HOUSE OF COMMONS
CHAMBRE DES COMMUNES
CANADA

Standing Committee on Government Operations and Estimates

OGGO • NUMBER 020 • 2nd SESSION • 41st PARLIAMENT

EVIDENCE

Thursday, April 10, 2014

Chair

Mr. Pierre-Luc Dusseault

Standing Committee on Government Operations and Estimates

Thursday, April 10, 2014

● (0845)

[*Translation*]

The Chair (Mr. Pierre-Luc Dusseault (Sherbrooke, NDP)): Good morning everyone.

Welcome to meeting No. 20 of the Standing Committee on Government Operations and Estimates. We are continuing our study on government's open data practices. We have several witnesses with us today, starting with Ms. Lyne Da Sylva, Associate Professor, School of Library and Information Science, Université de Montréal.

We also have with us via videoconference from Oxford, Mr. Richard Stirling, International Director, Open Data Institute, in the United Kingdom. From Paris, France, we have Ms. Barbara-Chiara Ubaldi, E-Government Project Manager, Organization for Economic Cooperation and Development, and via videoconference, from Sheffield, Ms. Joanne Bates, Lecturer in Information Studies and Society, at the University of Sheffield in the United Kingdom.

As is our custom, I will remind the witnesses that they may make opening remarks for a maximum of 10 minutes. Following that, committee members will ask questions of the witnesses.

With no further delay, I would like to welcome Ms. Da Sylva, who is with us in the room today. We are ready to hear your opening remarks as they relate to our study on the government's open data practices.

Thank you for being with us this morning.

Ms. Lyne Da Sylva (Associate Professor, School of Library and Information Science, Université de Montréal): Thank you for this invitation.

I was told that it may be a good idea for me to introduce myself first in order to assist you in your questions, which I will be happy to answer afterwards in either English or French.

I am a bit of a strange beast. My training has been in several areas. I completed a Bachelors in Mathematics and Computer Science, after which I did a Masters in Linguistics and a Doctorate in Linguistics, with a focus on artificial intelligence. This lead to my work on what is called natural language processing, that is, the use of computers to understand texts written in French, English, Italian, and so on, for the purposes of translating them, and automatically correcting or processing them.

I worked, among other areas, in the private sector as a natural language processing—or NLP—software developer. I am currently a professor at the School of Library and Information Science. I was hired under their digital information management envelop. That is really our main theme, that is digital information.

My current expertise is in two areas. I work in the area of natural language processing as it applies to document management. On the other hand, I'm focusing more and more on digital libraries for document collections, whether they be library documents, archives, museum document or other kinds of documents, and their access functions. Certain websites and databases would also fall under digital libraries. Collections and data sets are an example of digital libraries. I am particularly interested in these issues from that perspective.

I have based my opening remarks this morning on the five questions I received. I just wanted to give you an introduction first.

We talk about open data, linked data, linked open data, RDF data. They don't all mean the same thing. There are more or less open types of data. It is not enough to publish data for that data to serve as an excellent example of open data. An excellent example, the best format, is the RDF format which is user-readable and operable.

There are several jurisdictions that will publish data, but that data is not necessarily in an easily usable format. There are degrees of usability in what is provided.

Another term that is used is big data. Once again, that is something different. That term refers to research based on massive data. Even though it is different, one can only expect that the advent of enormous quantities of data will significantly change people's attitudes towards knowledge and the use that can be made of that knowledge. That will change everything.

The first question was how the Government of Canada compares to other jurisdictions, in Canada and abroad. I compiled some data in a table that is in the notes that I gave to the committee. It includes data on the availability of data from governments in Canada and abroad.

The results are quite variable both in terms of the number of data sets and degree of real openness. Some governments publish their documents as zipped PDF images, which is not necessarily the most desirable format for open data.

I am not going to go over the table in detail. I would say quite briefly though that the United Kingdom is known internationally for its extensive publication of data, including a large quantity of truly open RDF data. The number of data sets is approximately 17,000.

Canada's number of data sets is over 190,000, which is higher. On the other hand, Canada's data is less open. There are more zipped files, geographical maps, for the data. There is currently exactly one data set in RDF, which is a little sad. The table describes much of the data and it would be too long to go over that now.

I have also pointed out a website, Linking Open Government Data, which has ranked a number of countries. It puts Canada in second place for publishing data sets.

● (0850)

Clearly that ranking is based on the number of available data sets, but not necessarily on the ease with which those data can be accessed.

I am now going to answer the second question, that is, how does this compare with what the private sector is collecting and making available.

Obviously, public administrations do not publish the same kind of data. They publish information on the activities of the public administration, public services management, natural resources, etc. The private sector is much more reticent to share their data. The reasons for this are quite obvious. Businesses are afraid of losing their competitiveness. Many incentives are offered to the private sector to meet certain consumer expectations, because consumers want societies to be more transparent and environmentally responsible, among other things. The public sector acknowledges that this can lead to some risk sharing. For example, insurance companies and pharmaceutical companies can benefit from other businesses' data in order to improve their competitiveness.

The third question is how can proper use of public data stimulate job creation and economic added value? The availability of open data clearly encourages the development of various applications. However, one should not only think of the money that can be made. Rather, one should consider public data as a new public service, just like libraries. That's the parallel that should be made, rather than considering this as an economic added value for the purpose of immediately making money.

The fourth question is how we can make sure that there is accountability and transparency, while being prudent on privacy issues? The distinction must be made— and others do make this distinction—between collective data, that can be open data when it is anonymous, private or personal data, which should be available to the individuals but not to the public, and transformed data, which can be anonymized before being published. It's important to define a series of confidentiality principles in order to manage this.

