HOUSE OF COMMONS
CHAMBRE DES COMMUNES
CANADA

Standing Committee on Government Operations and Estimates

EVIDENCE

Thursday, April 3, 2014

—

Chair

Mr. Pierre-Luc Dusseault

Standing Committee on Government Operations and Estimates

Thursday, April 3, 2014

● (0845)

[*Translation*]

The Chair (Mr. Pierre-Luc Dusseault (Sherbrooke, NDP)): Order, please. We will begin immediately.

We have a busy day today. We are hearing from four expert witnesses for our study on open data. I will introduce them right away. We have David Eaves, open data expert, appearing as an individual. We also have Renée Miller, who is a professor at University of Toronto's Department of Computer Science. By videoconference, we will hear from Ginny Dybenko, Executive Director of the University of Waterloo's Stratford Campus, and Mark Gayler, from Microsoft Canada. Since Mr. Gayler is in Vancouver, where it is currently 5:45 a.m., we will show some indulgence toward him.

Without further ado, let's start with the presentations of up to 10 minutes for each witness. Afterwards, the committee members will have an opportunity to ask the witnesses some questions.

I yield the floor to you, Mr. Eaves. Thank you for joining us and for taking the time to testify before our committee today.

[*English*]

Mr. David Eaves (Open Data Consultant, As an Individual): Thank you.

Just so Mark doesn't get all the credit, I'd like to note that I also just came in from Vancouver, so it's also 5:45 in the morning for me.

Voices: Oh, oh!

Mr. David Eaves: My name is David Eaves. Since I'm listed here as speaking "As an Individual" and don't have the credentials of some of my peers, perhaps I'll start with a little bit of background on me.

For the last five, six, seven years, I have been working to make open data happen in Canada. I wrote the original motion that led to the creation of the open data portal with the City of Vancouver. I then worked behind the scenes with some of the provinces to help them create their open data initiatives. I gently applied pressure on the federal government to persuade them to adopt open data as a policy.

I also work with several governments. I ran the boot camp for the Presidential Innovation Fellows at the White House. I've worked for the State Department and the World Bank. I sit on Mr. Clement's open government advisory panel as well as Premier Wynne's open government task force, which recently released its results. I sit also as an affiliate at the Center for Internet and Society at Harvard University, at the Berkman Center. I also sit on the boards of several non-profits as well as several start-ups, both in the open data space and outside.

I want to share a few thoughts with you about what I think matters about open data, how we're doing, and some of the things we could be doing. Maybe just as a little bit of a backdrop—I imagine everybody is trying to explain to you what open data is and why it matters—I'll give you a simple metaphor.

I'm carrying around with me a Fitbit, a small device that tracks how many steps I take every single day. This is mostly because I have the potentially bad belief that if I take 10,000 steps, I can eat whatever I want. So I try to get to 10,000. When you look at this device, it's tracking some data about me, specifically my movement. Increasingly as you look around, all sorts of data is being tracked about you and created about you, from your bank statement to your mortgage to where you're going. This device happens to know where that happens to be all the time as well. It would be nice to think that you could harness all that information to tell you something useful about your life that could cause you to change your behaviour, or to do something different, or to save a little bit of money.

I'd like to apply that metaphor to the federal government. Right now there are probably about a billion of these types of devices. Whether there are people tracking expenses in Excel spreadsheets, or devices measuring the weather, the temperature, or something else around the country, around the world, all of that data is being collected. Wouldn't it be nice if we had access to it so that we could say something intelligent about this country and about our community, and maybe change some behaviour here, or figure things out that are not going well?

I think the open data initiative is trying to solve the same problem that many people are trying to solve on the consumer page: how do you harness all of this data that's being created, some of which you don't even know is being created, or where? Can you bring it into a central place where it becomes useful, actionable, and leverageable by a community of people?

Hopefully that gives you a metaphor that makes it a little bit easier to understand what open data is and how potentially it can be useful.

I think for me, there are one or two examples that strike me as the most interesting around how far we've come and how far we have yet to go. I'm sure this committee is interested in knowing, as everyone is, how we are doing internationally. I would argue that internationally Canada is doing relatively well. We're not what I would consider to be a front-running leader. We're not like the United States or the United Kingdom. But we're also not a laggard. Maybe only 20% of the countries in the world are thinking very, very hard about open data, and we sit very comfortably in that group.

The real danger I would flag around this is that I think using international comparisons, especially this early on, and any time in government, is always enormously dangerous. I get very frustrated when I see comparisons about how governments are performing in technology and then people becoming satisfied about being at the top of those rankings. Whenever you have a ranking of government performance in technology, what you are actually doing is you're taking all the slowest movers in a space, putting them into one category, comparing yourself with all the other laggards, and then sitting around and congratulating yourself for being really, really fast against the other slow-moving players in this space.

Leadership, for me, is not whether you're the second bull in a herd. The problem with leadership is that you're in the herd to begin with, and real leadership is how you actually break away and go do something that other people are not willing to do. Whether it's within a country or internationally, the herd mentality is so strong that I think it actually prevents leadership from happening. If you benchmark yourself against others, what you're really doing is saying that you just want to be inside the herd, and then asking where you rank in that herd, as opposed to really thinking about what leadership and transformation could look like and doing the things that could potentially really change society in a positive way.

● (0850)

While I think the international metrics matter, I caution you strongly about getting sucked into them and somehow believing that they're a magical metric that should determine whether we're happy with our performance or not, because they're almost invariably very, very poor.

The other question for me is, what's our goal? What are we trying to accomplish? When I look around I see different players doing different things, and they have, I think, a real vision for where they want open data to take them. I think that vision is less clear in Canada. Certainly, we're not realizing our full potential. I suspect that many on this committee are most concerned with the economic benefits of open data. I think those could be significant. I think there's a real risk of overplaying them, and I think there's an enormous amount of hype. I would be very, very cautious about believing every figure that passes by you, or why it's going to have an economic impact—and I'll talk about that briefly. There are also huge opportunities around making government more transparent, which actually has economic benefits in and of itself, as well as more accountable. I wouldn't want those to be lost.

The third is there are huge opportunities in reshaping how public servants work with one another and in using open data to vastly improve the efficiency and productivity of public servants. I wouldn't want that opportunity lost in the pursuit of economic goals.

With that said, I think there are four big ways that open data will serve to transform a society. Let me highlight each one of these, and what I think matters and doesn't matter about them.

The first, which I'm sure you've heard endlessly about, is the opportunity around apps. I'm not the first person...and I don't want to say that apps don't matter. I think apps are enormously important, and there is an enormous opportunity there, but I actually feel it's also the least relevant of all the opportunities before us, especially when it comes to federal government data. A great deal of federal government data is aggregated at such a level that it becomes pretty hard to do anything particularly interesting with it. In addition, the vast majority of the data is created for policy usage and not for day-to-day application usage. There are definitely datasets out there that are very, very interesting. Border wait times, which you heard Colin McKay talk about, I think, is a great example of that.

Operational data is very interesting, but the vast majority of your data is actually geared toward policy analysts, so it's geared to trying to do analysis and understanding what's going on in society or what's going on in the community.

That brings me to the second big place where I think we're going to have impact, and this is the one that I think is the 800-pound gorilla in the room, which is the opportunity for open data to dramatically improve analysis and productivity. The example I like to use is one that I think comes a little out of left field. I'm involved in something called the Canadian Boreal Forest Agreement, which is an agreement between the largest environmental groups in the country and the forestry industry. It's all about trying to ascertain where logging should take place in order to preserve woodland caribou and maximize the benefits for regional communities that make use of logging infrastructure.

This entire agreement is made possible because of enormous amounts of data about knowing where the woodland caribou is, about knowing where different types of trees are, and being able to layer maps over one another to figure out where the places are that we shouldn't be logging and where the places are that we should be logging.

There's no app that's going to come out this. But if you look at the total impact the CBFA could have on the Canadian economy, it could be in the billions. If you talk about no longer having protests against forestry companies by environmental groups, they're actually supporting logging companies as they try to sell their products internationally, having a wood product that is actually seen as ecologically viable and therefore more valuable, about the impact on local communities and the jobs it creates, the impact on shareholder value around all of the different logging companies, it is not hard to imagine that number very quickly running into the hundreds of millions, and even the billions.

That entire project is supported by government data. So for me, rather than focusing merely on apps, thinking about the much larger policy opportunities and the economic opportunities around analysis that open data can provide is the place where I think particularly federal government data becomes enormously valuable and interesting.

The third is the internal use of open data and the way I think it can transform how our government works. I've been around the world talking to people who run open data portals, and invariably you find that roughly 30% of the users of an open data portal come from computers that are located within the government that made that open data available. It's not hard to understand why. The data government creates is most useful to people who work in government. The problem was that before you had an open data portal, in order to make use of a dataset I would have to go and talk to you, and then your manager, and then maybe your ministry's lawyers had to get involved before they decided who was allowed to use this data or not.

● (0855)

You had nine meetings, eating up 10 public servants, 40 hours for a week, just so I could get access to data, and most of the time we're like, forget it; I don't want to even bother anymore.

Now all of a sudden, not all, but a significant amount of government data is available in a place where public servants can very quickly access it. All of that time spent negotiating over whether or not I should have access to something that's actually already a public asset has just disappeared. So the productivity opportunities within government, I think, are quite significant.

[*Translation*]

The Chair: I would ask you to wrap up your presentation, Mr. Eaves.

[*English*]

Mr. David Eaves: Yes.

If I were looking at this committee and I was trying to think about how we were going to assess the value of Canada's open data policies, I'd be looking at three things.

The first thing I'd be looking at is whether we are thinking strategically about the policy in economic areas that we want to be driving into and what the data is that we're releasing that might support those places.

The second is whether we are thinking hard about how government itself is using the open data that it releases, so it does what we call "dogfooding", which is that it uses its own information rather than sharing with others and expecting them to use it, but using something completely different itself.

