



HOUSE OF COMMONS
CHAMBRE DES COMMUNES
CANADA

Standing Committee on Government Operations and Estimates

OGGO • NUMBER 013 • 2nd SESSION • 41st PARLIAMENT

EVIDENCE

Tuesday, March 4, 2014

—
Chair

Mr. Pierre-Luc Dusseault

Standing Committee on Government Operations and Estimates

Tuesday, March 4, 2014

•(0845)

[Translation]

The Chair (Mr. Pierre-Luc Dusseault (Sherbrooke, NDP)): Committee members, good morning. This is our 13th meeting. As you know, this is our first meeting on our study of the government's open data practices.

We have with Treasury Board representatives with us this morning. We welcome Ms. Charette, Senior Director of Information at the Government of Canada, as well as Mr. Walker and Mr. Latour.

I will now give you the floor and I thank you for being here. A little later during the meeting, the members will ask you questions about your presentation.

Ms. Charette, you have the floor.

Ms. Corinne Charette (Chief Information Officer of the Government of Canada, Treasury Board Secretariat): Good morning and thank you very much, Mr. Chair.

It is a great honour to be here before the committee to speak about our success and our work on open data for the government.

[English]

I'm very pleased to be here with my two colleagues to talk about open data. I'll introduce Stephen Walker, who is the senior director for our information management policy sector as well as for open government at TB Secretariat. With him is Sylvain Latour, who is a director of our Open Government Secretariat at TBS.

The way we propose to cover the material this morning is that we have a presentation in two parts, and we propose to have a demo. We will go through the first part of our presentation.

[Translation]

You have in front of you a presentation which, I think, gives a good summary of the key concepts concerning open data.

[English]

We'll start off with essentially a primer on open data, what the key concepts are, and then we will stop and do a demonstration. You've noticed the screens in the room. We'll have a live demonstration. Stephen and Sylvain will go through our actual open data portal and show you some examples of the data and how the portal works. Then we'll revert to the presentation to give a summary of what different initiatives are going on within the federal government and with our colleagues across Canada in other jurisdictions, and in fact, on our initiatives internationally on the open data front. Of course, we'd be delighted to answer whatever questions the committee has.

That's what we propose by way of the three-section approach. Before I start into the first part of the presentation, I would like to say that we've just completed a very exciting weekend. On February 28, Minister Clement launched the Canadian Open Data Experience, which is an appathon challenge that brought together, finally, 927 registered participants from across Canada, from universities in all provinces across Canada, to try to see what kinds of applications they could develop using Canada's open data information published on our portal.

It was a very exciting weekend, and at the end of it, preliminary reports suggested that over 100 different apps were developed and will be validated and vetted and be the subject of tough competition. The finale of CODE will be March 28, in Toronto, where the 15 finalists will review their apps with the judges. The finalists will be awarded a prize.

This is very exciting because this would be our first national CODE appathon. Different provinces and cities have had a few, and there have been a number of efforts across Canada, but this is the first on a pan-Canadian basis. The success of CODE is a testimony to the enthusiasm and interest in Canada's open data portal and the information that we make available to Canadians.

With that, I will go into the presentation and hopefully help to demystify this. We'll be doing section 1.

Page 3 is titled "Open Data Fundamentals". I apologize that some of you may be well aware of this, but we weren't sure so we thought we'd bring everyone to a certain level of knowledge.

So what is raw data?

Raw data is machine-readable data at the lowest level of integration that can be reused alone, or mashed up—as the term is—with other data in innovative ways. The government either generates or collects and aggregates a vast amount of raw data. The best example of raw data would be weather data that we collect through sensors and radar and a variety of other means. We turn that into raw data, numerical data that is available for further processing and manipulation.

•(0850)

So what is metadata? Metadata is data about data. Metadata is key to the potential of open data. Without metadata, the vast numbers of data sets and information that are available are not as useful.

It's very important to describe the contents of a data set and to describe the specific kinds of information in each field of a data set that is presented, so that when application developers go to the data set, they know they're finding the right data set with the right kind of data and they know how to interpret the different fields. That's an important part of using the data effectively. In Canada, making our data available in an open data portal first involves producing metadata in both official languages so that app developers can quickly understand what the nature of the data set is and can use it appropriately.

Finally, what is open data? Open data is the practice that takes the raw data and the metadata and makes it available through a portal, as is the case of data.gc.ca. It allows users to search through the portal for the right data sets and allows them to browse and then to download the data in machine-usable formats so they can develop programs and information systems that can manipulate it and produce other uses for it and greater advantages.

The open data movement is quite well developed today. In October 2013, McKinsey Global reported that the potential for open data to generate economic value is significant. This is McKinsey's view. Certainly, through open data efforts in the U.S., in the U.K., in Canada now, and all over the world, we've seen the rise of many, many businesses through the generation of apps that basically use open data and are now widely available through different online stores and so on. Certainly, all of the large consultancies, including Deloitte, speak to the fact that data is the new capital of the global economy, and the ability to harness the vast amounts of data that we do generate is really a large potential for Canada and for society as a whole.

Just to give you a recap of the history, in Canada we have long been aggregators and generators of data. In fact, the concept of open data started around 1995 with the important stores of geophysical and environmental data that we already collect and manipulate through NRCan and Environment Canada.

In California, of course, in the U.S. in 2007, open data started to become an important movement. In fact, in President Obama's first term, there was really the first important national foray, I guess, into open data, with his mandatory policy on the release of open data. The U.S. launch of that direction certainly stimulated open data movements in the U.K. and internationally. Certainly, we watched in Canada and also thought that this was a valuable movement to embrace. It's really a movement that has grown very quickly, and it is, certainly in Canada, stimulated quite a bit by the work by our cities—cities are very active in open data in Canada—as well as by the provinces and by us in the federal government.

Open data is certainly well established internationally. As you may know, the Open Government Partnership, first launched in 2011 by the U.S. and Brazil as co-chairs, was a strong platform for further developments in open data and making governments accountable, open, and responsive to citizens. Similarly, the World Bank has opened its data, knowledge, and research, and is a strong supporter of open data and of all our efforts.

The Open Knowledge Foundation is a civil society organization dedicated to promoting open data and open content. The OECD has

also embraced open data and was present at the 2013 Open Government Partnership conference in the U.K.

● (0855)

Certainly the World Wide Web Foundation is, of course, a strong believer in open data.

Just to give you a capsule of open data in Canada, we're quite pleased with Canada's progress on this front. Four provinces—British Columbia, Alberta, Ontario, and Quebec—have open data portals, as do over 30 cities. In fact, certainly Vancouver was one of the first leaders in open data in Canada and continues to be very dedicated to that, but we're pleased with all of the municipal efforts, including the City of Ottawa, which is also working hard at open data.

Page 8 just contrasts what was it like, how civil society could access the data sets the government created and aggregated and made available, prior to open data. Before open data, the government was already publishing data, but in a different way and in a much smaller and less accessible way. Certainly weather data from Environment Canada has been available for some time, as were maps from Natural Resources Canada.

But what you see on the diagram on the right is one of the fundamental issues of the problem. Each individual department collected and prepared data and made it available on their own individual websites, but not always prominently, often without sufficient, or if you will, standard metadata that described the contents, and not always with appropriate search engines to access it.

So from a user's perspective, it wasn't easy to answer the question of what kind of data is available from the government on topic A or on topic B. The users would routinely have to go through multiple sites, go quite deep into the sites, and then the data was not necessarily in machine-readable format. So while they could visualize it, they couldn't really use it and create an information system.

Finally, an additional issue that users had to tackle at the time was that every individual website made the data available under slightly different licensing terms. The licensing terms are very critical to open data, and the ability to have an open licence that is recognized across Canada, that makes the data available for reuse without restriction on the same terms, is really key.

So that was the situation of data before open data.

Starting in 2009 we started to tackle these questions, and in fact started working on our first view of the licence and the first view of a portal that could potentially make this data available.

Why is open data important for the Government of Canada? Certainly we're strong believers that open data helps to reinforce accountability and the government's agenda. Certainly we are convinced that it does generate economic value for Canadians. It is aligned with our digital strategy, as we are working with our colleagues across government, and it is a key catalyst for innovation and science and technology. We are aligned with our international partners, and the success of CODE, I think, supports the fact that Canadians are equally aligned with it.