The last question is how we can make sure that public data serves the needs of the population of Canada? I have identified four potential ways of doing that. We can have new public officers, for example a chief data officer or something similar. Obviously there has to be a public and transparent official policy along with new structures, such as citizens' advocacy groups. Furthermore, we need to include the documentation sectors, that is, library scientists and archivists, who are used to managing data and taking into account user needs in order to improve their services.

Thank you.

● (0855)

The Chair: Thank you for your opening remarks.

I will now give the floor to Mr. Stirling, from the United Kingdom. He is the international director of the Open Data Institute.

[*English*]

Thank you for being here with us today, Mr. Stirling.

You have 10 minutes to make your presentation.

Mr. Richard Stirling (International Director, Open Data Institute): Good morning.

I want to start, in the same way as the other witness, by giving a little bit of extra context about me. I was instrumental in the U.K.'s rollout of open data, working in the Cabinet Office to write the initial policy and also doing the first 12 months of delivery and release of data.

To my mind, a political opportunity in open data has been created by the work and resolution at the G-8 for the G-8 open data charter, which was signed by all G-8 countries last year. This means that the biggest economies in the world will start releasing more and more data, and they're releasing more and more data in a way that is useful. They're releasing data around the core information assets, around such things as locations, times, environmental information, in a way that can be combined with other data sets and can also be combined across borders.

The first question this committee asked was what the value of this is. It's a huge opportunity. The McKinsey Global Institute published a report that put the value of this market at $3 trillion globally. Other reports cover smaller geographic regions and are of similar orders of magnitude. So the opportunity is enormous here.

The Open Data Institute, which I'm from, is a not-for-profit initially funded by the U.K. government. We were created to accelerate the benefits in the U.K. economy. We're here to bring economic, societal, and environmental benefits from open data, to answer the “so what?” question. We're here to make sure that there is some impact.

The way we do that is through training people, building capacity. We foster start-ups in our space. We have 10 open data start-ups as part of our program, employing 50 people—they were employing about 20 when they joined the program—and we convene academic, private sector, and public sector communities around particular problems and challenges and sectors.

In the last 18 months, because we've only been going 18 months —it's still a very new sector—we have a few examples of ways in which that $3 trillion number stands up. One of our observations was that there were a lot of enormous macro benefits and big numbers attached, and there were lots of tiny companies, but there was very little in the middle. So in the last 18 months we've worked with other people to identify £200 million cash savings in our National Health Service gross budget. We've mapped out the corporate structures of the investment banks in the U.S.A., drawing together information from three different regulators to provide insight in two months that none of those regulators had themselves. We've worked with the Bank of England, the major financial regulator in the U.K., to prove that you can take a data-rich, regulation-like approach to a market, in the new peer-to-peer lending market, which is now at $1 billion a year.

Many of these examples come from taking open data or data that was previously closed and combining. Many of the really interesting things happen at the intersection of open data and closed data, or open data and big data, or open data and personal data.

That leads into some of the questions you were asking. How are businesses approaching this? Are governments ahead of business? Well, they are, at the moment. This is one of the few sectors in which the government is slightly ahead of industry. Through our work of convening industry and through our corporate membership program, we talk to an awful lot of businesses about how they're approaching the open data challenge and how they view open data as an opportunity.

● (0900)

It feels as though the conversations we're having with them are very similar to the conversations people were having inside governments about five years ago. We're starting to see the first big businesses releasing open data as part of their business as usual.

There are some great examples from the U.K., often brought on by adversity. Tesco, one of the major retailers, is committed to publishing open data about every bit of own-brand food they create. They're doing that to show the consumers what they're eating so they can rebuild trust in their products.

One of our members, Telefonica, is looking to release some of the population data they know from the way mobile phones move around London during the day. We actually used that in one of our policy analyses to show the type of population in London and to show how that impacted on some of the resource allocation in public services and fire stations.

The next question you asked was around anonymization and how you can protect people's privacy in a landscape where open data is becoming ever more prevalent.

One of the organizations we're a member of is the UK Anonymisation Network. They do fantastic work to check people's work and they ask all of the questions around whether people have taken the right steps to protect people's privacy before any large data set is released. The £200 million savings that I mentioned earlier is drawn from a data set that contains every prescription written in England and Wales. That would possibly disclose personal information, but the NHS Information Centre has already taken the steps to check that they've done their anonymization well and also that it can then be checked by this peer-review process, the UK Anonymisation Network, through which statisticians check that all the right things have been done.

There is something called the open data barometer, which isn't quite large enough to be seen. You were asking how Canada compares to the rest of the world. Well, this is a nice visual representation of how Canada compares to the rest of the world on the release of data, particularly in terms of the data sets that are being requested and signed up to in the G-8. You can see that Canada is currently eighth in the world in the release of data. It has particular strengths for some of the core data that's being released, but it still has a little way to go on getting some of the social and economic benefits from the release of data.

I'd be very happy to send a link to this site to the committee so that everybody can see it.

In terms of how Canada could move up in the rankings and what my ideal ask would be, I think there are a couple of core data sets that could be usefully examined as to whether or not they could be released. We've done some work to try to make it easy for people to build services on the back of open data. An awful lot of work has gone into the technical standards around data release, and the previous witness talked about that.

We've put some work into the social side of data release. If you believe that open data is a raw material for the digital age, then as is the case with any raw material, you care about certainty of supply, you care about how often you're going to get a release, and you care about how much time and effort people will put into customer engagement, talking to you about how you use the data and what things are important to you. That's something we've tried to codify with open data certificates. We've given that away to the world.

The final thing I would leave you with is that this is a global market. It would be great if we could start tackling some of these challenges globally.

Thank you very much.

● (0905)

[*Translation*]

The Chair: Thank you, Mr. Stirling.