The third is whether we are actually sharing information about government itself. Where is the budget? Where are the things that make government transparent so that citizens themselves can better understand and make government more legible, so they can become more engaged in the political process and contribute in interesting ways in the policy debates?

Thank you.

[*Translation*]

The Chair: Thank you for your presentation.

I now yield the floor to Ms. Miller, from the University of Toronto.

[*English*]

Dr. Renée Miller (Professor, Department of Computer Science, University of Toronto): I would like to thank the members of the committee for inviting me to participate on this important panel.

I am a computer science researcher, and I study the problems and the opportunities that open data presents to the science of computers. In particular I study the problem of data curation, which can briefly be defined as ensuring that data maintains its value over its lifetime, ensuring there is still value in that data and that value can be used by humans.

I would like to present three points. I'm going to try to reduce the geek level. I realize I am used to talking to computer science students so please let me know if you don't understand anything I'm saying here.

I have three points that I think can help Canada become a leader in the open data revolution. My first point is that I think the open data portal should adapt the principles of open link data. So when we put a file up on the web we are using technology that's been around for even more than 20 years. Since the beginning of the web we've been able to share data in files over the web. The state-of-the-art data sharing is not just sharing these static, inanimate files. When I say "we" I mean scientists like myself, academics but also industry leaders. When we share data we share data that's linked and that means the objects we're referring to in the datasets are dereference-able, it's a fancy geek term meaning I can click on it. When I click on it I get important and interesting information about that object and among that important and interesting information I get relationships to other objects and important information about them.

So let me give you a concrete example of that. The most downloaded file from the open data portal in February 2014 was a file about charities. It's a static file. It just has strings and it has numbers in it. So it has strings naming different charities and it has facts about those charities, but it's just a dead file. What I would like is when I download that file, that when I see, say, the Rideau Street Soup Kitchen, I would like to be able to click on that link and get important information about the soup kitchen. For example, I'd like to get important information of where it's located, what community it serves, how many people it serves, some of the facts, some of the data. How much federal money it gets is in that file but other information like whether it gets provincial funding, private funding, who those private funding agencies are and information about them, that's not there.

But it's very easy to provide using today's technology to make the data linkable and to use the principles of open link data to enrich this data. So that's my first point, to embrace the principles of open link data.

My second point, which is highly related, is that open data is about information flow and that information flow can't be unidirectional. If the flow of information is solely from the government to the public then there's no incentive for people to do interesting and creative things with that data. So if we just make the data available, take a data file and throw it over the transom, close our eyes and hope somebody is going to do something interesting with it, we're not creating the incentives to get people involved, to change lives with the data, to solve problems, improve government, just for the economy.

Worse, I think it has the potential of creating this adversarial relationship. It gives the perception that the government's in control of the data, and is just handing it out. There is no ownership or investment in the data itself. So I think open data is fundamentally about creating participatory opportunities where people can become invested in that data and are incentivized to contribute to the data itself and incentivized to improve the data and to create new innovative ways of using the data. I think this investment creates trust and people will trust the data if they can contribute to the data. It also provides an information flowback into government—David was speaking to this as well—where the information itself is flowing back into the government improving government decision-making based on better data.

● (0900)

My third point is that opening up data is important, but it's equally important to create and curate participatory opportunities with this data. These are not just appathons. I think they are other ways in which the community can get involved in doing analysis over this data and improving this data.

I know the open data portal is already deeply involved in this. They have the Great Canadian Appathon, which is in its fourth year. It was just at the University of Toronto, so we're already doing quite a bit along these lines.

But I think there are two important outcomes of this. These activities are not just educational; you're not just teaching students how to use data. Rather, you're looking to find those visionary students—I call them students; everybody younger than me is a student—those visionary people who want to create new entrepreneurship opportunities with that data. I think that using open government data is an absolutely terrific way of getting those folks to stay in Canada. Too many of my entrepreneurial students go to Silicon Valley because that's where the air is filled with start-up culture. You can walk into a start-up and somebody will be there to help you with your start-up.

That kind of start-up culture just doesn't exist here in Canada, but if we have more people using community data, open government data, in their start-ups and building things around that, there is an incentive for them to stay within Canada. But, also, in addition to these visionaries who are going to spur our economy, I would encourage you to look to activate the power over the crowd; and by that I mean creating datasets where the community itself can contribute to those datasets. These are things like allowing the community to comment on the number of open beds in homeless shelters and activities like this where the community gets involved in improving the data and gets invested in the data itself. The power of

the crowd can be really important in leveraging government open data.

In conclusion, I'm absolutely thrilled to see this initiative in Canada. I think it is a tremendous opportunity here, and Canada has the potential to become a leader in the open data revolution, and I look forward to seeing much more.

● (0905)

The Chair: Thank you for your presentation.

[*Translation*]

We are now moving on to Mr. Gayler, who is testifying by videoconference live from Vancouver.

You have 10 minutes, Mr. Gayler.

[*English*]

Mr. Mark Gayler (Technology Strategist, Western Canada Public Sector, Microsoft Canada Inc.): Hello, and thank you to the committee for inviting me to participate this morning. It's bright and early in Vancouver.

My name is Mark Gayler. I work for Microsoft Canada. I've been working with Microsoft for more than 10 years. I'm a technology strategist for Microsoft Canada. I work primarily with municipalities. As part of that role, I'm a subject matter expert on open data and open source technologies.

I'd like to comment on a few things. First of all, I very much appreciate the comments by my colleagues David and Ms. Miller just previously.

One of the things I have experience with is working with different governments around the world, and so I've been engaged with open data projects in Canada, but also in the U.S.A.., Colombia, Japan, central and eastern Europe, and the U.K. I'd like to make some comparisons, even though I totally and fully agree with David's comment earlier on that it's dangerous to make comparisons in terms of a league table. But I think there are some insights we can gain from what other countries are doing compared with how open data has evolved in Canada today.

I'd like to start there, and then I'd like to pick up on a couple of other points that my colleagues have raised already.

What is interesting about the way open data is evolving around the world is that it's evolving in different ways based on the way that government agencies have chosen to engage it.

For example, in the U.K. and the U.S., we see a very top-down approach whereby the U.K. and U.S. governments at the very top levels of government have sponsored open data initiatives. They are driving adoption of open data throughout government departments and agencies, and we see this top-down approach as it flows downwards through the government infrastructure.

I would say that in Canada what we have seen is more of a bottom-up approach to open data. In early days it was adopted primarily by the cities, and then the provinces caught up. I think Vancouver started in April 2009, and we have seen other cities adopt open data initiatives. Then the provinces have come in, and I think the federal government has come in after some of these cities and smaller agencies had already adopted open data initiatives.

That explains why we see different countries and different initiatives at different stages of evolution, to a certain degree.

In the U.K. and U.S., I would say that open data initiatives across government are fairly mature and fairly consistent in the way open data is thought of. I would say that in Canada we see open data being adopted in different ways at different levels of government jurisdiction.

The second point I'd like to make around this is that as we look around the world, it's important to understand that open data itself is not an end point. Open data is a transition to something else. It's an enabler for other things to happen. It's an enabler for such things as economic stimulus, as we have discussed, and I'm sure we'll discuss more on that during the session. It's an enabler particularly for citizen engagement, getting citizens actively involved and participating in the business of government.

I think it also represents a cultural change internally for government and government agencies. When I've been around the world talking to national and provincial and state governments about their open data initiatives and the way we can use open data to engage citizens, particularly those parts of citizenry we may not already be engaged with, a big comment that I get at the end of my engagement with that particular government is: this is great, but now that we have this capability to share data and to collaborate, we want to do it internally as much as we want to do it externally. I think that point was made very well by my colleagues previously.

The opportunity for the Canadian government here is to provide guidance, to provide a framework to take the open data initiatives that already exist, to create opportunities to share more open data, to engage citizens and third parties and encourage them to share this data and use this data, and to enable the sharing of the data in such a way that it can easily be consumed by any of the actors in the ecosystem, be it a data scientist, a researcher, a citizen, an application developer, or a student.

● (0910)

But it's very important that we understand that this is a cultural change that will lead to other positive benefits; this is not just about sharing data itself. And so it's important that the government provide a framework to encourage parties to collaborate around the sharing and reuse of open data—private-public partnerships, for example— and particularly engage those parts of the citizenry with whom perhaps we are not already engaged and get them actively involved in the business of government.

Let me give you a very simple example. Two weeks ago we ran a teen hackathon in the city of Surrey. The City of Surrey is sharing its open data; they have an open data portal. They invited teens, young people from the ages of 13 to 19, to participate in this hackathon. For half a day we worked with them with technology and showed them how to produce applications. What was interesting is that at the end of it we asked for feedback and ideas, and it was amazing to see these teenagers come up with ideas about how to use transit data to better navigate through the city, how to use weather data to better understand when weather might affect particular tourist spots or landmarks.

You could look at that initially and just say that these are interesting ideas but ask whether they would ever come to any kind of fruition. But what was really interesting about the whole thing was that the city was stimulating students and young people to think about engaging the city in ways that had not previously been possible. These were young people who were thinking about actively working with the city—visitors to the city, citizens of the city. Getting them excited and engaged in looking at ways to improve city services both for visitors and for folks who already live in the city is quite transformational. This is a very simple example of transformational cultural change that can be brought about by sharing open data.

Another example I will give you, from a cultural aspect, comes from when I was engaged with the Government of Colombia. I was invited down there to provide some guidance to them about the way they would share data with their citizens. When I went down there I said I was surprised that the Government of Colombia was thinking about sharing open data, because they're not known, to an external person, for their openness or the way they might engage a citizen in a transparent way; that it might be considered to be a threat to the government.