Just to highlight the key milestones from a government perspective on open data, in March 2011 the government announced our first open government initiative, and at the time, our first open data portal. That was our first pilot. We launched it with much fewer data sets and with the first version of the licence.

In April 2012 Canada joined the international Open Government Partnership formally. We published our first action plan on open government at that time. The action plan on open government includes, of course, a number of commitments on open data.

In June 2013 the Prime Minister formally adopted the Open Data Charter with other G-8 leaders at the Lough Erne Summit in Northern Ireland.

Just to recap on this part of the presentation, and before we go to the demo of the portal, I'll just say that we have continued to work hard on open data since our joining of the Open Government Partnership.

● (0900)

In fact, this June we launched the second generation open data platform. We now have about 200,000 data sets from 27 departments. We launched with six departments and their data sets. Our search capability is state-of-the-art and we have incorporated social media features onto the site, so we're very pleased with our new portal.

In terms of GC resource management data, the expenditure database was launched in April 2013 to provide Canadians with financial information on departmental spending over the last three years, and we continue to add data sets through all topic areas.

We are working hard right now on a directive on open government, so this will be policy that will help departments and agencies to create a better inventory of their data assets and the information to be published, and provide an implementation timeline for them to achieve this. That will be an important part of our open government action plan commitments, and we're hopeful to see that in the new fiscal year.

Finally, our new open government licence, the second version of which was issued last June, is aligned with a Creative Commons licence. It's plain language. It clearly states the conditions for the reuse of data and aligns with all international best practices.

That's a quick primer on open data. Before we go to the demonstration, would you like to ask any questions?

[Translation]

I am wondering if the committee members would like to ask questions.

The Chair: Are there any questions from committee members before we move to the demonstration of the site?

Mrs. Anne-Marie Day (Charlesbourg—Haute-Saint-Charles, NDP): I have just one question and it concerns a word that is used in the supplementary goal No. 3 for 2015. In French, the word “mappage” is used. What does this mean?

Ms. Corinne Charette: To what page are you referring?

Mrs. Anne-Marie Day: I am looking at the appendix of the G8 Open Data Charter. At the third point for 2015, it reads as follows: “Contribuer à un exercice de mappage de métadonnées du G8”. The word “mappage” is often used in the presentation in French. I have never seen this word before.

Ms. Corinne Charette: It is a word that was used in the translation of the presentation. What it means here is that we need to assign metadata to different data sources and that international standards apply to the metadata that will be used.

● (0905)

[English]

Do you want to add to the concept of the G-8?

Mr. Stephen Walker (Senior Director, Information Management Decision, Chief Information Officer Branch, Treasury Board Secretariat): Further to Corinne's point, although metadata is very important to everybody who thinks that their goal is to provide metadata, especially from the public sector, many jurisdictions have developed their own approach to metadata.

If you're a user of open data or a developer, chances are you're going to want to bring data in from more than one jurisdiction. If we're all using different metadata, it can make that very complicated for the individual user.

Our goal is to work with other jurisdictions to map our metadata against each other in order to be able to provide potential users with easy-to-use tools that will make the data from different places more interoperable, more comparable to each other.

The Chair: Gordon O'Connor.

Hon. Gordon O'Connor (Carleton—Mississippi Mills, CPC): You're using the term “metadata”, and I wonder if you can give an example or two of metadata.

Mr. Stephen Walker: Absolutely.

Let's imagine we have a data set that was just crime statistics. The metadata would provide us with information on who the provider of that data was, so which department; whether or not there was a specific program or service within the Government of Canada that this data was created to support; the date of release of that data; a description of the data so you wouldn't necessarily have to go into the data to find out exactly what it contained, which is important because some of the data sets are very, very large; and the frequency of the data, so how often it is published and renewed.

This is the kind of information that goes into the metadata.

Hon. Gordon O'Connor: I'm not going to divert you into another topic, but there have been issues recently with CSE collecting metadata. What you're describing, basically, are all the details related to a certain subject area? Is that right? Metadata is basically the whole picture—everything.

Mr. Stephen Walker: Metadata would be more of a set of descriptive tags to describe the data, but not the actual data. For example, it would more likely be a set of factors; so time, frequency, title, provider, not the actual data that was held within the data set itself. So it's where it comes from, who it comes from, but not necessarily what the information is.

[Translation]

The Chair: Mr. Trottier, you have the floor.

[English]

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

Madam Charette, you use the terms “open data” and “open government”. I can see examples where open data would have little to do with open government, for example, publishing weather data isn't really about open government. So could you describe the relationship between these two concepts, which in fact are actually two different initiatives even within the Government of Canada, around open government and open data.

Ms. Corinne Charette: To us they're certainly interrelated and in fact our open government action plan has three streams of activity that we committed to in April 2012. One of them is open data, one of them is open information, and one of them is open dialogue. In fact, the open government thrust is about making information and data widely available to Canadians and civil society so that they can use it for their own benefit and derive economic advantage, specifically in the case of open data, foster engagement, and generally contribute to society.

For us, open data is one of the three key streams of activity in our open government action plan. In fact, most of the other countries that are part of the OGP have similar open data streams of activity in their own open government action plan: the U.K., the U.S., and many others. In fact, most of the governments have published action plans that reflect those three streams of activity, because they are separate yet complementary.

[Translation]

The Chair: Mr. Byrne, you have the floor.

[English]

Hon. Gerry Byrne (Humber—St. Barbe—Baie Verte, Lib.): I think Mr. O'Connor touched on something, as did Mr. Trottier, about what is it exactly that we're talking about here? I think there is a presumption when people hear the term “open data”, that it implies the essence of open government. The two, as you clearly defined, are not the same. As I understand it, open data is that which is already available or should be available to the public through normal access to information channels or through that which is already published but is just opaque in the way people can access it and should be made available in a broader, more accessible format. I'll ask you to comment about that if you could.

In addition, there is an assumption about cost recovery of data. The government, especially through Statistics Canada, in particular, has been providing information on a fee-for-service basis, on a cost recovery basis. I'll ask you, will this impact that by eliminating those costs? Does our membership in the G-8 Open Data Charter imply that those costs will now be eliminated and the information made more accessible?

Finally, how far is this going to go? I was perusing your deck a little while ago and it says, to give one example, “What happened to the fish in my lake?”, and it lists a whole lot of environmental, habitat, and other ecosystem observations and reports that would be available. The Department of Fisheries and Oceans has access to databases of who catches fish and how much, but they don't publish those. It's available under the Access to Information Act if you constantly probe and ask them for it.

Is that the extent of this? Is this where this is going? Will there be a PCO or a Treasury Board mandate that says to the Department of Fisheries and Oceans, you have to start publishing this data on a quarterly basis or on an annual basis and do so in a transparent and predictable fashion?

There are a few questions there and I hope you've absorbed them.

● (0910)

Ms. Corinne Charette: I'll try to address your questions the best we can. First, on your question about being available free of charge, that's an important part of our open data portal. Our open data assets are available free of charge. In fact, Statistics Canada had formerly been charging for access to their data and with our work with Statistics Canada on open data, or data.gc.ca, they have eliminated their fees.

Hon. Gerry Byrne: It's across the board now.

Ms. Corinne Charette: All of the data on our portal is available free of charge, and that certainly is an important part of the open data construct internationally and certainly one that we support. So if our data is available, we make it available. Now of course not all of our data is available all at once simply because there is a lot of work by departments on preparing the metadata and preparing these data sets in the format that makes them reusable and easy to understand, compatible with the search tool on the portal, and so on.

So departments are going at it steadily. In fact, in preparation for the Canadian open data experience appathon, departments really came together and made a great effort at preparing, under a very short timeframe, a lot of high value data sets that had been requested by users of the portal. They went ahead and did the metadata and prepared them in the right formats and they were available to the appathon developers. So free is definitely an issue; it's definitely part of the construct.

The next point you question is about privacy, and it's important to note that the open data portal does not present information that is personal in any way. There is no personal information. This is information that is completely impersonal, so to speak, and generally speaking is about topics, but it wouldn't tie a citizen's use to a particular topic or anything like that. We're very concerned of course with the protection of privacy across the government and we work very hard to ensure that our open data assets respect that commitment to the protection of personal information.