We will now go to France. We will be hearing from Ms. Ubaldi who, as was stated earlier, is the e-government project manager for the Organization for Economic Cooperation and Development.

Ms. Ubaldi, you have 10 minutes for your opening remarks. You now have the floor.

[*English*]

Ms. Barbara-Chiara Ubaldi (E-Government Project Manager, Reform of the Public Sector Division, Public Governance and Territorial Development Directorate, Organisation for Economic Co-operation and Development): Thank you very much.

I would like to start by giving you a very brief oversight of what we do at the OECD, what we've been doing with open data with the 34 member countries of the OECD and increasingly with the non-member countries. I would like to clarify that we work with governments for our open data project, which concerns the release of data in open formats by governments. So we don't work with the private sector.

Our project started about two years ago, and I think it's important to underline that we started the project at the request of the governments. We have a group of CIOs who represent the governments of the 34 member countries of the OECD, including Canada, who asked us to look a little more in-depth at the strategies, implementation efforts, and the impact of creation efforts that they were putting in place. We produced a working paper highlighting key issues, and we conducted a data collection in 2013 across the countries to be able to see in more detail what governments were doing in terms of being strategic, developing quotas, but also in trying to achieve the value they expect to get out of their open data strategies and initiatives, and to measure these impacts.

I think it's very important to underline that what we found out was that within the community of practitioners, both inside and outside the government, there was and there still is some confusion when it comes to definitions. This means there is much overlap with the activities, for instance, of the freedom of information movement vis-à-vis the open data movement, the discussion on access to information and open data, and how they complement each other. There is still some confusion between open data in the broader sense and open data applied within governments. There is still a little bit of confusion between open data and big data, and still some governments tend to confuse the discussion about data analytics and data mining and open data. We thought that it was extremely important, and still is extremely important, as governments progress in the implementation for open data strategies and initiatives, to work with them to clarify the definitions they refer to.

Briefly, I would like to share with you some of the outcomes of the 2013 data collection we ran that highlights some of the key challenges that governments still deal with. These challenges are of different natures. There are policy challenges when it comes to the strategy, for instance—what kind of strategy and how to make sure that the strategy for open data aligns or is better integrated with social and economic development strategies, open government strategies, public sector reform strategies, and digital agendas for governments, for instance. There are technical challenges—how to, for instance, enable interoperability and integration that didn't exist, how is it possible to foster the linkage of data sets to be released in open formats, and all the related technical issues that governments are still dealing with in many instances.

But there are also organizational challenges that, according to our survey, still remain some of the most important challenges that exist. For instance, administrations, unfortunately, are still very much silo-based in the way of functioning, meaning there is a strong sense of ownership that different public institutions associate with the fact that they are the ones responsible for producing, collecting, and distributing certain data sets. These represent a big challenge in some countries when they started thinking about the development of open data initiatives because they encounter a certain level of resistance within the public agencies.

Last but not least, there are challenges that are of a legal nature. The other witnesses, for instance, mentioned the relevance of privacy and security and how we deal with these issues. It is not only for these aspects that it is important to look at the legal constraints that exist in some legislations. For instance, I will provide two additional examples. First are access to information laws, or freedom of information acts, which were adopted by many OECD countries from decades ago. They are now going through revisions, for instance, to make sure that they also accommodate the need for open data, not just for access to information. There are also restrictions, legally speaking, that concern the sharing of data within the public sector. So at times, for instance, linked data sets can support their data analytics, which can help identify trends to improve policy-making and service delivery, but still some legal restrictions do not enable different parts of the administration to access the various data sets.

● (0910)

Now when it comes to value, we saw that there are three main sets of value that governments are trying to achieve. As an organization we do not advocate for any approach or for any value sets, but I think it's important to underline that there is economic value that can be achieved through open data in the wider economy.

The other witnesses mentioned for instance the ease with which business start-ups are created. I would like to add also the emergence of new private sector type businesses, for instance the so-called infomediaries that enable the relevance of the data being open to a wider group of citizens that, in many instances, would not know how to get the most value of the raw data sets being made available.

There is economic efficiency that can be gained within the public sector, improved service delivery, improved performance, and improved efficiency in the internal dynamics. There is also the social value, for instance in terms of empowering citizens to make more informed decisions on their own lives. It tends to do with a different type of engagement, for instance, and participation in policy-making and service delivery.

Last, but not least, there is a third sector value that has to do with what we call good governance value or political value. In other words, the fight for higher transparency, higher accountability, and higher responsibility of governments.

We at the OECD are now looking at the next step of what we would like accomplished in collaboration internationally with other organizations, with institutions like the ODI, and within contexts that are internationally collaborative like the OGP, the G-8, and the G-20. The big focus we have right now is on supporting the further strengthening of the strategic approach and implementation, but also focusing a lot on value creation impact assessment. Because we do believe that as investments keep being made by governments—and let's not forget that open data is not for free—there is a financial cost for governments.

It's important to keep an eye on the value being created and on the measure of this value. We are part of the working group on open data, part of the OGP, so we collaborate with other, not only international organizations, but governments and institutions to make sure that this effort moves ahead internationally, so not only working with individual governments.

So now I come to the questions that you asked. How does Canada stand in relation to other jurisdictions? Certainly we saw Canada being grouped among the countries of the OECD that we defined as quick followers, meaning there have been a group of countries that have been the pioneers, the U.K., the U.S. They have been excellent in being ambitious in this context right from the beginning.

Then we have other countries that have taken other approaches. We also have countries that have been, like I said, the quick followers. I can mention for instance France, Mexico, and Canada, which have caught up quite quickly, even if at different levels than the other countries, in following up what have been the good examples set by, for instance, the U.K. and the U.S.