They said that this was their entire reason for doing it. Whereas other governments say they're doing this for economic stimulus or doing it for better engagement with certain parts of society, in Colombia they are doing it deliberately to show that they're being open and transparent. This is part of their cultural change with their citizens.

The last point I would like to make is that I think the opportunity is huge for Canada to be a leader in this area. Even though we look around the world and see open data initiatives evolving in different ways, we have a long way to go with open data, to speak to David's point earlier on. There is much more that can be done and there is much more transformational benefit that can arise out of open data.

But I think the government can help. It can stimulate this by providing, for frameworks for working particularly in public-private partnerships, guidance in the sharing and openness of data, and also by providing ideas and guidance about the sustainability of open data and how it can be part of the ongoing business of government and citizen engagement, rather than just being seen as an end in itself.

Thank you very much.

● (0915)

[*Translation*]

The Chair: Thank you for your presentation.

We are now moving on to the last, but not the least, witness, Ms. Dybenko, from the University of Waterloo, who is appearing by videoconference from Kitchener, Ontario.

[*English*]

Ms. Ginny Dybenko (Executive Director, Stratford Campus, University of Waterloo): Thanks very much.

There have been great remarks already that have taken a lot of the points that I was going to make.

I would start off by saying that over the past 20 years, we've seen an awful lot of innovation, creativity, and disruption. Today we don't know exactly where open data will lead, but we do know that it will be very transformative. Some ways of doing business will start and some will evolve, and learning how to navigate them will be the challenge that lies ahead for us. But the potential that certainly we saw in the very early days of the web—and I lived through all of that —is what I see now with open data.

I believe open data is data with a mission. It will create jobs; it will fuel start-ups and launch new industries with revenue purportedly in the billions. However, every day untold numbers of people try one more time to figure out Facebook's privacy settings and wonder exactly what Facebook knows about them anyway, and most people have only one concern about their personal data, and that is that they want to keep as much of it as they can as private as possible.

So there is a central paradox here. Releasing personal data as open data can definitely benefit society and ultimately help the individual, but if the data is not controlled carefully, having it out in the open will damage individual privacy and may outweigh the benefit and slow the process down.

I was introduced as the executive director of University of Waterloo's new Stratford Campus focusing on digital media. Additionally, I have had an extensive background in technology in corporate Canada with 30 years at Bell Canada in IT and digital communications and I am currently a member of the board of governors for SSHRC, the Social Sciences and Humanities Research Council of Canada, so you can understand that I have a number of viewpoints on the open data opportunity.

I'd like to begin with the viewpoint of our wonderful young digital natives enrolled in our undergraduate and graduate programs at the University of Waterloo. Digital natives are young people who were born in 1997—ouch—who have literally never not had a device in their hands. And as Mark mentioned, we have engaged them on a number of occasions with the municipality.

We recently ran a project for Stratford on garbage, of all things. It was a hackathon run over a weekend. Essentially, when do you put your garbage out and how do we communicate with our citizens? We didn't think that our young students would be particularly interested in this. They dove in and produced some remarkable methods of connectivity that the city is now looking to continue to develop.

But that pales in comparison with Code 2014, run by Tony Clement. That was a hackathon run across Canada about a month ago engaging 900 young people that challenged them to develop apps around the open data that the Canadian government already has laid out. I am delighted to tell you that of the 900 applicants or participants, a team from our school actually won. Their application essentially delved into StatsCan employment and social development, Canada data, the Canada Revenue Agency, and the CMHC in helping immigrants choose the right place to come when coming into Canada. It was featured on the CBC this morning.

There are a number of other kinds of applications that were developed as well. Fifteen were finalized.

But, Mark, I loved what you said. From my point of view, what we're really doing here is engaging young people in the affairs of the government. That has been a huge challenge, I think, certainly at municipal and provincial government levels.

● (0920)

With regard to the expectations of these digital natives, as a consumer, they definitely want the personalization of their experience which comes from open data, but they also want to ensure that their data is very private, or that they have control over that. As an entrepreneur, they want ready access to the data, but they also want assurances of ownership once they have developed their idea.

At a corporate level, I think a storm is brewing. Corporations want access to unattributed personal data to examine trends by demographic group, for example, but they want attributed data to do specialized or specific targeted marketing, which is scaring a lot of people. The opportunity examples that McKinsey has pointed to look to billions, $300 billion annually in health care, in the U.S. alone. They go throughout the world and the opportunities are limitless, as David referred to earlier.

Finally, I do believe that the Canadian government is at the leading edge of governments around the world. Certainly we are seen as leaders, and we are regularly referenced, particularly in U.S. documents on the topic. I think we are well positioned for significant savings. The granting agencies, the tri-council, conducted extensive consultations around big data's role in the development of digital scholarship in Canada, which was conducted in the fall of last year, coining the term "open research".

The conclusions were threefold: first, that there is a culture of stewardship that asks for an establishment of clear policy for data sharing; second, that there is a coordination of stakeholder engagement, in other words, long-term planning—and remember that this isn't about data on colliding particles but mostly data on people—and therefore the involvement of SSHRC is very important; and third, they raised as an issue the developing capacity, so that engages funding and roles and responsibility among national, provincial, and institutional stakeholders.

In conclusion, I'd just like to say that I believe open data is our next natural resource. Canada has the digital infrastructure. We have the reputation for collaborative management. We have the respect of many in the world in this arena, and we have a hugely developing knowledge worker population, through programs such as ours in Stratford. Canada should make open data a priority, establishing policies, engaging in long-term planning, and developing capacity.

Thank you.

[*Translation*]

The Chair: I want to thank all four of you for your presentations.

As planned, we will now begin the question and answer period between the committee members and the witnesses. I would like the committee members to specify which witness they are putting their questions to. That will make things easier, since two of our witnesses are appearing by videoconference.

Mr. Blanchette, go ahead. You have five minutes.

Mr. Denis Blanchette (Louis-Hébert, NDP): Thank you, Mr. Chair.

Thank you very much to our guests.

The issue we are discussing this morning is very thought-provoking. Everyone approaches it in their own way, from a unique perspective. There may not have been enough time for this, but I would have liked us to put open data in its true context—in other words, the transformation of our society to a digital society. Whether we like it or not, the Internet now plays a concrete role. It's practically a right nowadays. People have access to the Internet almost as they do to running water.

The use of open data is part of that context. Some elements that were mentioned are already becoming a reality. Websites such as Facebook and Amazon are already partially personalizing users' and consumers' preferences.

Mr. Gayler, I thought your approach was quite noteworthy. You talked a lot about the approaches and the context. According to my understanding of your presentation, the Americans started with the federal government, while our approach was a bit more heterogeneous.

Have you looked into what approaches have been used outside North America? What other trends are out there? Do you have an idea of what is going on elsewhere? As representatives of the Canadian government, we would like to have a good idea of what is happening on the international stage.

● (0925)

[*English*]

Mr. Mark Gayler: First of all, I can only obviously comment on the jurisdictions where I have personal experience.

I would say a couple of things. If I back up a little bit and re-clarify what I said earlier on, I think where we see a top-down approach such as in the U.K. and the U.S.... And what we mean by a top-down approach in open data terms is where guidance is given by the national government in ways that data can be shared, how departments can share that data. They provide guidance and frameworks to enable that to happen as part of the business of government. That's what we mean by that top-down approach.

What we tend to see is that governmental departments, then, become more encouraged to share data because they have been given a mandate by the national government, if you like, and it becomes more baked into the process of government rather than being seen as, "Well, we do government, and oh, we also do open data."

I think there's some good learning there for Canada, certainly.

Canada is in the position where it can certainly exploit some of this learning that we see in the U.K. and the U.S., but that's not exclusive to other countries. If we look at Germanic countries—for example, Austria, Germany, Switzerland—again what we see there is that this is very much city-based. The national governments are looking at open data initiatives, they are looking at open data policy, but by and large, to date, the way that the citizens have engaged on open data is through city and provincial open data initiatives.

I would say the same for Italy, for example. If we look at some work that's been done with the Italian Ministry of Health to share data, if we look at the initiatives that are going on throughout Italy, they are largely city- and provincial-based. And there is a reason for this.

If you think about data—the value of data and its relevance to citizens—national data, of course, is interesting; statistical information, of course, is interesting. It's particularly interesting to data researchers and data scientists.

However, if you look at the average citizen, they're interested to know when their garbage is available, what the health situation is in their local school area, for example. Local data has a lot more relevance to the average citizen in many cases than, say, national trending data. That's why we see these initiatives evolving in different ways and citizens engaging and taking up that data in different ways.

● (0930)

[*Translation*]

The Chair: Thank you.

[*English*]

Mr. Mark Gayler: I hope that was....

[*Translation*]

The Chair: Thank you, Mr. Gayler. Mr. Blanchette's time is up.

We now go to Mr. Trottier, who has five minutes.

[*English*]

Mr. Bernard Trottier (Etobicoke—Lakeshore, CPC): Thank you, Mr. Chair. I'm very impressed with our panel of witnesses here this morning.

We started this initiative a few weeks ago, and it was recognizing that, as a federal government, we've signed a G-8 open data charter. There's a mandate now to develop a road map for the federal government, and I appreciate the different perspectives.

I'd like to get some input from all four of our witnesses this morning on this notion of the government as a publisher of data, very much a one-way flow of information from government to citizens versus the notion of the federal government being more of a facilitator or creating the public square where people can publish certain data, a way to engage them. Many examples come to mind where the government can't create the data. If you think of species at risk, for example, where there are eyes and ears all over the country, and people might be able to spot a rare bird and they can provide that information.

One of the challenges with providing that public square is how do you confirm whether the data is good or do you need to confirm? Some people could be there, not so much to publish data, but they have a certain point of view they try to advocate and they could hijack that public square.

Can each of you in turn, in the order that everybody spoke, talk about that 1.0 version of open data versus a 2.0 of more of an engaged version of public data?