The last point that you raised—and I may not have gotten all of your questions—is that what is also important about open data is that, from a federal sense, we have a lot of very valuable data assets, but our assets grow in value when they are combined with data assets at the provincial and potentially at the municipal level. So the federal government only collects data in the realms of jurisdiction that we have programs and services in, and the provinces of course have their own programs and services. They collect information of a slightly different nature. To open data enthusiasts, the greatest value is when they can mash up data that is free and that is under a common licence from all levels of government, and potentially one day internationally as the result of our work on the G8 charter.

● (0915)

Hon. Gerry Byrne: I think I'll clarify that, Madam Charette. I used a poor example because I used the example of a lake, which is a provincial jurisdiction. I was asking about a similar example, however, totally in the federal jurisdiction, which would be offshore marine seacoast fisheries. Just as an example, that would be data that would be totally collected by the federal government and would actually have no provincial government participation. There is information that is collected by the Department of Fisheries and Oceans similar to information that would be collected by Environment Canada and other things.

What my question really is—without muddling the question through the relationship of the federal-provincial jurisdictional issue on that which is strictly a federal issue—will there be a directive that will be given to departments as a result of Canada's participation in this initiative that says if you have data that would normally be collected and disseminated and made available through the Access to Information Act to applicants and you're not publishing it on the portal, you're not publishing it in a way that's transparent and clearly available, you have a responsibility and a mandate to publish it. Is that part of this initiative?

Ms. Corinne Charette: It is. In fact that is one of our commitments to our first action plan that was published, which is the directive. In fact we're working hard on that to essentially give departments guidance in how to conduct inventories of the data they could publish, how to identify what they have already published, and over what time periods they should be publishing the data sets that they're already collecting or working with as a result of different programs and services.

That's certainly part of it, but of course we have to do this in a way that respects departmental resource constraints and in fact their ability to maintain the integrity of the data, frequency of refresh, having the right curators, if you will, available to ensure they can respond to questions, and so on. Absolutely we are working hard in

that regard to give guidance that will require departments to publish more of their data sets.

Hon. Gerry Byrne: That's very helpful. Thank you.

[*Translation*]

The Chair: I will give the floor to Mr. Martin. Then we can move to the second part of the presentation which is about the demonstration.

[*English*]

Mr. Pat Martin (Winnipeg Centre, NDP): Thank you, Chair, and thank you, witnesses.

I think it is useful to take a minute or two now before we proceed to your main presentations so committee members have a better grasp of the scope of what we're trying to achieve. I still have questions. Just as I think I get it, I realize that I don't quite get it. There's more here than meets the eye. You mentioned the meeting in Northern Ireland where the G-8 members committed to the G-8 open government charter. Is that how you phrase it?

My questions are twofold. First of all, the open data initiative, there is great hope throughout the land among the access and privacy communities that maybe this is it. No more will we be frustrated with access to information requests where it takes a thousand days and costs \$10,000 to get a tidbit of information out of the government. It will all be there, and we can simply go and look for it. That's the best case scenario that everyone dreams of if it's true open government.

Who gets to decide what is revealed on the open government portals? Who will do the editorializing? Who will do the redacting, when you black things out? You must have a department of redaction that will be carefully redacting everything that the government doesn't want to release now, surely they didn't want to release then. Who, in your world, does the editing?

Second, Mr. O'Connor hit the nail on the head with his first question, this issue of metadata. I see that the third commitment in the Open Data Charter of the G-8 is to contribute to the G-8 metadata mapping exercise. That's where CSEC comes in. There are 2,200 employees in a building worth \$1.2 billion doing this metadata tracking. The budget and the scale of this initiative, if you include how many thousands of employees you have in Treasury Board who are engaged in this, what is your total budget? Why does it take 2,200 employees to track the emails that Canadians send to each other as per our obligation to this G-8 metadata mapping? Maybe you could expand a bit on what our commitment is, too. Are we involved in an international tracking of what everybody is saying to everybody else in the world? Is that what's costing so much money?

● (0920)

Ms. Corinne Charette: One thing that's important to clarify is that the mapping of metadata exercise referred to in the G-8 charter is about standardizing the type of metadata. It has nothing to do with what CSEC does or information that they do or don't collect.

It is about defining standards for describing data sets internationally. For instance, if we were going to define standards for how we would describe a lake and the geospatial qualities of a lake, all jurisdictions would describe those geospatial properties with the following four, five, or six attributes and these would be understood internationally. If we were mashing up data about a lake from data sets in the U.K., the U.S., and Canada, we would be working with the same kind of data but from individual data sets with information from different jurisdictions.

Mr. Pat Martin: Let me get that straight then, you're not mapping my correspondence to Mr. O'Connor during the night, after hours. The things that I say to him back and forth regularly is not what you're trying to track.

Ms. Corinne Charette: We're not tracking anything. In fact, metadata and this mapping exercise—and you raise a good point that we'll have to be a lot crisper with our terminology, so thank you—are really about aligning the data standards for how we describe these data sets internationally and ensuring that developers in civil society can mash up open data sets internationally because that's really their value. Be it in health, in environment, in economic terms, or whatever, do we describe this kind of data the same way? Because that is key. That's number one.

Number two, in terms of the number of employees we have doing this at Treasury Board Secretariat, I think it's about 10. That's the extent of our department, but I can guarantee you that we are not over-resourced to do so, and we work with colleagues across the departments and agencies in the federal government who, a lot of times, are not dedicated open data specialists. Quite the contrary, they're working in program areas, and they are doing this because they are equally committed to open data.

In the last point you raised, you highlighted the difference between open data and open information. Open data is not about access to information, although it certainly is in support of accountability and transparency and trust in government. Open data is about making the data that we do collect, this machine-readable fundamental data, available. Open information is about publishing documents or reports online, which would eliminate the need for an access request for those documents or reports.

For instance, right now, and also as part of our open government action plan commitments, we committed to publishing searchable summaries of ATI requests. In fact, ATI request summaries have been available now for over a year, I believe, since we started posting the summaries. People can go to the data sets and read through and select some of them and so on. Open information is another stream of activity in the open government action plan, but it is not the open data.

Open data is really about this great machine-reusable data. In the past, you could have received a data set through an access request; however, you certainly don't need an access request. You can go online anywhere in the world, search through the catalogue as we'll shortly see, and download it to your own computer or PC or your own CD, and work with it as you like.

•(0925)

[*Translation*]

The Chair: I will give you the time to give your presentation. It will then be possible for committee members to ask other questions.

[*English*]

Mr. Stephen Walker: I'm going to start with the page that you can see up on the screen now. This is the home page for data.gc.ca. This is our one stop shop for all of the open data that the Government of Canada makes available at any time to citizens, researchers, voluntary organizations, the private sector, the media. It maintains access and discovery of all open data. It also has some other open government activities, but I'm going to focus mostly on the open data.

The page has been designed with large tiles, as you can see, so that you can easily find what all the features on the site are, so that you can quickly jump to the information that you're looking for. Key, I think, for our conversation today is that tile in the top left-hand corner, the search data.

I'm going to proceed as if I was an average open data user. I click on “search data” and I'm going to pretend that I'm somebody who's looking to buy a new house and I'm interested in what the safety is in the neighbourhood that I'm considering buying a house in. I would type in crime, for example, hit submit, and all the data that is currently made available from the Government of Canada related to crime would come up.

Oftentimes there is a lot of data. We need to be able to help the user filter down those results to a smaller amount, so that they can find what they're looking for a little bit faster. I could reorder the data sets alphabetically or by the date that they were actually created or last modified, and their relevance. I'm going to leave it at relevance for now.

The left-hand side provides a whole variety of filters that can be used to narrow down the search results. I'm going to keep proceeding as if I'm looking for neighbourhood safety information, so I'm going down and see that under subject there's law. I click on law. The number of data sets comes down, still a fair amount, but as I move down the data sets looking for the information I'm looking for, I see crime statistics for Canada, the provinces and the territories, and I know that that's the information that I'm looking for.

I click on that data set and this is the metadata. This slide provides all of the information that we have on that data set, who the publisher is—in this case it's Statistics Canada—what subject it falls under, the date it was last published, and a short description and title of the data set. Those for us are the mandatory metadata fields that must be completed by any department or agency that is making data available.

Before I go into the specific information related to this data set, because I want to point out some of the features for each data set, I want to highlight the licence, which is right here. This licence is a significant point of progress for us working within open data.

As Corinne mentioned, it used to be that individual departments would make their information and data available under a variety of licences, most of which were several pages long and written in legalese that was very difficult to understand by the layman and oftentimes there were specific restrictions on the reuse of that data. For example, most often you couldn't reuse that data commercially.