In that sense, I think, an extremely positive value-add of Canada has been the one of linking open data with open government, the one of linking digital government strategy with the open data strategy, the effect of having adopted an approach that nurtures collaboration internally, the fact that a committee was created to gather various representatives from the various jurisdictions.

I think a big focus has been on improving the portal, the first version of the portal, to in June 2013 the release of a new version that increases not only the accessibility of the data sets but also the use of social media features that focus very much on increasing the engagement of the citizens.

Because when we come to value creation—I think this is one of your questions also—how do we make open data valuable for the Canadian community? I think that a key point where we see the need for strengthening the efforts of OECD member countries and maybe Canada could be strengthening the focus on knowing the demands of the data.

If you consider the three sets of value mentioned, there are different data users in the community of users, which may have different needs. So knowing the demand is important. Nurturing the demand is important. Nurturing the engagement in the use of the data is essential to produce the value.

● (0915)

In that sense, I think it's important down the line. For instance, in the data collection we conducted last year, Canada ranked as one of the governments that had the highest number of data sets available. But as one of the witnesses mentioned as well, I think it's very important now to move ahead in the level of openness and the visibility of these data sets, which have an important impact on the value creation.

Last but not least, I would like to refer to the point on privacy that you were asking about. In addition to what the other witnesses mentioned, I think in order to protect privacy it is extremely important to have clear guidelines for the public servants. Remember that public servants are key actors in the ecosystem, and therefore, keeping the focus on training civil servants and raising their awareness of breaches of privacy that may emerge from a number of actions they can do in relation to open data is essential.

It is essential more and more as social media efforts are combined with open data efforts and mobile government-supported efforts such as, increasingly, the use of mobile technologies within government, because all of a sudden we start merging the value domains that are relevant to produce the value for open data. But I think it's very important to remember that civil servants need to be aware of the risks for security and privacy that emerge from the linkage of these three different domains.

Last but not least, yes, I agree with the previous witnesses, in the sense that I think governments are ahead of businesses in these aspects, in a sense. But I wouldn't be unfair and compare government with the private sector in terms of how much they are opening up, because I think there are important concerns in terms of privacy and security that relate to data sets owned by governments, which are very different from data sets owned by some entities in the private sector. I think comparing the two is important, but I think it's even more important to keep high the comparisons across governments in the world to make sure that the best practices are shared and replicated.

Thank you.

The Chair: Thank you for your presentation.

Now we'll go back to the U.K., with Ms. Bates, from the University of Sheffield.

You have 10 minutes for your presentation. Thank you for being here.

● (0920)

Dr. Joanne Bates (Lecturer in Information Politics and Policy, Information School, University of Sheffield): Thank you very much for inviting me and hello from Sheffield.

I'm a lecturer in information politics and policy. I've been researching the politics of open government in the U.K. for the last few years now. What I've decided to concentrate on in my opening presentation today are the two themes that I saw emerging in the questions that were presented to the panel by the committee.

First of all, I'm going to talk a little bit about how Canada compares to other jurisdictions; and secondly, I'm going to talk about this issue of generating value from open government data.

The first question, then, is how does Canada compare to other jurisdictions? There's a number of different methods that we could use to compare different countries' open data initiatives. A very simple approach would be the one taken by the open data index, which is an Open Knowledge Foundation supported project. This basically just compares a number of different data sets that have been opened in different categories by different countries. In this kind of method, Canada comes out 10th overall out of 70 countries, so it's doing pretty well there.

A more complex approach is the one that Richard mentioned, the open data barometer project, which was supported by the Open Data Institute and the Web Foundation, and published last year. This more complex methodology looks at open government data readiness implementation and impact across different countries. In this methodology, Canada scored eighth out of 77 countries, so it's doing a little bit better in this sense.

Now, the researchers behind the open data barometer project used a number of different methods to collect the data. One was an expert survey that they did across all the different countries, and they used quite a robust methodology here to gather and to analyze this data. I think this is the best sort of comparative data that we have at the moment. What this data suggests is that Canada's is a very well-resourced open data initiative, but in terms of government support, in terms of incentivizing reuse, for example by competitions and grants and things like that, Canada is perhaps a little bit lower compared to some other countries. Also in terms of the training that's available for potential reusers in Canada...[*Technical difficulty—Editor*]...from the experts that Canada is a little bit lower there as well.

That's how Canada fares in terms of the implementation coming out of the open data barometer. In terms of impact, as Richard also said, Canada seems to be doing pretty well comparatively in terms of the political impact, and even the economic impact of open data. Although scoring only 3 out of 10 through this survey, that does actually compare quite well. It brings in Canada to joint eighth overall. But in terms of social impact, and this includes things such as environmental sustainability and the inclusion of marginalized populations in policy-making through using open government data, Canada is scoring relatively low, scoring 0 out of 10 for environmental impact and 2 out of 10 for social inclusion. Now relatively speaking, that means that Canada is doing quite poorly in terms of environmental impacts, but is about average for impact on socially excluded populations. There's been very little impact from open government data on improving social exclusion issues.

What this study also highlights is that this is quite a similar pattern to what we're seeing in the U.K. In the graph that Richard showed earlier, the U.K.'s pattern is very similar as well. The social impact of open government data in both the U.K. and Canada is a lot lower relative to the observed economic and political impacts. This suggests that perhaps not enough is being done in both Canada and the U.K. to enhance that social impact from open government data.

This pattern is not the same in every country. For example, in the U.S.A., Sweden, and New Zealand, those countries are scoring much better relatively on the social impact in relation to the political and economic impacts, which suggests that there might be interesting best-practice cases and similar things that you could use from those countries if you're interested in increasing the social impact of open government data.