Mr. David Eaves: Absolutely. I agree there's opportunity there, but for the reasons you mentioned, I'd be fairly conservative about how I would try to engage in that opportunity. One of the things that the government has done, and I think has been quite effective, has been a canonical source for data that is highly trusted. Statistics Canada creates data that is highly trusted by people in the non-profit, for profit, and government space. Having something whereby people can all point to a dataset and say they believe that and they use that as the foundation for their conversation is enormously useful and cannot be underestimated.

To talk about how you crowdsource the creation of data creates an enormous number of methodological problems that I would be wary of rushing into, especially when we have so much data that is canonical and is verifiable that we already are not sharing and I would argue are not leveraging as effectively as we could. I'd much rather solve that initial problem first before thinking too much about that second problem.

Mr. Bernard Trottier: Ms. Miller.

Dr. Renée Miller: Let me give you an example of where this has worked. In the U.S. They have a portal, I forget the exact name, but it's something like peer to patent database, where they have opened up the patent process to input from experts, recognizing that the expert on a particular patent topic is often not in the government itself. They invite scientists and inventors to comment on patents that are under review. They have found that they get much higher quality information from that process than they could just adjudicating the patents themselves using their own experts. They didn't abdicate the responsibility for making the final decision, for having somebody

making sure there wasn't somebody with a bias inputting data, but they were able to get much richer information on which to make their decisions.

I think you can walk that line. I think you do have to be careful with it, and still have somebody adjudicating the information itself.

Mr. Bernard Trottier: Mr. Gayler and Ms. Dybenko, could you comment briefly on that? One of our objectives of this study is to give our own government, the Treasury Board specifically, that direction on how to create that road map for the federal government.

Mr. Mark Gayler: I think one thing that's very important to understand with this topic is that all data is inaccurate to a certain level, so you can't wait until the data is 100% accurate to share it. That's something that some government departments feel very concerned about. Again I agree with David's point. I think the emphasis here is on sharing data that's not being shared today and setting an expectation for the integrity and accuracy of that data as it goes out either to public or to commercial entities for that matter. I think as long as you're clear about what that might be, that would be where I would place the emphasis first.

● (0935)

Mr. Bernard Trottier: Okay, thank you.

Ms. Ginny Dybenko: If your eventual goal is engagement, then there is nothing like asking someone to contribute to the dialogue in a real way to engage the individual constituents. The accuracy aside, I think it's a process that's well worth pursuing.

The Chair: Thank you.

Thank you, Monsieur Trottier.

[*Translation*]

Mrs. Day, the floor is yours for five minutes.

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

My questions are mainly for Mr. Eaves or Ms. Miller.

As you know, our study is about improving the government's open data practices. More specifically, we want to look into how Canadian companies can have better access to high-value information with strong economic potential. So this study is part of an economic perspective. We are trying to find out how all this can be used within an economic vision.

Mr. Eaves, earlier, you talked about cross-sectional data—environmental data on the caribou and on logging potential. That was very interesting. One of our recent witnesses asked questions about that.

Earlier, for fun, I used my iPad to research something as simple as taxes. We are currently in the midst of the tax season. The search engine ranked Government of Canada data 13th, and Revenu Québec data 4th.

A Treasury Board representative was saying that, in terms of open data, Canada was doing fine and was well-positioned compared with other G8 countries. Yet I remain skeptical.

How can we increase data accessibility and people's interest? Why would people go on the Canadian government's website data.gc.ca, instead of using a search tool? How can we position ourselves to ensure that our data is used regularly? We will have very detailed big data, mainly with regard to universities, and research and development. How can Canada become a world leader? We are being told this is already the case, but is it really?

Mr. David Eaves: I will answer in English, as I cannot explain all the nuances in French.

[*English*]

There are a few things I would say.

First, I think we have to look at some timelines here that are going to matter. I love the point about students using open data in their research. Prior to the release of the open data portal, you had to pay for StatsCan data. That meant every single student in this country who was doing an undergrad paper or doing research used American data to do all of their work. All of their case studies were American-based, because the American data was free. Up until three years ago, everybody in Canada who did any kind of studies in university tended to gravitate towards American data.

Some of the economic benefits, then, will come from having a population that becomes more and more familiar with Canadian data and what's available. That will take us a process of several years, to have students who are going through college and in their studies beginning to familiarize themselves with what's possible and what's available and then entering the workforce and bringing that to the companies where they work. I do want us to make sure that we have some expectations about how long some of the transformation will take.

That would be the first piece. The second is I think to have a really strategic vision about what the industries are that we want to support that we have data around, and what the policy goals are that we think we can pursue that would enhance those industries. One thing we do know is that data, in and of itself, even when it almost never gets used, can have a transformative impact on how industry operates.

One of the best examples of this was the release of the TRI, which is the pollution data in the United States. In Canada we have something familiar, called the NPRI data. This is the data that every facility in the country must release about how much pollution it released. The very creation of that dataset caused a huge number of facilities in the United States to lower the amount of pollution they were releasing. They became more efficient and more environmentally sensitive just because they now knew that everybody in the world could come and look at what they were releasing.

As a government, it would be interesting for us to think about what was the data that, if we knew we had it, would enable our economy to become more productive and more effective, and then we had a pursuit around how to gather that data and how to share it in a way that industry could leverage or community groups could leverage.

In fact there was an article about that just this morning. There's enormous concern about a potential housing bubble in Canada. At the end of the day, as the lead economist on this issue at CIBC said, we don't actually gather data that would allow us to assess whether or not there is a bubble.

So if we're looking at the various industries that are out there and where the deficiencies on data are, economists and industry experts, people in the industry, are already telling us where we're deficient. I think the question we need to be asking ourselves is this: what role does government have in creating those datasets and curating them to help the economy reach its maximum potential?

● (0940)

[*Translation*]

The Chair: Thank you, Mrs. Day.

[*English*]

Mr. Aspin, you have five minutes.

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

Welcome to our witnesses. Obviously we have a wealth of information to help us with our study.

I was rather intrigued, David, by your analogy of the bulls. I liked that analogy that you're in the pack or you're the leader.

If we want to assume a leadership role in Canada in this data argument, what is the number one factor that we should pursue? Maybe I could get a priority from each of you. Or should we in fact be the leader at all?

Maybe we'll start with you, David.

Mr. David Eaves: I was really hoping we were going to go in reverse order for once.

I don't know whether this takes us out of the pack. This is going to be boring and technical, but the danger we have with open data right now is the thing we're tacking on at the end of the process. You have a government that creates data, analyzes it, does interesting things, and then at the very end we tack this thing on saying, by the way, you have to make it public with the rest of the world.

As a result, our open data initiative has a compliance problem. It's something ministries do.... It's rather like access to information: they don't really want to be doing it. They have to be doing it because the government has asked, but it doesn't actually support a business need right now at the ministry.

My argument would be that if you want to be a genuine leader and want to be thinking about what a government looks like in the 21st century, you have to stop thinking about the data as being an end product that sits at the top or at the end of the process, but rather as being core infrastructure for running government and as the platform upon which all good decisions and all government rests.

I talk about the term "dogfooding", which is when you use your own materials. You don't just publish data and hope other people are going to use it; you dogfood it: you create it and then you build your own infrastructure on top of it.

If we expect industry to be using government data, they're only going to start using it and really believing that we're committed to it when we're using it as well and build our own infrastructure on it.

So the number one thing I would do is go from here to there.

Dr. Renée Miller: I'm responding to both questions too.

I think we shouldn't worry that the government data is not necessarily what is returned in search engines and so forth. I think what we should do is understand to what extent government data has been taken up by researchers and by industry and made into higher quality data.

David alluded to the fact that most researchers in Canada use data. gov data to do their research, and I can attest to that: my graduate students use data.gov data to do their research. But we republish it as richer data, using what we have done because we have gone in and found data that is interesting to us. In terms of that information flow, we have to both understand what data has been taken up by the community and use that understanding to motivate what additional data we provide through the open data portal.

So we use the expertise of the crowd to come back and say that actually we can improve the data we're putting out, to better spur economic growth.

Mr. Jay Aspin: Thank you.

Mark?

Mr. Mark Gayler: One of the areas I would focus on is data that is locked in siloed government data stores.

I'll give you a very simple example, I was working with the Government of Slovenia and I met with their bureau of statistics. One of the challenges they had is.... I don't know whether anybody here has worked with statistical data, but government statistical data is often locked inside very specific, very narrow, and niche statistics systems and is made available in very strange and, even from a technology point of view, almost impenetrable data formats. What was interesting was that the Slovenian bureau of statistics people were very familiar with the Canadian bureau of statistics, so from one statistics bureau to another they had a relationship and were familiar with each other's work. However, from a broad citizen perspective, the citizens really couldn't get easy access to this data.

I think the point was made before that much of the data that's locked in some of those siloed government data stores is really rich, valuable data for citizens, analysts, researchers, and even private entities. It's worth looking at how we get that data out of locked systems and make it more available to end users, citizens, and consumers using common tools and access methods that they have today.

● (0945)

Mr. Jay Aspin: Thank you, Mark.

Now, Ginny?

Ms. Ginny Dybenko: As highlighted in the SSHRC consultation document produced in October of last year, I believe we would benefit hugely in Canada from developing a more forward-looking digital research environment. Specifically that document calls for the development of a coordinated plan to establish and operate a number of world-class centres specializing in data management. Indeed, I think $3 million was already targeted toward the Open Data Institute in Waterloo at last budget.

[*Translation*]

The Chair: Thank you for your answers.

[*English*]

Mr. Bryne, you have five minutes.

Hon. Gerry Byrne (Humber—St. Barbe—Baie Verte, Lib.): Thank you very much, Mr. Chair.

Thank you to our witnesses for giving us a really excellent presentation and perspective on this.

I want to get some feedback on the reconciliation of data integrity within an open data environment.

