the world, because we were a little late to the game, and found that a lot of jurisdictions were taking a lot of flak because they were just dumping data sets that were easy to dump out there, but weren't really of use to people. That's why we went with the voting tool, so that the public could tell us what's useful to them. So that's the quality, "quality" meaning value to the individuals, the developers, or the community.

We also did it because we have a restriction in terms of how much money we can spend on opening up data sets. "Open by default" means all data will be released, and over time it will all be open. We'll get there. For us this is staging so that we can get the most valuable data sets out first.

Mr. Jim Hillyer: Okay.

You say you're late to the game. If you're late, then we're even later. Who else is doing this, and are other federal governments doing it?

Mr. Ron McKerlie: Yes, the U.S. federal government and a lot of the states, and the U.K. were probably earliest off the mark with the open government initiatives. But there are lots of countries and jurisdictions that have done well. There are now over 400 jurisdictions, national and subnational governments, that are involved and have signed off on the principles of open government. Data is the easiest for a lot of people to start with, so a lot of them have already moved down this path.

In Canada there are lots of good examples. The B.C. government is doing quite well. The City of Edmonton is doing quite well. There are lots of success stories to point to.

Mr. Jim Hillyer: Robert, since you're more on the technical side of this, how do we make open data accessible to non-developers? Once the techies get their hands on it, would their apps, whatever applications they come up with, make it no longer open data?

Mr. Robert Giggey: One way we look at it is always trying to guess what the public wants to see and what they want to do with it. We found—and I know the municipalities are closer tied to the app space—that it's quite typical that when we want to provide a service, an online service or an app, we're trying to guess what the public wants and how they use it.

But the people who are doing something with the data now, they can do a better job because they're surveying, say, a market. So that app developer is not going to create an app unless there are people there who want it and are going to use it, so they're letting the market drive what will be done with the data, and it takes some of the guesswork out.

In some cases we'll be told directly what data people want, or what they want to do with that, and maybe there is a gap. Maybe in some cases there's nobody out there who wants to do anything, but the public is still asking for that or demanding that. Then instead of spending our time, effort, and resources on those things that somebody will create and do something with, we can instead focus on those areas that nobody's picking up and nobody's doing anything with. Whether it's research, community groups, special interest groups, or app developers looking at certain topics, they have kind of a constituency. They have people who are demanding stuff or asking for stuff, and that's why they're creating it, so they're kind of taking some of the guesswork out for us.

● (1030)

[*Translation*]

The Chair: Thank you. I'm going to have to stop you here.

This concludes the testimony. Thank you once again for having been with us and for having shared your expertise on this with us. I am sure that this will help the committee to continue its study and later to draft its report.

Before we leave, the members of the committee have some business to consider. Before we do that, I am going to suspend the meeting for a few minutes.

● (1030)

_____ (Pause) _____

● (1030)

The Chair: We will now resume our meeting.

I have a few details to share with you concerning the next meetings, but first of all, as you saw on the agenda, we're going to deal with a notice of motion from Mr. Byrne.

Mr. O'Connor, you have the floor.

[*English*]

Hon. Gordon O'Connor: Mr. Chair, I would ask that we go in camera if we're talking about motions. They should be in camera.

[*Translation*]

The Chair: A motion has just been introduced asking that we continue our meeting in camera. That motion cannot be debated.

[*English*]

Mr. Mathieu Ravignat: May we have a recorded vote?

[*Translation*]

The Chair: A recorded division has been requested. I'm going to let our fine clerk proceed with the vote.

(Motion agreed to; yeas, 6; nays, 3.) [*See Minutes of Proceedings*]

● (1035)

The Chair: Since the motion was agreed to, we are going to suspend the meeting for a few minutes in order to allow the technicians to do the necessary work.

[*Proceedings continue in camera*]

Published under the authority of the Speaker of
the House of Commons

SPEAKER'S PERMISSION

Reproduction of the proceedings of the House of Commons and its Committees, in whole or in part and in any medium, is hereby permitted provided that the reproduction is accurate and is not presented as official. This permission does not extend to reproduction, distribution or use for commercial purpose of financial gain. Reproduction or use outside this permission or without authorization may be treated as copyright infringement in accordance with the *Copyright Act*. Authorization may be obtained on written application to the Office of the Speaker of the House of Commons.

Reproduction in accordance with this permission does not constitute publication under the authority of the House of Commons. The absolute privilege that applies to the proceedings of the House of Commons does not extend to these permitted reproductions. Where a reproduction includes briefs to a Committee of the House of Commons, authorization for reproduction may be required from the authors in accordance with the *Copyright Act*.

Nothing in this permission abrogates or derogates from the privileges, powers, immunities and rights of the House of Commons and its Committees. For greater certainty, this permission does not affect the prohibition against impeaching or questioning the proceedings of the House of Commons in courts or otherwise. The House of Commons retains the right and privilege to find users in contempt of Parliament if a reproduction or use is not in accordance with this permission.

Also available on the Parliament of Canada Web Site at the following address: <http://www.parl.gc.ca>

Publié en conformité de l'autorité
du Président de la Chambre des communes

PERMISSION DU PRÉSIDENT

Il est permis de reproduire les délibérations de la Chambre et de ses comités, en tout ou en partie, sur n'importe quel support, pourvu que la reproduction soit exacte et qu'elle ne soit pas présentée comme version officielle. Il n'est toutefois pas permis de reproduire, de distribuer ou d'utiliser les délibérations à des fins commerciales visant la réalisation d'un profit financier. Toute reproduction ou utilisation non permise ou non formellement autorisée peut être considérée comme une violation du droit d'auteur aux termes de la *Loi sur le droit d'auteur*. Une autorisation formelle peut être obtenue sur présentation d'une demande écrite au Bureau du Président de la Chambre.

La reproduction conforme à la présente permission ne constitue pas une publication sous l'autorité de la Chambre. Le privilège absolu qui s'applique aux délibérations de la Chambre ne s'étend pas aux reproductions permises. Lorsqu'une reproduction comprend des mémoires présentés à un comité de la Chambre, il peut être nécessaire d'obtenir de leurs auteurs l'autorisation de les reproduire, conformément à la *Loi sur le droit d'auteur*.

La présente permission ne porte pas atteinte aux privilèges, pouvoirs, immunités et droits de la Chambre et de ses comités. Il est entendu que cette permission ne touche pas l'interdiction de contester ou de mettre en cause les délibérations de la Chambre devant les tribunaux ou autrement. La Chambre conserve le droit et le privilège de déclarer l'utilisateur coupable d'outrage au Parlement lorsque la reproduction ou l'utilisation n'est pas conforme à la présente permission.

Aussi disponible sur le site Web du Parlement du Canada à l'adresse suivante : <http://www.parl.gc.ca>