Now what I would also point out is that both of these studies, the open data index and the open data barometer, are very quantitative studies that are interested in ranking countries against each other. My research is interested in the political drivers behind open government data.

● (0925)

I'd say there's a real need for further comparative political research in the drivers behind open government data across different countries. I think we need to really be asking, who is benefiting from specific decisions in different jurisdictions? Who is being empowered and disempowered as a result of where the boundary is being drawn between open and closed data in different countries? Who's being empowered and disempowered as a result of where the investment is being made, where the reuse of open government data is being incentivized? As well, what do the regulatory contexts in different countries allow in terms of what is allowed or prohibited in terms of open government data reuse?

That takes me on to thinking about the potential value to be generated from open government data. I just want to state quite explicitly there's no simple linear trajectory from opening up data to generating positive societal impact. A lot of other things go on within that space as well.

In terms of economic value, lots of claims have been made based on economic modelling. Richard referred to the McKinsey report. There has been other research done as well, such as Rufus Pollock's work in the U.K., but there are still a lot of uncertainties in terms of the conclusions this research comes to.

In terms of the headline figures that research like this promotes, such as *x* trillion pounds can be added to the global economy, £6 billion can be added to the U.K. economy, I think we need to remember that all economic growth is not necessarily good growth. Open government data can lead to the production of all sorts of exciting, innovative, socially beneficial products and services. Equally, open government data can be used to develop products and services that could have negative social implications even though they generate substantial profits and might contribute a lot to GDP.

One example I'm thinking of here is the weather derivatives market, which is heavily dependent upon open weather data but has a very questionable relationship with climate change mitigation.

So that's the economic value.

In terms of generating social value, which is an area that the open data barometer project suggests Canada is relatively weak in, I think what we need to see really is an investment in the development of an infrastructure that brings together organized civil society, local communities, researchers, and other domain experts, with open data, to both source data sets from public bodies to advise on their collection of data that is useful for them to be using, and to develop methods of data analysis and create tools and resources that can engage and critically inform common concerns.

We're starting to see a little bit of this in some of the work that the Open Data Institute does, but I think that could go further and be more widespread as well.

In conclusion, I just want to reiterate really that we need to avoid the assumption that there is this simple linear trajectory from opening data to generating positive societal impact. When making policy decisions, I think it's important to think about what specifically you're aiming to achieve with open data, and then think about the wider policy ecology that needs to be thought about in order to make that happen.

Thank you.

[*Translation*]

The Chair: Thank you all for opening remarks.

We will now move on to questions from the committee members.

Mr. Ravignat, you have five minutes.

I would just like to remind you that it's better if you state the name of the person you are speaking to, especially for those individuals testifying via videoconference.

Mr. Ravignat, you have the floor.

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

[*English*]

Thank you to all the witnesses for being here from so far as well.

Perhaps I'd like to address my first question to Mr. Stirling. It's a question that surrounds what can be done about setting standards across ministries, across departments, to ensure the data sets that are compiled, in this case on a portal...and basically just making sure that the coordination within a very large apparatus like a federal government can make sure what winds up getting posted or available, in this case to Canadians, is actually useful.

● (0930)

Mr. Richard Stirling: Okay, thank you.

There are a couple of lessons or things that I want to talk about here. One is examples from my own experience of how not to do it, which is to be absolutely dictatorial to the departments; or indeed to accept the standard once and expect everybody to be able to follow it, even with what you might at the time assume to be simple guidelines.

What we've been trying to do at the Open Data Institute is to create tools that enable people to know whether or not they're meeting the agreed-upon standards. One of the data formats that I think a number of your witnesses in this and other sessions will talk about is something called CSV. It's publishing simple tables with certain columns in a standard format every month.

In the U.K., every local council—so 454 different authorities—publishes the same data set to the same standards every month, in theory. In practice, you have around 400 elegant variations on the theme. That's not because people don't know what they're meant to be doing. It's because they're following a process and they don't understand it, and they can't see what good looks like. So they get the file at the end that looks like a CSV, and they're happy.

We've built a simple validation service, which enables you to check against a schema. That poor desk officer in his local authority, who thinks he's doing a good thing, can check. He can upload the file and say, okay, that hits the standard.

That's how we're supporting this sort of federated approach towards data setting standards. The mechanics of that, you can do elsewhere.

Mr. Mathieu Ravignat: Great. Thank you very much for that. It was very interesting.

My next question is for Madam Bates.

You were talking about social inclusion. We got a sense from this government that it began with a very broad understanding of what open government was, touching on basically three components: open dialogue, open data, and open information. It seems to me that it's begun to be narrowcasted on just talking about open data and kind of dropping the open dialogue piece.

I wonder if you'd be willing to comment on the relationship between social inclusion and open dialogue, and open data and the generation of that data. In Canada, we have first nations and aboriginal communities across our country who we need to take into consideration, particularly with generating data issues on cultural ownership and so forth. The open dialogue piece, to me, seems to be pretty crucial, and I wonder if you might have some helpful comments for us.

Dr. Joanne Bates: Thank you for the question.

I think this is a really important issue. There was a piece actually written by a Canadian researcher called Michael Gurstein. I'm not sure if you're aware of him. He's a community informatics expert who has researched in this area. He was particularly concerned about how this kind of shift to open data, perhaps in the prioritization of data over other aspects of the democratization process, could lead to an empowering of the empowered and a disempowering of those people who are already socially excluded.

Most people have very limited skills and ability when it comes to numeracy, never mind data analysis and using complex government public sector-produced data sets. When we're thinking about how all of this connects together, if open data is on the agenda, there will always be some sort of what's been termed infomediary. That's somebody standing there in between certain population groups and the data, to help them make sense of that data and to use it in ways that are beneficial and useful for them. They will incorporate dialogue and understanding and a more complex social understanding, rather than the more technical approach, to thinking about open data and democracy.