I think some of the perspective here is that there's a single portal, a single channel, and a single standard for the integrity of the data, so that when you plug into the portal you are getting a set of data that has been tested and that you know to be authentic and for the integrity of which someone is accountable.

Renée, in your presentation you included the aspect of having a discussion about how many beds are available at a homeless shelter. That almost seems more of a blog. If there isn't credibility and authenticity of the data; if it is not tested and someone is not accountable for the data.... Undoubtedly, there will always be mistakes, no matter what standard you create, but there has to be a relatively highly certifiable standard for inclusion into an open data project; otherwise, it could be termed just a blog.

Could I get some perspective on that notion of the single data integrity concept? Governments have one perspective on all of this: they are accountable, or at least they have the capacity to be accountable. A group of community-based organizations with limited funds in a municipal environment has a lesser standard, I think it's fair to say.

Could you give a little bit of perspective on that?

Dr. Renée Miller: Sure.

I would take the example of Wikipedia. From a broad community with a broad set of expertise, you can come down to finding good, high-quality information. Is everything in Wikipedia true? Absolutely not.

I think there are certain things you can use the power of the crowd and the aggregate opinions of the crowd for. Allocating resources in real time as to where you see the resources should go is, I think, a good use for that information.

If you're trying to do longitudinal studies, you probably need some oversight over the meaning of the data and need some curation over the data itself. I think we shouldn't, though, dismiss community-provided data just because it's not curated and may not have the same level of integrity, because it can still provide incredibly valuable information for people, particularly for public workers on the ground. It can give them a sense, a signal about where their resources, their information should go.

That is very different from an historian's trying to pin down exactly what happened. I think we have to weigh the differences that exist.

● (0950)

Hon. Gerry Byrne: Thank you.

David?

I would like to go to our teleconference as well, so just—

Mr. David Eaves: I'll try to be brief.

I agree with you. I think there needs to be accountability, especially around datasets that government is using to make decisions.

I am interested in crowdsourcing, but I think there are incredible limits around how to do it. Even in the example of peer to patent— it's a wonderful example—there are very tight constraints around what it makes it work. It's very easy to use crowdsourcing to disprove things. You may be identifying cases in which something is actually not true, such as identifying patents that are not valid; it's also great to identify datasets that are in error. It's much harder to use it to identify what is actually truthful or is actually a fact.

So one of the nice things is that we should be treating our open data portals as an engagement tool because they're actually a wonderful way to crowdsource errors, not because we want to find errors and make people accountable. There are always going to be errors in the data, so let's surface them more quickly so that we can then get to better quality data faster, so that governments make better decisions with more reliable datasets.

Hon. Gerry Byrne: Chair, could we go to Ginny and Mark?

[*Translation*]

The Chair: I could give one minute to each of you.

[*English*]

Mr. Mark Gayler: Sure.

The comment I was going to make on this is that first of all it's important that you have attribution: who is accountable for curating a particular dataset? That's very key here.

The second thing I would say is that it's very important to have an agreed feedback loop whereby, if you choose to crowdsource the accuracy of the data and you invite third parties to participate in it, you have a feedback loop that enables them to do it effectively, so that people see that the data gets updated within that authority—the authority of source of that data—and that the more accurate data is then reflected on a timely basis.

If you have that feedback loop, I think you then give people confidence that this is a real and a sustainable thing and that the quality of the data is improving over time. It's not something you can do as a one-off or a "let's try it and see"; I think you have to have that feedback loop and sustainability effort on top of it.

Ms. Ginny Dybenko: Mark said it perfectly. I have nothing to add.

[*Translation*]

The Chair: Thank you.

Ms. Ablonczy now has the floor for five minutes.

[*English*]

Hon. Diane Ablonczy (Calgary—Nose Hill, CPC): We wish we had a day with each of you because this is a very rich discussion.

As you know, this is part of a G-8 initiative, and there's been commitment by a number of countries to move in the direction that we're talking about, so I'd like each of you to focus on the internationalization or the global collaboration of the open data initiative. Although Canada may or may not be a leader, there is a mastermind principle that we want to tap into of sharing best practices and learning from others. I'd be interested in your observations on how Canada can improve its collaboration on the open data initiative and where we should put the most focus with our partners.

Mark, why don't you start?

Mr. Mark Gayler: I think one of the things that I would point to, as a way of responding to this question, is that we're starting to see some interesting relationships outside of the traditional, say, government/industry relationship pattern, particularly around open data. One example that I would give you here is how the World Bank is starting to allocate some of its investments in stimulus funding. It now requires countries, nations, and states that it's working with to have an open data policy and to be able to provide evidence that they are being more transparent with their use of data and providing data services to citizens. That's happening sort of globally.

If we look at that as an example, I think Canada can learn from these examples and encourage similar relationships between government and industry participants because the more you join these collaborations together, the more participants you get working together, the richer the data becomes, and I think the impact of the data is more powerful on the community.

● (0955)

Hon. Diane Ablonczy: Ginny.

Ms. Ginny Dybenko: I talked a little bit earlier about establishing roles and processes, and that's hopefully one of the underpinnings to the ODI, but once you have processes and roles established, then sharing with more international participants is facilitated.

Mr. David Eaves: I have about three pieces of advice today. I'm really glad you raised the G-8 because there are some things going on there that I think are quite interesting.

The first thing I would say is that at a very tactical level I worry about some of the ways we might be slipping around our G-8 commitments. In fact, I was very disturbed to realize two months ago that Industry Canada, which shares a database of corporate entities in Canada, used to share who the directors of those corporate entities were. Now the list of those directors goes beyond a $5 pay wall. Rather than being able to see who the corporate directors are, you now have to pay $5 per company. That actually runs contrary to the spirit of the G-8 agreement, which was to make corporate data more transparent to the public.

In fact, if you wanted to spot global corruption, tax evasion, or problems at a corporate level, having a corporate database that is downloadable and accessible is critical to doing that. The G-8 agreed to that, and yet we've gone in the opposite direction. From a tactical place, I would encourage this committee to be looking very closely at Industry Canada's move, to understand why they made this choice.

The second thing I would say is that there is an opportunity in looking at something like corporate data. The opportunity exists in how we harmonize this data across jurisdictions. The question would be on places where we think there's policy importance, corporate transparency, for example. How do we harmonize how we release the data with how the U.K. releases the data, with how the United States releases the data? This would make analysis across jurisdictions much easier, so spotting things like fraud, tax evasion, those types of things, would become significantly easier because the data has all been harmonized.

The third thing I'd say is that if we want to take a leadership role, one of the things that I don't think our G-8 partners are doing, and one of the things that makes everybody in the open data movement very, very nervous, is that there's no protection to our access to most of this data. Our only protection is to do an ATIP request.

If the country wanted to do something that was truly transformational, it would try to figure out whenever it was passing legislation what the core datasets are that make the legislation work. What are the core datasets that allow for the transparency so that the public can assess whether the legislation is working?

The NPRI, which is the data about pollution, is a wonderful example, and that dataset is protected by legislation. Government is required to collect it. It's required to share it, by law, and it is almost unique in that way. I would love to see how we are building datasets that we think are critical to infrastructure or to accountability in this country being protected by law. Users, whether they're corporate or just citizens would know that this data is going to be around—they can actually build infrastructure around it—and they would not be taking enormous risk because the government might get uncomfortable in the future and simply pull that data back the moment it does something that it doesn't like.

Hon. Diane Ablonczy: You're so cynical for someone so young.

[*Translation*]

The Chair: Thank you, Ms. Ablonczy.

Ms. Ablonczy's time is up, but do you have something to add very quickly, Ms. Miller?

[*English*]

Dr. Renée Miller: I think we shouldn't underestimate the difficulty in integrating datasets, and David alluded to it.

The difficulty in taking two open datasets and aligning them, figuring out if two records actually refer to the same corporate entity —if the data that they represent is actually consistent with each other, if one is in metric and one is in imperial units—is an incredibly difficult problem. It is one that requires today, with current technology, many human years of intervention in order to really align two datasets. It's a very difficult problem. My advice would be to strategically pick areas where we can see real economic benefit in aligning the data with other G-8 countries.

[*Translation*]

The Chair: Thank you.

Mrs. Day, you have five minutes.

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

I want to come back to the look of the website, which is very statistical. It looks like a table of contents. Its front page is still displaying information on the CODE event that was held over the weekend in Toronto.

Something Ms. Miller said, I think, caught my attention. She said that our students were leaving for Silicon Valley. We know that Silicon Valley was built from scratch. That's not a utopia. In the beginning, there was nothing there but desert. Yet an amazing computer technology hub was built in Silicon Valley.

Here, some things have been done for the film industry—be it in Ontario or in Quebec—that have contributed to its success. Similarly, a site dedicated to open data could be created. Our students could go to some wonderful locations, such as Gaspésie or Lake Louise, in Alberta. Time will tell whether this is a utopia or not. They could perhaps be gathered in one place and be provided with the necessary tools, similarly to what was done for the people who created Silicon Valley from scratch.

What assets could the University of Waterloo provide to the Open Data Institute?

Will the University of Toronto partner with the Open Data Institute?

Would it be possible to create winning conditions, either in terms of tax credits or something like that, to keep our researchers and students in that field in Canada—at the University of Toronto, among others?

● (1000)

The Chair: We will begin with Ms. Dybenko, if that's okay with you.

[*English*]

Ms. Ginny Dybenko: It's an excellent question, and an excellent observation as well. It's very early stages for the ODI, the Open Data Institute. There was very definitely an inclination to do this inter-institutionally. I think the reason that Waterloo was selected is just because of the amount of success that the Waterloo region has achieved over the past number of years in and around digital technology development and incubation of initiatives.

So as I said, it's early stages yet, but there's an inclination to not only share what has been developed but also to work together with other universities.