Our new licence is written in plain language. It's extremely simple. It's based on best practices for open licensing internationally. We are sharing this licence with other jurisdictions within Canada, promoting adoption on a pan-Canadian level, so that data users will be able to bring data together from multiple jurisdictions within Canada at both the provincial and the jurisdictional level, and that can be combined and mashed together.

I'll just go back to the data set and point out a couple more features.

- (0930)

Back before we launched the most recent version of data.gc.ca this past summer, we held a series of round tables across the country with the open data community to hear what it was they would most like to see in the revised, revamped open data portal.

They wanted the ability to rate the data sets themselves and tell us what they thought of the data. They wanted to be able to provide individual comments on that data in the hopes that we could perhaps improve that data. They wanted to be able to share the data easily with others.

All of those features have been incorporated into the new data.gc.ca. You'll see up here on the right side that you can rate the data. It's a five maple-leaf scale. You would simply pick the rating that you're giving the data. You can provide individual comments below and then submit, and it becomes part of the ongoing consumer rating of that data. You can share the data via Facebook, Google, or Twitter, and you can provide comment on the data and share those comments with all other users of that data.

If I want to download the data, I simply click on one of these buttons. The data is made available in different file formats to ensure flexibility of use by the individual users. The data sets are made available in French and English, and there is additional supporting documentation to help the users use the data, and all of these are one-button downloads. Press the button and the data downloads—I won't do that right now.

I should just mention before I leave this page that there's an openness rating down at the bottom. We've incorporated the use of an international openness scale that's used by other jurisdictions to indicate the level of openness of the data sets. It's based on the five-star scale. Most of the data that we hold is three stars and above. This speaks to whether or not the data is being made available in a well-structured format, whether or not you require proprietary software in order to be able to open the data instead of an open software program, and we, the U.S., the U.K., and a variety of other jurisdictions, are using this scale.

If I go back to the search page and pretend that I wasn't able to find the data set that I was looking for, there's a button that says "Can't find what you're looking for?" up at the top. We're very keen to get feedback and information from potential open data users on what data they'd like to see that we haven't yet made available. That helps us to prioritize our work, by working with individual departments to have that data made available. If you didn't have the data and you clicked on that, you'd see a variety of data sets that have already been requested. So you'd look at that first to see if the data that you're looking for has already been requested, and if you see it, you could add your voice to those who have already requested that data.

Behind the scenes what we do is we work with this list and individual departments to find that data and to try to make that data available, and then we update here, on this page, when we've been able to make that data available.

So, for example, here, with the national household survey released in May 2013, now when you click on that it would take you to the actual data set. If you couldn't find the data that you're looking for, you could submit a new data set and it would become part of this list, and again, other individuals would be able to come in after you and add their thumbs-up or their support for getting that data set as well.

Now I'll go back to the home page to show you a couple of other features specific to open data. I'll start with the showcase. We use this area of the site to provide examples and illustrate the use and the utility of open data. We keep a whole section called open data in action, which provides information on specific projects within the Government of Canada, most of which are collaborative, working with other jurisdictions, that use open data specifically to inform a particular policy area.

- (0935)

The oil sands monitoring portal is a joint initiative between Environment Canada here within the federal government and the Alberta government. It specifically focuses on open data. Together the two jurisdictions make more open data available out to the academic world to support greater research.

Also available through the showcase is an apps gallery, which provides access to a comprehensive listing of the apps that have been made available and developed by the Government of Canada using open data. These apps are downloadable for mobile devices.

If I click on the left-hand side, for example, to find a specific app for my phone, I can click on "mobile" and see all of the apps that are currently made available by the Government of Canada for download into mobile phones. "Recalls and safety alerts", for example, is an app using open data that can be downloaded. I can download the app straight from the site.

I'll give you just a couple more features of the site. About data.gc.ca, I want to point out that a variety of the information resources that are put on this site are for the open data layperson, designed to get them interested in open data and explain what can be done.

We talk about the licence, making it clear that the data that is available can be released and reused on an unrestricted basis. There's a section on frequently asked questions. There's also "Open Data 101", a handbook on open data to get people who have yet to start really using open data off the ground with the basics of what open data is, how it can be used, and how to work with open data.

At the other end of the scale, the site has a developers' corner, which is really for the open data user who has some experience already. These are potential developers, for the most part, people who are interested in building applications using federal data or federal data combined with data from either the private sector or other public sector jurisdictions.

Here we have a little bit more sophisticated information around working with data sets; how to use an application programming interface, a software tool that makes access to data that changes frequently within the federal government more easy to use if you're building an app that will want to access that data on an ongoing basis; information about our metadata element set; and then Open Data 101.

That really brings me to the end of the tour.

Ms. Corinne Charette: With that, before we embark on the third and concluding part of our presentation, are there any questions about the demo?

[Translation]

The Chair: Ms. Ablonczy, do you have any specific questions about this presentation?

[English]

Hon. Diane Ablonczy (Calgary—Nose Hill, CPC): Yes.

You mentioned on page 15 that since the launch of the open data portal in June of 2013, there have been 88,000 downloads.

I'm wondering if you triaged the kinds of downloads that have been made and if you can tell us what kind of data is in hottest demand.

Mr. Stephen Walker: I should have shown this to you before, because there is a page that provides information on the current top 25 downloaded data sets. There's also information on how much traffic the site is getting on a month-to-month basis, the total number of departments that are contributing data sets, and how many data sets from each department, and so forth.

This page is updated on a monthly basis because it continues to change. Data sets that are very popular can move up and down that scale from month to month.

• (0940)

Hon. Diane Ablonczy: Thank you.

[Translation]

The Chair: Ms. Day, you have the floor.

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

I have two supplementary questions. First of all, I checked the word "mappage" and it does indeed exist in French. It refers to cartography, or *mapping* in English. I did not know this word, but I will try to use it in the future.

Do you have connections with the private sector, universities and research centres? Are links created with these groups in order to constantly improve the open data file?

Secondly, if an average person types a word into a search engine, such as Google or Bing, will this appear in their choices to obtain the information that they need? Obviously, if it is Statistics Canada, it is a given. If the words are "delinquency", "situation", "environment" or any other word for which one might wish to obtain data, is the site well placed to give this information and can it be accessed easily?

Ms. Corinne Charette: As for links with the private sector, we do not have any formal links as such. We were happy to see the Open Data Institute, which is found in Canada, the United Kingdom, and the United States, mentioned in the last budget. This is a growing movement. It is a non-profit that essentially brings together the private and academic sectors as well as all stakeholders who are interested in working in a concerted way throughout Canada.

So we will have our Open Data Institute located in Canada. It will be created by this organization that will bring people together. We will work with them in cooperation with the provinces and municipalities. They also have links with the academic sector and with other sectors.

As for the private sector, we do not necessarily have direct links. These are more indirect. Private sector businesses support this movement. To that effect, there are also certain enterprises such as ESREA. It is a longstanding organization that promotes tools and software that allow the best possible use of scientific data. So this is really informal cooperation, but across a wide range of sectors to provide in the best way possible more open data for everyone's benefit.

As for your question about Google, Mr. Latour can answer it.

Mr. Sylvain Latour (Director, Open Government Secretariat , Treasury Board Secretariat): I can answer this part of the question.

Yes, we ensure that all information contained in the portal and everything that can be directly searched is also available for search engines like Bing and Google. That means it is possible to get exactly the same results from the Google web page by using the same keywords.

When we talk about mapping with other governments, the advantage is that it makes it easier for Google to find data from Canada and also from the United States and England when making a search. By using the same metadata, we are making it possible for these private search engines to search on a level playing field. This increases our ability to make information available to citizens.

The Chair: Thank you.

Mr. Trottier, you have the floor.

[English]

Mr. Bernard Trottier: Thank you, Mr. Chair.

I appreciate that all the data on the open data portal is provided free of charge. I was looking at the app gallery. Presumably all the applications there are also provided free of charge. Are any mechanisms whereby, say, a private application developer develops a great app, the government buys that application, and then in turn provides that to citizens free of charge?