I don't have any simple answers in regard to social inclusion. It's a very complex thing. I think—

● (0935)

[*Translation*]

The Chair: I am going to have to cut you off; time's up. Ms. Bates, perhaps you can continue your answer on another question.

I will now give the floor to Mr. Trottier for five minutes. Remember: that includes the time for questions and answers.

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

My first question is for Ms. Da Sylva.

You referred to data sets. A lot of that data is provided by the government, but it's not necessarily in a usable format. It is not in RDF format. What are the technical barriers that prevent us from producing more data sets in RDF format? Does that mean Rich Data Format?

Ms. Lyne Da Sylva: No, but...

Mr. Bernard Trottier: Perhaps you could explain what that acronym stands for.

Are other levels of government in Canada in a better position to provide that data in a usable format? Are there other examples elsewhere in the world where data is provided that way?

Ms. Lyne Da Sylva: The acronym RDF stands for Resource Description Framework. It's an extremely simple format, which allows the computer to manipulate it. However, it is highly structured. So it is painstaking for a person to write and read it.

In Canada, there are several data sets in CSV format, which Mr. Stirling referred to. These are Excel spreadsheets recorded line by line, with commas between columns. Producing this type of data in CSV format is quite easy. There are no technological barriers. You just have to develop the corporate culture.

Some types of data do not lend themselves to this, such as significant quantities of geographic maps or information on geographic maps. The interest in this data lies in the graphic/visual aspect. You are obviously not going to capture that in an Excel spreadsheet. In that sense, there's a limit to the information that can be disseminated.

However, CSV formats can be readily manipulated by computer and can be converted into RDF format. Some governments have outright decided to put everything in RDF format. In the United States, as in the United Kingdom, there's a desire to go the way of RDF.

As I explained, the barriers stem from the fact that the nature of the data does not lend itself to this format in some cases. The alternative would be to set up automatic conversion systems for certain types of data.

Mr. Bernard Trottier: Thank you.

[*English*]

My next question is for Ms. Bates. I think you made a very valuable comment about the need for the organizational infrastructure to make sure there's continuous improvement when it comes to open data, for all the good reasons you mentioned. This need to bring researchers, academics, government, and citizen stakeholders together where they can talk about how they can improve open data. I'm not even sure to what extent that's part of our government's plan right now; hopefully it will be part of the plan.

Can you give some examples of best practices when it comes to making sure there's that ongoing dialogue, as opposed to different government departments doing things in isolation or different researchers doing things in isolation?

● (0940)

Dr. Joanne Bates: Yes, I think it is very important. I was just reading a piece of research that is yet to be published here in the U.K., looking at barriers to open government data fulfilling the promise that was made at the beginning. One of the issues, it turns out, is around the implementation side of things, the numerous barriers that have been mentioned by some of the other speakers around implementation.

Another category of barriers is around the reuse space and increasing that demand for open government data. At the moment there's a demand from certain sectors that have been interested in seeing the potential value in open government data, but there are a lot of people out there working in local communities, in organized civil society, researchers who could potentially get a lot of value from open government data but have never heard of it, don't really understand what it is, don't understand the “open” about open government data—that it's reusable, rather than just something they can access—and things like that.

So increasing that knowledge within the broader community, I think, is really important.

Thank you.

[*Translation*]

The Chair: Thank you.

I will now give the floor to Ms. Day for five minutes.

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

Thank you to all the witnesses for being here.

My questions are going to be in French, and there is simultaneous interpretation available.

Mr. Stirling, I will start with you.

To your knowledge, do you, in the United Kingdom, have the equivalent to our Access to Information Act? In the United Kingdom, are there any problems with access to information that get in the way of implementing an open data policy?

In Canada, 35% of the 50,000 annual requests were not answered within the established timeframe. At a previous meeting, we also discussed the problem of the $5 cost.

The NDP is of the view that you cannot be all about transparency on the one hand, and secrecy on the other. Apart from an open data policy, should the Access to Information Act also be improved?

[*English*]

Mr. Richard Stirling: I think that you need to look at the legislation alongside the culture of the organization. In the U.K. we've done both. So there has been a big culture shift inside governments towards the presumption that data can be published and that where it can be published it should be, partly to help your colleagues in government when they're looking for data but also as a generally good thing to support innovation and greater accountability in the U.K.

The legislation that we've been working on came in under the Protection of Freedoms Act, which amended freedom of information legislation in the U.K. to make changes to how the data could be used and published, enabling reuse and also enabling people to ask for data in technical formats.

[*Translation*]

Mrs. Anne-Marie Day: Thank you.

Ms. Bates, my next question is for you.

You talked about societal risks. You mentioned that open government data policies are used with a view to reaping the full benefits of commercializing public services, growing capitalism and exploiting societal risks.

Can you explain what you mean by “societal risks” in the context of that research?

[*English*]

Dr. Joanne Bates: Thank you.

I think there are two questions there, really, that are related to my research. One of the arguments that I made in some of my written work around open government data in the U.K. is trying to analyze how open government data policy connects into the U.K. government's open public services policy, which is really an effort to further marketize public service provision in the U.K., opening up provision to the third and private sectors.

What I was doing there was thinking about the ways that open data fits into that agenda. It is quite explicitly laid out within the policy, but in terms of how the data can be used by business intelligence—analysts, for example—to see where there might be profitable public services to bid and run and things like that, and also in terms of this notion of the public service user as a customer of public services and being able to use open data based apps to make decisions on which services they ought to use.