Dr. Renée Miller: Yes, and to follow up on that, I think it's a very needed initiative to create something like this in Canada. I have had students who have done start-ups in Toronto, and they eventually all go out to Silicon Valley because they eventually hit a wall and there's just not the culture and expertise and enough people here to draw on to sustain their endeavours. I think we do need something— an initiative as you're saying—that is cross-institutional and cross-provincial to bootstrap that process, to create the critical mass that you need to be able to sustain a culture like that.

[*Translation*]

Mrs. Anne-Marie Day: In your opinion, what would be the winning conditions? What should be implemented?

[*English*]

Dr. Renée Miller: It's the critical mass of expertise. So I think you need the significant investment of existing entrepreneurs. You need it on the business side, you need it on the technical side, you need it on the marketing side. You need a holistic approach to this; it's not just focused around students. We have incredibly bright students but we need the infrastructure that's holistic for creating that entrepreneurial culture.

Ms. Ginny Dybenko: Excuse me for coming in, but MaRS has done a terrific job in Toronto, as has Communitech in Waterloo. We're working together, but what typically blocks us is the lack of venture capital to really take the initiative to the second stage of development.

[*Translation*]

The Chair: Thank you, Mrs. Day.

Mr. O'Connor, you have five minutes.

[*English*]

Hon. Gordon O'Connor (Carleton—Mississippi Mills, CPC): Thank you.

The Government of Canada has identified 14 areas for data that they're supposed to produce, and I think when we talk about government, I believe we're talking about the bureaucracies because they're the big monsters out there that make the data. I'm a bit skeptical. For instance, one of the areas chosen is government accountability in democracy. I can't imagine any government of any stripe is going to pour data out on that, but maybe they will.

We're talking about the government because I don't think private industry provides much information in the sense that they're commercial. My problem with all this data is what compulsion can we give to a government to make them produce data? Because as I said, there are 14 areas here: education, justice, energy...it goes on and on. Governments are only going to provide the data they want to provide.

I'll ask each of you in turn to answer the question. I'll start with David.

● (1005)

Mr. David Eaves: I don't disagree. It's actually enormously difficult to get governments to provide data, especially data that might make them uncomfortable.

My hope is that longer term we could end up in a world where we have a more iterative review of government where we're less driven by the scandal that we can nail someone, particularly a public servant, to the wall against over a particular error. I'd rather end up in a place where we try to iterate around solutions. Actually spotting errors in government is seen as a good thing because it allows us to iterate and make it better rather than something that drives the scandal, particularly if it's the type of problem that's really non-strategic but can end up eating everybody's time.

This would certainly be the place I would love to go. The way I think you've got to go about it, as I mentioned earlier, is that you have to think about how you're going to draft this stuff into legislation because, when you draft it into legislation, it forces parliamentarians and it forces the government to think about what is actually high leverage data and what is the data that will cause us to behave in ways that we want to behave. They'll accept a longer-term plan.

The NPRI dataset, the one around pollution, I think is a wonderful example of a government's potentially embarrassing dataset, and yet, now encoded in legislation, it causes all the incentives both in the private sector and in government to be wonderfully aligned around how we minimize pollution. So I think finding those types of leverage points is going to be critical.

The other thing is, ultimately we do have access to information legislation that should allow us to have access to this data. So if I was going to be thinking about how we prod governments along, I'd just tweak the access to information legislation so that it says that when I make a request, I'm allowed to request a dataset, and you're not allowed to hand it to me in PDF or on printed sheets; you actually have to hand me a disc or send me a file that gives the database in a machine-readable way. At that point I can get the data either way. Aren't you better off making it accessible to me so that I don't eat up a whole bunch of time making requests over and over and over again? Would it not actually reduce the burden on government?

So we actually have access to the data either way. The questions are: how do we make it painless and how do we make it easier for government? So let's put it in the legislation where we have to and then let's improve the access to information legislation.

Dr. Renée Miller: I show my stripes as the bohemian academic here, but I don't think that the stick approach always works. As an academic, I'm required by my funding agencies to always provide any data that I come up with that's federally funded, and I have to make that available.

There's a tremendous number of scientists who feel that their data is their own, and they don't want another scientist to make a breakthrough based on the data that they spent time collecting, right? So we find ways around that. We don't publish the stuff that we think could be valuable. We publish enough to satisfy the law but we keep to ourselves the things that are going to give us that Nobel Prize.

So I don't think just mandating it is enough. Rather, I think what we want is creative government employees who understand that, if they need certain kinds of data and certain kinds of expertise, publishing a dataset may be a way for them to get the expertise that they don't have in-house and may be a way to get somebody else to solve problems that they have. I think that we should view open data as a way of providing solutions into government, which is what I was saying about the flow back into government. Can we provide data out there that, if somebody did something with it, I, as a government employee, could benefit from that and improve my processes? If we can get governments to start thinking that way, I think we'll get better data out there.

[*Translation*]

The Chair: Thank you.

Mr. Gayler and Ms. Dybenko, you have 30 seconds to answer.

Let's start with Ms. Dybenko.

[*English*]

Ms. Ginny Dybenko: I love the carrot and the stick idea. I think in general carrots work way better, so there have to be incentives. The thing that leapt to my mind is efficiencies. So if indeed the individual government department can actually see utilizing that data to drive efficiencies that are required because of reduced funding to their organizations, that would be ideal. Then ultimately they can look for fraud and theft and other kinds of things that are going awry within their departments.

● (1010)

Mr. Mark Gayler: I think I would build on earlier comments. I think this is about reducing friction. Many governments around the world build their open data policies on the freedom of information legislation or access to information legislation that they have in place. I think this issue is really about reducing friction and making access more available.

If you can publish the data and make it consumable and easily available, then you should do that and not hide it behind a lot of bureaucracy that may be unnecessary in a particular case. That's not necessarily always the case, of course, but I think you should reduce friction where you can, and that encourages people to publish. It encourages citizens to consume.

The other thing I would say to finish is that it's important to understand that open data is not WikiLeaks. These are separate things, so I think it's important to make that distinction as well.

[*Translation*]

The Chair: Thank you.

Mr. Blanchette, go ahead for five minutes.

Mr. Denis Blanchette: Thank you, Mr. Chair.

This time, I will address Mr. Eaves.

At the heart of your presentation is really what Mr. Gayler calls organizational culture. In other words, data is not simply being produced for the sake of data production, especially in the government. If that were the case, we would say that taxpayers' money was being wasted. This has more to do with a way to work, an approach.

Do you have an example of what should be changed in organizational culture? For instance, what trends and methods within the federal government would help gradually build a data sharing culture? Of course, I am not talking about confidential data, although this is a very important issue we could come back to.

As it has already been said, departments often have to operate in isolation. However, what is happening in the public administration is also happening in private companies. You would be surprised to see how isolated companies can also be in terms of their operations, if they are even remotely large.

I would like to hear your thoughts on this matter.

[*English*]

Mr. David Eaves: That's an excellent question.

I agree with the carrot and the stick. I'm a big fan of carrots and I've used them with this government, but I'm also a believer in sticks. So the question is how do you manage both.

And the question of culture is enormously important. It doesn't actually matter how many rules you have in place. What I'm most interested in is how you create incentives for public servants to want to share more. And I think there are a couple of things that we have at our disposal on how to make that happen.

First, most public servants have grown up in an era where ministerial orders or deputy ministerial orders have been to not share anything because there's only risk involved in sharing. And how do we begin to crack that? How do we begin to change that culture?

The first one is there are lots of examples around how transparency can advance a policy agenda. But my most favourite example, and the one I always give in talks, is around restaurant inspection data. It turns out in L.A. they decided they were going to publish restaurant inspection results on the front doors of restaurants. The moment they started doing that, you had more people going to restaurants that had better results and fewer people going to restaurants that had worse results.

And wouldn't you know it, but it turns out that as a result, you also had fewer people ending up in the emergency room with food-borne related illnesses, which is the most expensive point of contact in the health care system.

So if you want to drive a policy outcome of reducing health care costs, it turns out publishing restaurant inspection results in a useful manner is a great way of driving that. So we have a whole bunch of examples where transparency and sharing data actually advances policy agendas. So driving those stories through the public service and causing public servants to think about where transparency is actually strategically in your interests would, I think, cause people to begin to re-evaluate why they should be sharing.

And then if we could get them to be thinking about that, we might begin to crack the door open a little bit more where they see that actually the risks of sharing this other information doesn't quite feel as high as it did before. Now they see that there are actually benefits in these policy areas where sharing created these outcomes they liked.

So that's probably the direction I'd try to go.

Mr. Denis Blanchette: Of course, that is part of the issue, but we need to look at all the aspects. You used an example where the stick method is being used in restaurants.

Earlier, you talked about four points with regards to open data. In sum, points two to four illustrate how difficult it is for public administrations to truly use all the data they have. You are basically saying that people could help with that a bit.

We could perhaps take things a bit further by saying that data sharing could help the government fulfill its own mission. It's amazing to see how much diversity human curiosity can create. This applies in terms of the economy, but since all governments also have non-economic duties, the overall government mission could benefit from this.

● (1015)

Mr. David Eaves: And I would say that the product recall data.... The federal government has an app on product recall data, which strikes me as a great example. When you look at the data that underlies that app, only some of it is available in the open data portal. But you have to be the most OCD person in the world to fire up an app to check whether or not a product has been recalled before you buy it.

But if you made the data available, there's kind of a long tail of users. There are people who have dairy allergies, people who have wheat allergies, or people who have kids. They're not all going to go to this app. You've tried to aggregate all of those users into one app, and they're not going to use it.

But if you actually had the data available, organizations and associations that represent those people or that serve those people might grab that data and provide it to them in the places where they actually look and they actually read. You'd actually end up having a much higher policy impact around that dataset than having the government create an app.

The Chair: Thank you, Mr. Eaves.