Mr. Stephen Walker: It hasn't happened yet, but I suppose it is possible. Our biggest wish is that the data made available by data.gc.ca will trigger new development external to government and that some of those tools, some of those new apps, will be useful to citizens. Most of the time, I think they'll want to distribute those themselves, but I certainly will be watching carefully, as we will be with the results of the CODE contest that just finished this week. We expect to be looking at over 100 new apps. I think one of the things that we're really interested in is whether or not any of those apps would be good for use by the federal government.

Mr. Bernard Trottier: If I can just add to that question, if there were a great private application developed so that there could be a transaction between a user of the app and the app developer, are you allowed to promote that application on the website? Really, you'd be promoting a private interest.

• (0945)

Ms. Corinne Charette: The apps on our website are free apps. Developers that create apps using our data and anybody else's data are free to decide to sell those apps and put those apps on the Apple store or any other app gallery out there and actually charge money for them, but we would not necessarily promote those or put them on our app gallery.

Mr. Stephen Walker: No, not yet for sure.

Mr. Bernard Trottier: Thank you.

[Translation]

The Chair: Mr. Martin, you have the floor.

[English]

Mr. Pat Martin: I appreciate all of this, and it's helping to give some definition as to what type of information will be put up for open.... If the default is to be openness, that's an important directive. Currently the default seems to be secrecy. It's like pulling teeth to get sensitive information out of the government through the access to information regime. Even though you say this is not set up to replace or to do the job of ATI, you mention in the opening page of your website, data.gc.ca, that this is really an extension of the spirit, if not the letter, of the Access to Information Act.

I'm still suspicious and I'm still interested, but you didn't answer my question as to who screens. Who ultimately gets to say what goes up and what stays down in terms of the portal? Is it the minister, is it the government of the day, or is there some overarching, independent authority, such as the Information Commissioner, who says that cutting the hair of Afghan detainees should be public information and should go up on the portal, and that you shouldn't have to wait a thousand days and go to court to find out whether or not you cut the hair of the Afghan detainees?

Who is your boss who says what goes up and what does not go up?

Ms. Corinne Charette: Well, there are two questions there.

First, in terms of what data sets are made available for publication, the first thing that's important is that what's on the open data portal is not classified information. Clearly it is not information that is classified secret or confidential. It is information that is public, so that the data sets themselves should be made freely available.

Now, within departments there is a wealth of data and of data sets that are not yet published. While we're very proud of the 200,000 data sets we have on the portal today, we believe there are a lot more.... That is the focus of the directive, to require that departments conduct an inventory of the data sets and then prioritize the data sets with our help. We work with them to identify high-value data sets and to help them make them available.

Ultimately, the departments have to be in a position to maintain and assure the integrity of the data and ensure that the data they are promoting—

Mr. Pat Martin: Who is going to catch something?

Let's say that the government department thinks it should be released. We've had cases where ministers' assistants dive into dumpsters or into postal boxes to retrieve an access to information request that the minister decided they had better not release.

Who is supervising what gets put up and released and what is held back under whatever justification the minister might see fit?

Before you answer that, the final question I have is this. What's to prevent the data mining of the users of your new open government portal? We were all horrified to learn that during the interactive tracking of the Aga Khan, whoever signed on is now part of a Conservative Party fundraising mailing list or something.

Who's to stop the data mining of the information of people who use your service now? There are going to be patterns developing. Is the privacy of Canadians who use this portal going to be protected and shielded so that it doesn't get harvested into the SIMS database?

Ms. Corinne Charette: To answer your last question first, the use and the downloading of these data sets is totally anonymous. That is part of the privacy notice on the site. There are no accounts. You don't have to create an account to download a data set; you can download it, and we do not track that or keep that information. We're very vigilant about that.

In fact, we also have privacy requirements when we get feedback, such that we can only keep the information of the person requesting feedback for enough time to reply, obviously, to the person who may request some information. We have very strict privacy requirements, and the use of the data is totally anonymous.

That is to deal with your last question.

To distinguish between access to information and open data, departments ultimately will determine the speed and the schedule of publishing data sets. This is not access to information. Access to information will still go on; Canadians may still pose an access request. Most access requests are not about machine-readable data. They're usually about reports, documents, memos, etc. Access to information will continue.

We have, in our open government action plan, a few other commitments that support the open information stream of the open government action plan. In particular, since the first version of the plan, Library and Archives Canada have moved on our commitment to declassify and make available—I don't have the exact statistics—almost 10 million pages of information that was formerly classified and not available. That information is now freely available to Canadians through their website. That is in the stream of open information, in addition to the searchable summaries of ATI requests.

We also have another commitment, working towards a virtual library that will allow us to publish more information online in order to avoid unnecessary access to information requests. We're working through these commitments one by one.

● (0950)

[Translation]

The Chair: I will now give the floor to Mr. Adler.

[English]

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

I have two questions, but I want to preface my questions by making a couple of comments.

First, Mr. Martin talked about ministerial assistants jumping into garbage bins. If he has any evidence of that, I'd like him to present it to the committee.

Second, it is our government that has been the most transparent in Canadian history, and the numbers demonstrate that in and of themselves.

Mr. Martin is playing fast and loose with the figures, but I'd be very interested to know details of any ministerial assistants jumping into dumpsters. If he could table that with the committee, it would be very helpful.

Mr. Pat Martin: I'd be happy to, Mark. We'll get that information over to you.

Mr. Mark Adler: You mentioned earlier, Mr. Walker, round tables that you held. There is no need to be specific in terms of who was around these tables, but I'd be curious to know what types of companies were involved in these discussions.

Mr. Stephen Walker: Our approach was to work with municipal governments in a variety of individual cities.

In Edmonton for example, we would work with the office of the city clerk. In Edmonton they have quite a robust open data initiative of their own. We asked them, if they would, to help us contact local users—local companies, the post-secondary institutions, civil society organizations, and government representatives as well. We created

the widest possible invitation list we could and sent it out and invited them all to come for a day.

Mr. Mark Adler: These would be both public sector and private sector institutions. Thank you.

My second question is about the data sets themselves. How are they organized? For example, if I were looking at crime statistics, is there a way of determining by geography, for example, if I'm interested in just a specific geographic area? Would it be possible to mine the data for that kind of information, on either crime statistics or any other kind of data sets, or are there just broad statistics that you have to take with a national perspective?

Mr. Stephen Walker: There is not one answer to that question.

For a significantly large number of data sets, including the one that we're talking about, there is geo-specific information included into the data set, so that you are able to find which statistics are being applied to which particular area of Canada. We can do that for all of the data sets for which there is some kind of geolocating reference within the data set, which is a large number.

For StatsCan specifically, for example, and for Environment Canada and NRCan, most of this information has some kind of geo-specifying element. So yes, you could map the data across Canada.

● (0955)

Mr. Mark Adler: So it is possible. That's interesting. Thank you so much.

The Chair: Mr. O'Connor.

Hon. Gordon O'Connor: Thank you very much.

I know these sites are called open data and they're open sites. However, I think also about the security side. Are these sites protected in some security way from hacking? You can have a hacker get in there—we have histories of it—and provide a whole bunch of data that is not open data and it's out into the public. Is there anyway of protecting our system against this kind of thing?

Then I have a second question.

Mr. Stephen Walker: I think most of the Government of Canada external facing sites are subject to the same protection regime. The protection around data.gc.ca might be a little bit higher than other sites, but it is maintained outside of the firewall, not within the firewall of the Government of Canada. We've been up and running for three years now through the various versions of data.gc.ca and there have been no incidents so far. As Corrine mentioned, we just ran the CODE hackathon over the weekend and we put in place specific protection regime in order to bump it up a little bit, and again there were no incidents.

Hon. Gordon O'Connor: My other question is about Crown corporations. Are Crown corporations involved in this activity? If they aren't, why aren't they?

Mr. Stephen Walker: They're included in the request. We have gone out to Crown corporations specifically when a data set has been requested from an external user. Our job is to see whether or not we can actually get that data. We've also sent out general requests for data out to Crown corporations but they won't be covered under the directive on open government.

Hon. Gordon O'Connor: What you're saying in effect is that you don't have anything from Crown corporations.

Mr. Stephen Walker: We have a few data sets from Crown corporations now, but we also could conceivably make data available that we have not yet received from Crown corporations.

[*Translation*]

The Chair: Thank you.

Would you like to add something, Ms. Charette?