The second part of the question is around societal risk and how open data might fit within that. One of the areas of open data release I've been looking at in some detail is the opening of weather data and how this fits in with efforts within the financial markets to develop weather derivatives products. These have been popular in the U.S.A. for a number of years and then spread elsewhere, and the U.K. financial markets want to be competitive with the U.S. markets.

Open weather data, as weather data is already open in the U.S.A., is very valuable for these financial market trades around weather derivatives. But they do have a very questionable impact upon climate change mitigation because basically when businesses are buying these products, they are essentially removing the financial impact of weather instabilities on their businesses. So it gives them less incentive to demand action on climate change mitigation. So there are various complex relations going on there that I think need to be thought about when we're looking at why different data sets have been released in different jurisdictions.

Thank you.

● (0945)

[*Translation*]

The Chair: Thank you, Ms. Bates, for your answer.

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

[*English*]

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

Thank you to all our witnesses for lending your expertise here today.

I do want to pursue my line of questioning with Mr. Stirling, if I may.

We're really looking at a new frontier here aren't we, Mr. Stirling?

Mr. Richard Stirling: Yes, I think so.

Mr. Mark Adler: With a huge amount of potential and opportunity. You mentioned the McKinsey study, which looked at seven sectors and said there is the potential of adding $3 trillion annually in economic value. Wouldn't that also encourage—remember the Klondike days—a lot of unsavoury folks and draw them into the marketplace too? Does it have that potential?

Mr. Richard Stirling: I don't know about that. The $3 trillion figure comes from adding up all of the benefits across the world from being more efficient. The seven sectors that they looked at were things like education, health care, consumer credit, where the benefits were from the markets operating more efficiently and more effectively, and in the case of consumer credit then in a more fair fashion.

Mr. Mark Adler: We all know and realize that there's a tremendous amount of good that can come of this in terms of the delivery. I'm just looking at it from the public sector side initially. It's empowering citizens. It's changing, really, how government can work. It's propelling innovation, and what's most important is it can improve the delivery of public services.

Are we looking at a day when we can deliver public services that are tailored to the individual as opposed to broad, umbrella-like public services? Are we heading towards that kind of a regime in your view?

Mr. Richard Stirling: In my view we are heading towards that type of regime. I think it's quite a long way away, and it will require more than just open data. That's the whole digitization of public services. It requires you to solve the problem of identity, which is a thorny one, for getting identity in the public services context, and using data that shouldn't be made public in a consensual way to then personalize and tailor those services for the individual who's in front of you. You can see this happening a little way in a completely analogue fashion where a job seeker goes in and speaks to an adviser, and the adviser looks at them and gives them tailored advice. What the technology would enable you to do in a few years' time is start automating some of that advice and sending people towards specialists.

● (0950)

Mr. Mark Adler: You indicated earlier that the public sector is ahead of the private sector in the collection of data?

Mr. Richard Stirling: In the publication of data.... The private sector is, depending on the sector, slightly ahead of the public sector in collection. But the business case for the public sector in opening up its data is clear, and in the private sector—

Mr. Mark Adler: The private sector realized the advantage of this way in advance of the public sector. Is that correct? Because they realized that the more data they can get on their customers the better they can serve them. Is that correct?

Mr. Richard Stirling: The private sector is ahead of the public sector in customization and profiling. I'd just draw the distinction between open data that tends to be information around a cohort or around a place, and the personal private data that shouldn't be made open in general.

Mr. Mark Adler: But there will be some public data that should remain proprietary, right? All public data should not be open data, correct? In your estimation....

Mr. Richard Stirling: There are good reasons why some public data wouldn't be made open, but I think the presumption should be that public sector data should be made open where possible.

Mr. Mark Adler: Right.

How much time do I have left? Am I done?

[*Translation*]

The Chair: Your time is up. Thank you.

We will now go to Mr. Byrne, who has the floor for five minutes.

[*English*]

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

Thank you to our witnesses for taking the time to provide us with your expertise and knowledge. We particularly value the fact that we can get international participation in our own thoughts and deliberations.

One question I would have is surrounding the notion of behavioural modification within government institutions related to open data. It's a relatively known phenomenon that when circumstances change, legislation is passed, it requires more open.... Whatever may be the driving force, institutions will indeed to change and adapt to the evolving circumstance. With open data, is there a potential...?

There's an obvious potential for governments to modify their behaviour to allow the increased outflow of information, but there's also the potential that governments may retract. Knowing that there is a driving force, an impetus to collect data and to disseminate it, there may be a driving force to stop collecting it so that they do not have to disseminate it.

In any of your experiences—I'll start with Mr. Stirling and then lead to Ms. Da Sylva, and Ms. Bates—if you could then comment relatively briefly, do you have any experience that you could relate to that particular instance or example?

Mr. Richard Stirling: I would say that the culture change comes from people seeing benefits themselves in how they're using data. The NHS Information Centre is now much more gung-ho on the open data agenda because someone else has come in and provided them with a service and some analysis that was useful to them in doing their day-to-day job. That's the thing that really switches the culture change.

In terms of the behavioural response that you just highlighted, of some people not wanting to collect data because they're worried that they then have to publish it, I haven't seen that. That's partly because, as in the earlier questions we were talking about, the public sector is slightly behind in the collection of data so it tends to only collect data it needs for operational reasons. So you can't stop the collection of the data without also stopping the operations.

● (0955)

Hon. Gerry Byrne: Ms. Da Sylva.

Ms. Lyne Da Sylva: I don't know whether I should answer in French or English.

Hon. Gerry Byrne: As you choose....

Ms. Lyne Da Sylva: I think because the discussion has been going on in English I will pursue this way.