Now we'll go to Madam Brown for five minutes.

Ms. Lois Brown (Newmarket—Aurora, CPC): Thank you, Chair.

Witnesses, I'm only a visitor to this committee. I'm filling in for one of my colleagues, but I find this discussion really exciting, and my word, it's phenomenal where it can go.

Ms.Dybenko, I had the opportunity to visit Communitech in Waterloo, and I visited the MaRS building in Toronto, and so I have been exposed to some of this. I don't consider myself any sort of a technical wizard by any stretch of the imagination, but this absolutely fascinates me.

Ms. Dybenko, in your opening comments you talked about this being our next natural resource. Canada is incredibly blessed with natural resources. If this is our next natural resource, what kind of transformation is this going to have for our economy? We know that when we started building cars we stopped building buggies, so the opportunities are there.

You talked about the challenges with getting venture capital. What incentives can we put in place to help venture capitalists become engaged?

Mr. Gayler, maybe you have some comments from the perspective of Microsoft. I'm sure you have a long history of looking at how new technologies come on board and how that is aggregated into the business.

Could you both comment ?

Ms. Ginny Dybenko: I'll start by talking about what I think is really necessary for the transformation of this opportunity for Canada, and that is our young people. I believe it's a matter of doing the best we can to produce the knowledge worker of the future. By that I mean not necessarily just engineers or just computer scientists or just mathematicians, but also bringing in arts and an understanding of the humanities and social sciences as well to create individuals who are fluent and creative on all levels. Only by doing that are we going to create the natural cycle.

Renée spoke very compellingly about this. We have so many bright young people in Canada, so the ideas are not the issue, they truly are not, but if we can prove, as we have in the Waterloo area with the kind of universities and colleges we have, that a huge labour force is being created in that region, then the rest will come. Venture capitalists from Silicon Valley are already turning their eyes to Canada. Unfortunately, they now attract the young people to come down—my son among them, I should say—to work in California, but over time, if we're able to prime the pump enough, my belief is we can create our own ecosystem here in Canada.

Mark.

● (1020)

Mr. Mark Gayler: Yes, I think I would totally agree with that. I think there are already examples of where we've seen transformational things being done with government data that's being used by commercial entities.

One of the examples that David will be very familiar with was back in the early days of the City of Vancouver's open data. A local firm of architects here used some open data to predict water levels and the impact that would have on the local downtown area in Vancouver. The comment that I loved about that was that it wasn't that we couldn't get access to this data before as a commercial entity; it was that we didn't know who to talk to. We didn't know how to get this data. It was just too difficult. So the fact that the data was published enabled us to do something new and innovative with that information that helped our business and ultimately benefited our customers.

I think this is a very simple example of how this data can be used in partnership with commercial organizations. Of course data has been commercialized for years. This is nothing new. There are whole industries built on it—advertising, marketing, demographic data, retail analysis data, and organizations around that. This is something that the commercial industry is very familiar with and makes money with in different ways.

I think the key here is to establish the partnerships between the government data sources and those third party commercial sources. Once you start combining the data in this way, transformative things start to happen.

Let me give you another example, again from Vancouver. When Vancouver shared parking data initially, it shared parking data around the use of city-owned meters. It didn't have the commercial parking data because it didn't own that data, so it didn't publish that. The commercial parking companies weren't publishing their data because they were effectively in competition with each other and the city. It actually took a third party developer, an independent developer, to actually take the published City of Vancouver data and the commercial data and combine them. This is something that naturally neither the city nor the parking entity would do off their own bat. It took a third party independent developer to do that.

I think this is the value of these partnerships. By sharing the data in the first place, you can create these chains of reaction. The more partnerships you bring to bear on this, the more valuable the data gets.

[*Translation*]

The Chair: Thank you.

[*English*]

Mr. Mark Gayler: The last thing I would say—sorry, I have just one more point—is that we have a sea, an ocean, of data coming down the line. If you guys think the data we're sitting on today is it, it is not. Even the data that the government alone will produce is nothing compared to what we call "big data", which is the wealth of research, scientific, and commercial data coming down the line. The potential for combining that and creating more economic value is absolutely huge.

[*Translation*]

The Chair: Thank you.

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

[*English*]

Hon. Gerry Byrne: Thank you, Mr. Chair.

I was interested to hear, Mr. Eaves, one of your comments. One of the objectives, I guess, of this study would be to assist the government in moving the initiative along. You mentioned that there was one practice, at Industry Canada, that you found contrary to the spirit of the G-8 commitment that Canada was making.

Are there any other circumstances or practices that you may be aware of that you could share with the committee, where you sort of question or want to raise whether or not the Government of Canada is running contrary to the G-8 charter or to the spirit, generally speaking, of what an open data portal, an open government, is supposed to be about?

Mr. David Eaves: I think there are numerous examples and I don't say this to pick on this particular government. As previously said, all governments get cautious around sharing data and information. One of the reasons we have a competitive political process is to keep people honest.

There are two points I'd love to make.

One is I think there are all sorts of macro examples. For example, people were curious about the F-35 spending. Parliament demanded documents, and then those documents were produced in 100 boxes and printed out. What this meant is that if you were someone who was actually interested in learning anything about this, you couldn't do keyword searches of that. You'd actually have to go through and read every single piece of paper. This is what we call hiding things by making them available but in formats that are completely useless or very hard to use.

I think those types of behaviours are examples of a government where they're not actually interested in transparency and they're certainly not actually interested in sharing information. I'd be looking for ways that we could curtail governments from doing that in the future, so when I ask for a document, I get it in a machine-readable way so I can do keyword searches and go and find the interesting things.

I'd love to see more around actual budget data being made available for downloads, so that people can actually.... How do we make the government more legible to the population so they can see where money's getting spent and they can see how their tax dollars are being used? I think all governments have a long way to go, but we, in particular, have a long way to go. And the U.K. is actually a very interesting example around this. They've made all spending data down to £500 downloadable and publicly available. So you can actually go and see how each department is spending its money. This has been interesting to the public, but I think it's actually been very interesting even to the people in government, because they can actually now access how the money's being spent in a very direct way. Their staff can, and they can do their own analysis.

I think even government officials, elected officials, have found this dataset to be very interesting.

Another example would be the access to information requests. I see no reason that when someone makes an access to information request, that document is not being put in a publicly available database so that if I now want to do the same access to information request, I'm not going through the whole thing all over again. I can go and scan the ones that have already been done and you can save me a whole bunch of time, but more importantly, you can save government and taxpayers a whole bunch of time by not having people running around and gathering up the same documents and doing the same assessment all over again. I think it's in everybody's interest to make that happen.

So those are cumulatively the things that I would say. At a high level, a recommendation that you could make.... One of recommendations we made in Ontario with the open government task force was actually creating rules around procurement. So you say any system that is bought by this government that is going to produce data must have in its procurement demands, as part of the specifications, the ability to extract that data easily. So if someone comes along and asks for something, there's no longer this question of, "Oh, well, the system's old or it's hard to use and we can't extract it." We've actually built that in as a requirement so we make it easy to extract information.

Changing procurement rules is one of the most powerful tools that you have at your disposal to think about how we can make data more accessible to people.

● (1025)

Hon. Gerry Byrne: On that specific subject, would you be able to expand upon that through a letter to the chair of the committee? Because I think that's something that, unfortunately, 30 seconds is not going to cover well.

Mr. David Eaves: Yes, agreed.

Hon. Gerry Byrne: Also, in that same vein, if there are other specific items that you really want to highlight to the committee... and I'm saying this to all the witnesses. If there are any other specific items that you may wish to highlight to the committee, you would have an opportunity, and I'm sure the chair would agree, to be able to write a letter to the chair, and that would be considered, I'm sure, by the committee, to be part of the witness testimony for the purposes of our report writing. I think you'll get a consensus on that.

Could I, then, just share the floor with some of the other witnesses, Mr. Chair?

[*Translation*]

The Chair: Yes, certainly.

I remind you that you can always submit documents to us through the clerk, and we will take them into consideration during our study.

Would anyone like to add something?

[*English*]

Mr. Mark Gayler: Sorry, are you asking for final comments? Is that your question?

[*Translation*]

The Chair: No. I was giving you an opportunity to answer Mr. Byrne's question.

There will be other questions anyway. Just before we adjourn the meeting, I will give you an opportunity to add a few words.

Mr. O'Connor, it's your turn.

[*English*]

Hon. Gordon O'Connor: Thank you.

For 40 or 50 years, technology has continued to generate machines that provide more and more information to people. If I think back to 3,000 years ago, in Athens they had the Acropolis. All the citizens would arrive at the Acropolis and discuss issues and pass judgment, etc. But as populations got bigger and more organized, we went into silos and categories, etc.

Nowadays, we have trillions and trillions of pieces of information that citizens can't get at. I'm looking at the trend lines of what's going on here, and it seems to me that if open data actually becomes open data throughout this country and other countries, we'll be moving back towards the Acropolis again. That will affect governments, the organizations of governments, etc.

For your final thoughts, I wonder whether you have any opinions on how open data is going to affect governments.

I'll start with Ms. Miller and run the other way.

● (1030)

Dr. Renée Miller: That sounds good.

I think open data is going to lead to an open government revolution. It has the potential of making governments more participatory, down to a much finer grain than we have right now.

This is my call to ensure that, even at the beginning of the open data revolution, we make channels for the flow of that information back into the institutions. I think we can make better decisions based on data. We're seeing that revolution in education. We're seeing data-driven education. We're seeing data-driven medicine. We're going to see data-driven governments where we're using past best practices in order to bring that back into government. I think it's an exciting opportunity.

Hon. Gordon O'Connor: Ms. Dybenko.

Ms. Ginny Dybenko: I would like to point once again to engagement. I think lack of engagement of the citizenry is the biggest challenge that government faces today. I would see open data as a very useful tool to not only speak to the electorate but also to get opinion from them, and in doing so to get them involved in government affairs.