[*English*]

Ms. Corinne Charette: I would say that Crown corporations are quite collaborative. CMHC has reached out to us and is interested. We haven't put a focus on getting data sets from Crown corporations yet, but certainly that is something that we will work on over this fiscal year to try to expand beyond the government departments and agencies on schedules I and II that are the focus of our information management policy, and directly the focus of the directive. But the Crown corporations are by and large very collaborative as they see value in making their data available to Canadians as well.

Hon. Gordon O'Connor: Just to respond to that, it would seem to me that Crown corporations are more likely than departments to be the sources of the information that the public wants. They have particular functions that relate to the public and it would seem to me that they should be at the forefront of this information.

[*Translation*]

The Chair: Thank you for your comments.

Before we move to the third part of the presentation, I will give the floor to Ms. Day.

Mrs. Anne-Marie Day: I found Mr. Martin's question interesting and relevant. So, I had some fun. I took my iPad, I went on Google and I searched for the words "airports Russia". We know that Russia is a member of the G8 and is often present on the websites. So, by searching, I found airports, a diagram indicating where they are located, a map and colour pictures. There is everything one could want. Basically, there was a lot of information.

I went back to data.gc.ca and I searched for the keywords "airports Russia". It gave me "Haven't found what you are looking for?" and the words "relevance" and "search".

So, with countries that are part of the G8, be they France, Germany or another country, are there links to their data that we can get? If these links are available on Google, why are the hyperlinks not available? Will we have them later? Is it necessary to go through the structure as presented, by entering, for example, the following request: "I am looking for available water quality indicators"?

So it is basically the same question. You told me earlier that if I search in a search engine, like Google or Bing, for something like water, it will bring me to the site. Is the opposite true? If I search for words in the file in question, will it take me to external links that contain the information that I want?

•(1000)

Mr. Sylvain Latour: At the moment, the answer is no. The Government of Canada portal can only find data from the Government of Canada. However, through our international

collaborations and through our national cooperation with the provinces and municipalities, we have discovered that this is something that users want. They have an interest in that.

As such, in the coming years, we will work to create as many links as possible between the different portals as well as between the different access points in order to permit relatively transparent navigation between these sites. At the moment, these sites are grouped by jurisdiction. Currently, our portal only gives access to Government of Canada data.

Mrs. Anne-Marie Day: Are there agreements with the G8?

Mr. Sylvain Latour: The current agreement with the G8 aims to standardize data and metadata in such a way as to facilitate that sort of thing, but there is no agreement or commitment to create transparent links and navigation between the participating countries' data.

Mrs. Anne-Marie Day: Mr. Martin asked a good question. He wanted to know who decides which data is included on this site. This site is of great value and I don't want us to forget that. It is truly wonderful to have this information, which is very relevant. For example, if one were to deny that climate change exists and thus decided not to put data on that on the site, we are stuck with that. There are no other links, but you have said that that will come.

Thank you.

The Chair: Ms. Charette, you have the floor for the third part of your presentation.

Ms. Corinne Charette: Thank you, Mr. Chair.

[*English*]

With that we'll go on to the last part of the presentation. We'll jump to page 18 to just give you a wrap-up of the specific initiatives we're working on with our federal government colleagues.

There are three important international initiatives. We've talked a bit about the G-8, the Open Government Partnership, and the International Aid Transparency Initiative.

The Open Government Partnership was by far the first international initiative that we participated in with our signing of the charter and joining the partnership. Then in June there was the G-8. What's to note is that the U.K. was the co-chair of the Open Government Partnership in 2013 and was also, of course, the host for the G-8 summit. They are very strong open data enthusiasts, and we saw the work on the charter, which was quite fruitful.

On the IATI initiative, CIDA has been working with this international initiative, and we're very happy to showcase their work in this regard as part of the open data portal. These are the topics, on page 18, that internationally are of the greatest value, so in fact the sharing of scientific data is very important on an international front.

I'm very pleased to say that Canada is amongst the leaders on open data internationally. We are working on the open data working group. It's co-chaired by us. It was launched in October 2013 by the OGP to support open data work by all OGP member countries. I don't have the exact count of OGP member countries today, but there are over 60 of them. Not all of them are as advanced in their open data efforts as Canada, the U.K., and the U.S. are. Some of them are smaller developing economies and so on, so they benefit quite a bit from our work with them in this regard. The work is focused on the four themes.

As part of the G-8 open data action plan, we have a sub-action plan for the Open Data Charter that was published in February 2014. It commits us to the proactive release of high-value data sets identified by the G-8 countries. We're very pleased that most of those high-value data sets identified by the G-8 were already published by Canada, and we continue to work down the list of the high-value data sets identified by the G-8 to close that gap and work with civil society.

As a result of strong support and sponsorship by Minister Clement, we continue our work on open data. The Canadian Open Data Experience, we are hoping to repeat on an annual basis, and we will certainly also collaborate with our provincial and municipal colleagues whenever we can on their own appathons. We continue to make progress and achieve the commitments we stated in action plan number one and are currently working on the second version of our action plan on open government, which is a requirement of our membership in the OGP. That will be published towards the fall.

Open data in Canada is an important, ambitious commitment. As part of our work on action plan number one, we were asked to make an ambitious commitment for open data in Canada. What we proposed was that we work to get to common licensing terms across Canada in all provincial, territorial, and municipal jurisdictions on the use of open data. We already have four provinces and a number of municipalities. Quebec and its municipalities also announced their own initiatives, in February in fact, in the same form of licence as ours, so they're now in the same fold. We continue to work with the remaining provinces and territories and municipalities to make open data in Canada a reality, which will allow us to really mash up data with the same licensing terms across all jurisdictions in Canada.

The work of the Open Data Institute has not yet begun as it was just announced, but we are working with the founders and will be strong collaborators for their work to ensure that we, as the federal government, benefit and that all citizens benefit in this.

We will continue to chair the OGP open data working group, as well as continuing our work on the G-8.

I talked about CODE, so I will jump over that.

- (1005)

We are in the process of developing, with input from civil society, action plan number two for the Open Government Partnership. We are going to focus, in the second version of the action plan, on implementing our directive on open government during the upcoming fiscal year, certainly accelerating the release of other high-value data sets and more appathons.

If we're able, over the next fiscal year we will probably come to a third version of our portal. It is an open source portal. It is supported by work that is done internationally and with our partners in the U.S. and India. We're the source of the portal but we're also working internationally with the U.K., I think, on greater search capabilities and other features in the portal.

Of course we'll continue our work on international standards for data and metadata and the interoperability of the data sets between jurisdictions.

On page 23 there's just a bit more description of what our ambitious commitment to open data in Canada is all about. But again, it's about harmonizing our work across Canada, and it would be quite significant for us to achieve that. It would certainly be very useful to civil society to be able to tap data sets across Canada without any challenge to usability or licensing. We're certainly looking for a federated discovery approach that will allow a Canadian to perhaps enter through B.C.'s open data portal or Ontario's open data portal and find the right data, be it from Ontario, B.C., or the federal government, or any municipality, and work with that.

This movement has seen a very strong collaborative philosophy. It's really the basis of open data. It's all about collaboration among all levels of government and civil society and the private sector. It's really very exciting, and we're very pleased with our results.

We talked about the Open Data Institute, and we will look forward to its progress. It does have partners including OpenText, Communitex, and the University of Waterloo, so it certainly includes the private sector as well as the academic sector in developing our work, similar to what has been done in Britain and in the U.S.

On page 25 there is a little bit of work on our OGP open data working group. We've already shared our open data experience with a number of OECD countries. I think we shared it with Colombia, as I recall, and we're also sharing with Mexico. In fact Canada is quite determined to continue to collaborate in this regard. On page 26 I've mentioned this.

With that I think we've touched on all the key points about our work on open data. I certainly hope that we have advanced your understanding of the key concepts and the value of this, and of our efforts and our achievements from a federal government perspective.

With that, we're happy to answer any further questions you may have.

•(1010)

[*Translation*]

The Chair: Thank you. I will now turn the floor over to committee members. We will follow the usual speaking list for committee members. We will start with Mr. Martin. You have five minutes.

[*English*]

Mr. Pat Martin: I didn't realize we were going to formal rounds, but I'm happy to perhaps pick up where I left off.

I'm still interested—and I don't think we've received a concrete answer—in what this is costing us. What is the total cost to date invested or associated with implementing the federal government's open government plan?

I've already cited how it compares—or how we're kind of shocked at the amount of money we're spending on spying. I would hope that we're spending a corresponding amount of money on the inverse, in openly and freely sharing information. The public has a right to know what their government is doing with their money.

So if we have 2,200 employees down at CSEC, with a building alone worth \$1.2 billion; and you have 10 employees at Treasury Board Secretariat, what is the corresponding budget in this open government initiative?

Ms. Corinne Charette: Well, the reality is that it's a very small budget and a very efficient operation, but I'll explain why.

Open data and our open government work is really about leveraging the assets and the work of all government departments and agencies. For instance, Environment Canada has been publishing weather data for many years and has the infrastructure, the program, policy support, all the public servants who work at Environment on weather data, and so on. As does NRCan, with the geospatial survey, or the—

Mr. Pat Martin: You're not doing the original research. You're just pulling it all together.

But I only have five minutes, so could you give me a dollar figure for what you've spent to date and what the project cost is to December 2015, when you expect to be completed?

Ms. Corinne Charette: We don't really measure it that way, because it's not a specific program. It's not an official program. What we have is a Treasury Board cost, which I would say has been, in this fiscal year, less than \$1.5 million to cover employees and the other expenses we have. We foresee probably similar expense levels over the next five years from a Treasury Board Secretariat perspective.

Departments and agencies do not segregate their costs on open government because they're part of their program and policy costs, so we don't dissect them and we don't roll it up. It's not a formal program. It's really leveraging the assets, knowledge, and information that are available, but leveraging it smartly.

Mr. Pat Martin: Okay. Thank you. It's about \$1.5 million.

How much money have you allocated to advertising to promote open government and the open data portal, data.gc.ca? What's your advertising budget?

Ms. Corinne Charette: To my knowledge, we have not had any advertising. We promote it online. We promote it through consultations—

A voice: Search engines.

Ms. Corinne Charette: —and search engines, but certainly we're not aware of an advertising budget.

Mr. Pat Martin: Okay.

On the international component, when you developed the Open Data Charter, that was just at the 2013 G-8 summit. Is that what I'm to understand? That's how recent it is.

The United States was on board. That was one of Obama's first announcements, I think, when he was elected, wasn't it? How do our efforts compare to those of the United States? How far along are they? Have they finished their program? Are they in full disclosure of all they're going to disclose, or are they still putting it together, the way we are?

•(1015)

Mr. Stephen Walker: Although the G-8 charter is from just 2013, we and U.S. both have been involved in developing open data for a bit longer than that. The G-8 charter had a couple of more specific commitments for G-8 members over the course of the next couple of years.

In general, the U.S. had an early start on us. They definitely had a head start. We've pretty much caught up now. In a system that is even more complex than ours, they have a lot of work to do in order to be able to dig the data out of individual departments and agencies. They have a lot more of them.

I would say that the early leaders were the U.K. and the U.S. Now we would be considered up among those leaders.

Mr. Pat Martin: Do you expect Russia to be as compliant as the U.K., the United States, and Canada as far as full disclosure, etc., goes? What do you have to promise to be able to belong to this club?

Mr. Stephen Walker: It's difficult to say. They have just barely started. We've been at it for two or three years already. They have a ways to go to catch up.

Mr. Pat Martin: Thank you.

[*Translation*]

The Chair: Thank you, Mr. Martin.

Mr. Trottier, you now have the floor. You also have five minutes.

[*English*]

Mr. Bernard Trottier: Thank you, Mr. Chair.

One of the things we're going to be doing with this study is summoning certain witnesses who have some information and insight to add to our study. Hopefully, we can then assist you with your planning.

I'm trying to get a sense of the various departments. You mentioned that they're doing certain things. The reason you're small is that other departments within the Government of Canada are doing some of the effort. Are there certain departments within the Government of Canada that are well advanced in providing open data? Is it worth bringing them in as witnesses for them to describe...?

I'll give you one example. There are certain data that I'm very interested in that have to do with immigration settlement patterns. In Toronto, those patterns are shifting over time, and when it comes to providing settlement services, it's useful to know where the immigrants are going. It's more frequent than every five years of a census. Every year or every few months, these patterns change.

For a department like Citizenship and Immigration, would it be worthwhile to bring them in as witnesses or are they just getting their feet wet right now with respect to this initiative?

Mr. Stephen Walker: No. They're well on their way. In fact, the CIC data sets are among the most popular metadata available on data.gc.ca. If you look at the list that I showed you before, the top 25, you see that eight or nine of the top 20 are CIC, and most of them are related to immigration statistics in one way or another.

There are certain departments that are way ahead because they've been doing it longer or because the demand has been greater and they've had to get their feet wet a lot faster. Statistics Canada is an expert. NRCan is an expert. Environment Canada has a lot of expertise as well, as does CIC. The other major actor and contributor would be Agriculture Canada. Although there are almost 40 departments contributing now, those five departments are probably the federal leaders.

Mr. Bernard Trottier: Among other levels of government, you mentioned Alberta, B.C., Quebec, and Ontario. What about municipal governments? Are there witnesses that you would recommend that we talk to, to gain some insight?

Mr. Stephen Walker: Certainly. A group of five municipalities have been working together for longer than we've been working with the provincial government to share best practices on the delivery of open data: Vancouver, Edmonton, Ottawa, Montreal, and Toronto. They're all leaders in this area.

Mr. Bernard Trottier: Thank you.

Those are all the questions I had. I know Mr. O'Connor had some questions.

Hon. Gordon O'Connor: Yes. What is clearly not open data?

Mr. Stephen Walker: We've developed specific standardized criteria that departments must apply to data sets before they can contribute them to the open data initiative, to be able to identify data sets that cannot be distributed because of personal information, security information, privacy, confidentiality. We must own the data. If it's third-party data then we don't make it available as open data.

Hon. Gordon O'Connor: Okay. You talk about how this open data is going to create a larger economy, etc. What proof do you have that this will contribute to improving our economy?

Ms. Corinne Charette: We don't have direct proof, but we do have, if you will, indirect confirmation. There was a report by one

institute yesterday about the apps economy and certainly in the U.S. and in the U.K. there are multiple media references to this.

We're not tracking the economic result of open data. It's very difficult for us to do that. We're still wrestling with performance indicators that we would use to consistently measure the success for our program, as are the provinces and the municipalities. It is very difficult to make the link. However, the easiest link is the whole notion of GPS, which has made quite an economic impact, and the whole industry of technology based on government data.

• (1020)

Hon. Gordon O'Connor: If you can't prove there's any economic benefit, I don't think you should be touting an economic benefit until it can be proven.

Mr. Stephen Walker: For Canada it's a little more of an extrapolation. We can certainly point to the 15 different companies that are using open data to make money on their bottom line.

In the States and in the U.K. that information, that analysis, has gone a lot farther because they've been at it for a longer period than we have.

A list is about to come out south of the border, much like the Forbes list of the top 500 companies involved in open data. We're interested in getting a chance to look at that, so that we might be able to extract that same type of approach in Canada.

[Translation]

The Chair: Thank you, Mr. O'Connor.

Ms. Day, you now have the floor. You have five minutes.

Mrs. Anne-Marie Day: Thank you, Mr. Chair. I have a few questions to ask.

In my opinion, and going back to what Mr. O'Connor was saying, a good example is cartography in the High Arctic. I believe that the cartography for that region is on the website in question and that it would be enormously useful to companies that wish to do business, move forward, or develop something in that area. This makes it possible, among other things, to know where deposits are and where local population groups live. This also makes it possible to know the environment, the ground plan and the whole area, which means that if someone wants to invest, thanks to this open data, they have a great deal of information at their disposal. That is one of the goals of this study, to highlight the impact of information on economic growth.

Now, in terms of accountability, how do you check this? Is it done annually or biannually? How do you know if you are heading in the right direction, that you are respecting the five initial goals and the eight others that will be added in the context of the G8? What do you analyze to determine if you have the right goal relative to what you should meet and if you are on the right track to meet the goals set by the G8?

[English]

Mr. Stephen Walker: We have two. We have mandatory reporting requirements associated with the OGP and with the G-8. In the OGP it's an annual reporting process to define exactly what progress we've made on our specific commitments. That reporting process is templated. It's standardized internationally around the globe and requires specific public consultations to be able to feed into the development of the report. Then the report is subject to third-party review.