Hon. Gordon O'Connor: Mr. Gayler.

Mr. Mark Gayler: I think this is a very important point.

Going back to my point that I made earlier, the amount of data is only going to increase. I think where government can help is by providing easy access to that data, based on the way that consumers and citizens want to access it. Don't force them to go to a particular portal. Don't force them to use a particular type of technology to access that data. Give them the data in a way that's open, as a service that they can consume in the way that they want to.

Second, I want to give you two more examples about how transformative open data, combined with the sheer growth in the amount of data, really can be. The first one is with the City of Barcelona, where they have a program where they share bikes across the city. They provide bikes that can be taken from one transit stop to another. The City publishes data on the availability of these bicycles, which obviously fluctuates depending on time of day and also whether they are having events or not. The City combined that data with social media data. By tracking sentiment on social media, through Twitter, Facebook, and stuff like that, they're getting instant dynamic feedback about the citizenry, regarding the availability of these bikes, if there are enough bikes in a particular area, whether the bikes are of good quality.

What you can see now is government combining their data with the huge volume of big data that's out there and growing every day, and—to the point that I think Ms. Miller made earlier on—providing this to improve decision-making. I think this is where we see this going: an increase in data, the increase and ubiquity of technology, engaging consumers and crowdsourcing to enable government to engage and make better decisions.

Hon. Gordon O'Connor: Mr. Eaves.

Mr. David Eaves: You guys actually have a much bigger responsibility than I think people would let on. I'll maybe share that more in closing.

As an illustration of that, I love what Clay Christensen, a famous author and Harvard Business School professor, says, that when you destroy the value in one part of the value chain, it migrates to another. If you make software free, the value doesn't disappear; the value shifts over to the services, and now servicing the software is where the money will be located.

I think there is something similar around politics. When you knock the politics out of one part of the political chain, you don't destroy the politics, it just migrates to a different part. This is one place around data where you have an enormous responsibility. In a world of open data, you can presume that the information or data that government is creating will be made public. That used to be a political decision. It used to be a decision where a minister could say whether or not they'd share this or that data. But if we now presume that all data will be made public, we've now taken the politics out of that part of the chain. That doesn't mean the politics disappears; it just moves to a different part of the chain.

So I think one of the questions politicians will increasingly be asking themselves is, "If the data gets created, if that means it will be shared, I will have a lot more scrutiny over what data will get created in the first place." Some people would argue that this is what happened around the long-form census, that we actually didn't want to have the data created in the first place because that meant there would be questions asked that government didn't want to be asked, or policies pursued that people didn't want to get pursued.

The data will become more political the more open it becomes. This committee needs to think about what the ramifications of that are.

I would follow it on to say that I would be very careful about presuming that data will lead to better decision-making. Having data-driven decisions does not mean a better decision. I only need to show you a map of a congressional district in Chicago, that looks like a tiny little filament running through nine different neighbourhoods, that makes absolutely zero sense. The reason that congressional district was created was to produce a very specific outcome.

That was a data-driven outcome; I want to be clear. You could never create that congressional district unless you had phenomenal data about who was living in what types of neighbourhoods and who you thought those people were and how they were going to vote. That was a data-driven decision. We could argue about whether it was a better decision or not. It was a better decision if you were trying to create the outcome that the decision created.

So we're not about to depoliticize any of this, and we're not about to end all of this. You guys have an enormous responsibility to be thinking about what the politics of data are, even if you're just talking about economic data. I don't want you to lose sight of that.

●(1035)

The Chair: Thank you, Mr. Eaves.

[*Translation*]

Mr. Blanchette, go ahead.

Mr. Denis Blanchette: Thank you, Mr. Chair.

Mr. O'Connor's approach is very interesting, as is Mr. Eaves's answer. That raises other lines of questioning, such as in terms of usage guidelines for open data. This goes beyond data sharing.

Mr. Eaves, I understand that you are in favour of full data sharing. However, there are some obstacles that cannot be ignored. For instance, Ms. Miller said that intellectual property needs to be protected. When it comes to companies, it would be difficult for them to share data that shows their weaknesses relative to their competitors, and that's perfectly understandable.

If the federal government decided to go ahead with data sharing, what kind of safeguards should the government or a large company implement to protect intellectual property and respect privacy, among other things?

I would like each of you to make a quick comment.

[*English*]

Mr. David Eaves: I agree; I think there are criteria that we need to be worried about. The first would be making sure that we're not releasing data that has personal information in it, or at least not releasing data that has personal information that we don't want to be releasing. I'm quite happy to release who the directors of companies are, but I certainly don't think the government should have a role in releasing my personal health care data. That would be one.

Around IP, intellectual property, I get that it's a concern, although my personal feeling is that if something is government-funded, then tax dollars went to create that. It is already a public asset. By assigning the IP to a private company, we're actually doing a disservice to Canadians. So I'm not so concerned about the IP element of this.

I'll leave it at those two. I want to make sure everybody else has a chance to share something.

Dr. Renée Miller: In terms of privacy, I think it's important to note that privacy is not just about releasing individual data, releasing the name of somebody and their medical status. Aggregate data actually can also release personal information. It's important that statisticians and so forth are very careful about the aggregate data, that you can't learn information about individuals from aggregate data. There are well-known cases of this where you can infer information about a public figure's medical records from the release of aggregate data. I think that's incredibly important in terms of privacy and what we release.

In terms of IP, I think there are very good open licences out there. Just as we have open licences for software, there are open licences for data that retain the IP on the information and let it be used for the public good. I think those licences are evolving still, but there is a good trend in those licences that lets you release your data while still maintaining ownership over that data. That's also an important point.

●(1040)

[*Translation*]

Mr. Denis Blanchette: What do you think about that, Ms. Dybenko?

[*English*]

Ms. Ginny Dybenko: Obviously privacy issues are huge with big data, never mind just open data. I have often felt you almost need a magna carta along with big data, and that would include notice that predictions based on big data had taken place, and what's been predicted, and how; an opportunity for feedback or a hearing, if you like, if the individual has an issue with that; and then finally audit trails that record the basis of predictive decisions.

I think although not all open data is obviously personally attributable, there is a need to consider those three basic tenets, if you like, across big data as well.

Mr. Mark Gayler: I think my colleagues have said it very well. We have to be careful, obviously, about sharing open data with personal, private information. By and large that does not happen except in specific circumstances.

But there is data that can be shared that would historically or traditionally be considered sensitive that has tremendous value, and I'd like to give you a very quick example of that, particularly around health data and crime data, for example.

Citizens are very interested in crime data. Where is crime occurring, and how frequently is it occurring? You share the data in terms of trending. In statistical data you don't obviously give intimate details around who committed the crime, or who the victim of the crime was.

A real benefit of this happened recently in the U.K. where in one of the cities they analyzed the influx of patients in the city hospitals over a weekend. They combined that with violent crime in the area around bars and restaurants in that city. Between these different agencies—the hospital, the police, and bars and restaurants—they took a collective decision to start serving plastic glasses in bars and restaurants because they realized I think it was up to 40% of the emergency cases in the hospital were caused by fights and violent activities occurring around bars and restaurants.

So here's a great example where you're sharing data in a way that's not sharing personal, private information, it's anonymized data, but it enables a decision to be made that probably wouldn't have happened had that data not been shared in the first place and shared in particular among different government agencies.

[*Translation*]

The Chair: Thank you for your answers.

Would anyone like to add something?

Go ahead, Mr. Eaves.

[*English*]

Mr. David Eaves: Just to keep it brief, I would say I hope I have impressed on this committee the responsibility it has when thinking about this issue.

A lot of people are going to come before you and talk about the huge opportunity and the money. This stuff matters. I want you to be thinking holistically about this, not just about software, but about analysis, and about how we can be using this to change the way government works and increase the productivity and effectiveness of government.

And also about the responsibility around the politics of data, and that you can't escape that. You need to be thinking about that.

I am here at the disposal of this committee. You should feel free to call upon me any time, and if you want information or ideas, you would be more than welcome to write me an email in French or in English, and I will respond as quickly as I possibly can to get you the best possible information to inform your decisions.

[*Translation*]

The Chair: Thank you.

Ms. Miller, did you want to add anything?

[*English*]

Dr. Renée Miller: Thank you very much for the opportunity to testify on this panel. I think it is a very important issue. I'll reiterate that. And I think you are going to be amazed at the kinds of things Canadians can do with your data.

We have a lot of creativity out there in the community, and I think you will have some wonderful surprises in the future and great things.

●(1045)

[*Translation*]

The Chair: Mr. Gayler, did you want to add anything?

[*English*]

Mr. Mark Gayler: Thank you very much for inviting me to participate. Obviously the area I specialize in is technology so if I can be of assistance to committee members around technology issues, I'd be very happy to do that.

In closing I think it's a tremendous opportunity. We have a long way to go in the world with open data, the way it's being published, the way it's being consumed, and how transformational it can be. I think Canada has a wonderful opportunity to be even more of a leader in this area, and so I'm looking forward to what happens next.

[*Translation*]

The Chair: Ms. Dybenko, you have the last word.

[*English*]

Ms. Ginny Dybenko: As the others have said, thanks so much for including me in this discussion.

To reiterate, I think that although there will be huge challenges associated with grappling with the obvious benefits that could come out of open data, I strongly believe that one of the most important benefits will be the engagement of a populace, particularly a younger demographic, that today feels very disconnected from government processes.

Thank you again.

[*Translation*]

The Chair: I want to thank the committee members, as well as Ms. Miller, Mr. Eaves, Ms. Dybenko and Mr. Gayler for their time, especially those of you who are on Vancouver time. I am sure their expertise will greatly benefit the committee during the rest of its study.

The committee will meet again next Tuesday.

The meeting is adjourned.