The G-8 is a little less mature at this point because just this past summer the G-8 Open Data Charter was established, but there was a commitment made by all G-8 countries to involve themselves in annual reporting so that later on this year the expectation would be that individual G-8 countries would have to update on what commitments within their G-8 action plan have been met, and the same thing again for 2015 but the reporting template has not been determined yet.

[Translation]

Mrs. Anne-Marie Day: Page 29 of your document refers to Canada's Action Plan for on Open Government. It refers to an open dialogue with Canadians, which means that there will be consultations with Canadians and open regulation. How will this dialogue be organized? Earlier, you said that people can criticize, make suggestions, ask questions and you will read and follow up on those messages with the goal of improving the site. I imagine that you are not receiving a large number of messages because if you were, it would be too much.

Will you hold other kinds of consultations? Those who do not know that the site exists will not consult it. They will not send you potential questions. How are you proceeding?

Ms. Corinne Charette: Last year we used various means to solicit feedback on the site from Canadians. They were invited to participate in online consultations and were asked questions over a certain period of time. We asked them to answer by submitting their comments and their questions.

As well, the minister used TweetChat. He established a dialogue on Twitter during several hosted sessions which were announced on our website to encourage participation. The minister then participated in a series of open discussions throughout Canada, in several municipalities, to promote open data in various regions. I believe that he continues to do so.

Finally, we are currently developing our consultation plan for the second phase of our action plan for the Partnership for a Transparent Government. We hope to finalize it over the next few weeks. Our plan will determine when we will continue to encourage Canadians to participate and the means that we hope to use for that.

•(1025)

Mrs. Anne-Marie Day: Thank you.

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

Mr. Aspin, you have the floor.

[English]

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

Madame Charette, you mentioned quite a number of times that it's all about leveraging assets. Both of you mentioned and referred to the U.S. and particularly the U.K. as being leaders in this field.

Are there any other countries that are off the mark on this and doing well, and could you name and describe them?

Mr. Stephen Walker: Sure. There are a number of other countries, perhaps not quite as close to us, so that we don't work with them quite as frequently. Brazil is significant. All of the Scandinavian countries are very advanced in terms of open data. Australia and New Zealand are also active in open data, and in open government in general.

As Corinne mentioned, there are almost 70 countries now that are formal members of the Open Government Partnership. Most of those countries would have specific open data initiatives as part of their open government action plans.

I should also mention that some of the really advanced open data work is at the subnational level, so in the United States, for example, at the state level, and here at the provincial level. That holds true within some other jurisdictions as well, such as Australia.

Mr. Jay Aspin: Where would you rank us, generally? Would it be in the top echelon, the middle?

Mr. Stephen Walker: We're at the top echelon.

Mr. Jay Aspin: We are?

Mr. Stephen Walker: Yes.

We weren't. The U.K. and the U.S., and other countries had definitely hit the ground running a little bit faster than we did, but we've caught up.

In many instances that's helped to bind the relationship we have with those two countries. The work that we continue to advance is now work that we're doing with them, and that has helped us to close the gap. Internationally, other countries—and not necessarily the U.S. and the U.K. as they think of themselves as being in the upper echelon—will be coming to us for advice.

Mr. Jay Aspin: Is there much data set interchange among the countries?

Mr. Stephen Walker: The G-8 Open Data Charter is specifically designed to support more comparison of data across the G-8 jurisdictions in very specific subdata domains, such as budget, procurement, and legislation. They want to be able to look at the data sets from different countries across the G-8 in order to compare and contrast.

Mr. Jay Aspin: Thank you.

[Translation]

The Chair: Thank you.

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

[English]

Hon. Gerry Byrne: Thanks, Mr. Chair.

You reference, with some vigour, the notion that privacy must be protected, which I think is reassuring to us all. But at the same course, a lot of data sets will indeed include elements of personal privacy. There will be a reference to an individual, whether it be my example of fisheries or the registrar of Indians or any of a countless number. That which government measures often includes the measurement of people, so there will be a reference to personal information. Do you take the decision to just simply not include that entire data set into open data, or do you redact, as Mr. Martin had alluded to?

• (1030)

Mr. Stephen Walker: One of the most interesting and sort of revolutionary parts of open data going forward with our colleagues in the U.S. and the U.K. is around new technologies that allow for the anonymization of data. I can say without hesitation that there's no personal data of any kind in any of the data sets that have been made available. Right now, when we're going out to individual departments to look for data, there can't be any personal data of any kind, not even a blip. But in the future, as we try to expand the amount of data being made available by a single department, we hope to be able to work with them to anonymize it so that it can still be made available.

Hon. Gerry Byrne: Right.

Just to expand on that a little bit, under what parameters would that be the case? Just clarify that for me, you want to expand on it so that...?

Mr. Stephen Walker: Within the government, we define personal data very specifically. It's a long list. None of those data elements can be contained in any of the data sets that we currently make available.

Hon. Gerry Byrne: Some data sets already have highly personal information. For example, if you look at a Government of Canada telephone directory, it lists not only the individual but their position, and then you can cross-correlate with pay scales and so on and so forth. Is this a concern to the open data project, that data miners could correlate information provided through open data with other data sets that are not protected from a personal information point of view, and very purposefully so?

Ms. Corinne Charette: First of all, the government directory with names, titles, and telephone numbers is not personal; it is available online.

Mr. Stephen Walker: Yes, that's public information so that one is not really a concern. I think if you're talking about Canadian citizen client information being included in individual data sets, the reality is that right now because privacy is such a paramount concern to us, we have to be exhaustive about working with the departments to ensure that no aspect of personal data is contained so that it cannot be correlated with any other data set. That's part of our work day to day with the individual departments.

Hon. Gerry Byrne: How did the G-8's assertion towards this initiative assist in crossing boundaries between jurisdictions? There is some information that Canadians would like to have access to that currently is not available because it involves a second jurisdiction, an international partner. Are there attempts being made to ensure, for those that are members of this charter, that those boundaries be removed?

Mr. Stephen Walker: There really are. The charter specifically listed a large appendix of very precise data sets that each individual member state has agreed to provide by the end of the two-year period. We all, as well, have to provide an update at the end of 2014 on how we're doing with those.

In terms of us, as Corinne mentioned, and the U.S. and the U.K., the bulk of these data sets are already available. You can already compare them under the existing licensing regime across those three jurisdictions. Some of the other G-8 countries are not quite so advanced. The monitoring that will be done will be to help them to catch up to the rest of the G-8 numbers.

Hon. Gerry Byrne: I have one final question. In terms of procurement, other governments have a tendency to be very covetous over internal procurement data. Would there be a schedule or chapters that would include that which would be included in the open data within the G-8 concept?

Mr. Stephen Walker: Procurement data was one of the areas defined.

Hon. Gerry Byrne: Thanks very much, I appreciate it.

[Translation]

The Chair: I would like to thank the witnesses for accepting to appear before our committee.

We will suspend the meeting for a few minutes and we will come back to discuss the second point on our agenda.

I will take this opportunity to thank you for your presentation on open data. This will be very useful to the committee in its future work.

• (1030)

(Pause)

• (1035)

The Chair: I call this meeting back to order.

As I announced at our last meeting, we will vote on the supplementary estimates (C) 2013-2014. Pursuant to standing order 81(5), votes have been referred to the committee.

[English]

PRIVY COUNCIL

Department

Vote 1c—Program expenditures.....\$2,062,421

(Vote 1c agreed to on division)

[Translation]

PUBLIC WORKS AND GOVERNMENT SERVICES CANADA

Department

Vote 1c—Operating Expenditures..... \$68,822,039

(Vote 1c agreed to on division)

Shared Services Canada
Vote 15c—Operating Expenditures..... \$102,331,630
Vote 20c—Capital expenditures..... \$1

(Votes 15c and 20c agreed to on division)

TREASURY BOARD
Treasury Board Secretariat
Vote 1c—Program Expenditures..... \$1,424,230
Vote 15c—Compensation Adjustments..... \$73,303,264

(Votes 1c and 15c agreed to on division)

[*English*]

The Chair: Shall the chair report the supplementary estimates (C) 2013-14 to the House?

Some hon. members: Agreed.

● (1040)

[*Translation*]

The Chair: Those are all the votes, but I would like to add something before we adjourn. We have asked the clerk to check the dates when the President of Treasury Board would be able to come and appear before our committee. However, he is not yet available to appear on the main estimates 2014-2015.

We are adjourned until Thursday, March 6th.

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>