If I understood correctly you want my opinion on whether organizations may stop collecting data so they won't have to disseminate it. They can't disseminate what they don't have.

Hon. Gerry Byrne: There have been some that would think—maybe in terms of the Canadian example, the elimination of the mandatory long-form census. There has been some criticism by civil society that the elimination of the long-form census prevents data from being used to apply public pressure to establish public policy priorities. That's a Canadian example, which may or may not be true, but I throw that out to help focus your thoughts or answers. If you like, maybe we could move to Ms. Bates and then Chiara Ubaldi to....

Ms. Lyne Da Sylva: I will try a tiny answer, however.

There's actually a link between increasing open data and the data-intensive research, referred to as big data. There'll be a point when there will be so much data, it won't be that we won't know what to do with it but we won't know how to handle it. There will have to be decisions made as to what is kept and how. I think that actually is linked to your question.

We may not necessarily stop gathering data or the data collection may end up being a little more focused if we get archivists into the equation, saying, "Well, think about what you want to collect and how you want to keep it and the reason why you want to keep it." That may determine the policy.

Hon. Gerry Byrne: Thank you.

[*Translation*]

The Chair: I have to cut you off because your time is up. However, Mr. Byrne, you may get another turn a bit later.

We will now go to Mr. Aspin, who has the floor for five minutes.

[*English*]

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

Welcome to our guests from near and far. Thank you for helping us with this study.

I have a couple of questions, first of all to Mr. Stirling.

From the information you've given us, Mr. Stirling, Canada, you say, ranks tenth or eighth, depending on the criteria and the method. I think you've given a little bit of a hint in this regard, but could you give us maybe two or three primary factors that could enable us to be maybe amongst the leaders like your own country?

Mr. Richard Stirling: Yes. The rankings that I showed put Canada eighth, and the biggest areas of difference between where you are and where the countries at the top are currently are around social impacts and also some of the core data sets, which haven't been released but I think were specified in the G-8 communiqué around land use, budget use, company identifiers, and legislation, although there are also some others around health, education, and crime. If you think about wanting to get the benefits flowing through into health, education, and crime, you need those sorts of core data sets to be available. Otherwise, the innovation can't flow.

Where the other countries are leading to is supporting the innovation and their economic and social impacts, taking things from an idea and supporting them through to the point where they're sustainable organizations.

Mr. Jay Aspin: Thank you.

The second question is to Ms. Bates.

Ms. Bates, I'm very interested in the commercial aspects of this. You know that I like my government focused on jobs and the economy. I'd like to ask the question: which databases have been proven to provide the best return on investments? Second, how can the new Canadian Open Data Institute help with start-ups in Canada?

● (1000)

Dr. Joanne Bates: Thank you.

I can't give you a specific answer on which data sets lead to the best return on investment. I'm afraid I don't have access to that information.

Perhaps, Richard, that might be something that you have.

The kinds of data that I've been looking at, looking at specific data sets where there is a commercial interest, have been around things such as prescriptions data and the pharmaceutical companies' interest in that data for marketing practices, the weather observations data that the financial markets are interested in, and there's some arms trade data as well that's not been opened but that is very valuable for the arms trade as well.

In terms of how to incentivize and help start-ups in this area, I think the Open Data Institute model is a great model. The shame is that it's in the middle of London. It would have been great to have an institute like that in the regions as well—like Manchester or somewhere else, maybe up in Scotland—rather than just incentivizing start-ups and growth in London. That would be my recommendation, that things like that need to be in the regions as well as in the capital.

Mr. Jay Aspin: Thank you.

Thank you, Chair.

[*Translation*]

The Chair: You have one minute left, Mr. Aspin.

[*English*]

Mr. Jay Aspin: I'm fine.

[*Translation*]

The Chair: Thank you.

I will now turn the floor over to Mr. Ravignat.

Mr. Mathieu Ravignat: Thank you, Mr. Chair.

My question is for Ms. Ubaldi and Mr. Stirling.

When the decision was made to centralize access to data and to take the data away from various departments, we saw some concern being expressed, by the scientific community in Canada, in particular. Their concern was that this might be a move to make it easier to control which data would be open and which would not.

Could you give us your comments on the need for a healthy relationship between those in power, particularly those in cabinet, and those responsible for ensuring open data. If you have any examples of best practices in your political system, I think it would be very useful for us to hear about them.

Let's begin with Ms. Ubaldi.

[*English*]

Ms. Barbara-Chiara Ubaldi: If I understand the question correctly, it applies only to old data sets and not to data sets applicable today in the scientific domain. I think the answer is transparency, meaning that certainly there is a need for transparency in the actions taken, in the case of Canada, by the cabinet office in relation to which data sets to open and in which format to open them. So, there are two points.

First of all, open data requires an ecosystem of actors who work together. You have the organization that sets the policy but then you have the other parts of the administration that also produce the data sets and share the data sets that need to be brought on board.

Second, there is the need to be transparent. In order to overcome some of the resistance—and you mentioned some of the negative impressions—there is a need to be transparent about what is going to be done in terms of which data sets will be opened, in which format, at which point, and for use by whom. So there might be scaled-up approaches. Not all governments have taken the approach of getting data out there. For a couple of governments—in the Netherlands and in Denmark, for instance—the approach has been less adventurous, but there's been clarity and transparency about it.

● (1005)

[*Translation*]

Mr. Mathieu Ravignat: Mr. Stirling, do you have anything to say about this?

[*English*]

Mr. Richard Stirling: I think this is one of the reasons that any such activity needs to be underpinned by strong principles around a presumption in favour of publication of the open data, because that helps build trust in the idea that nothing adverse is happening through the centralization.

The other thing I would just observe is that open data has actually been the underpinning of a shared research base for hundreds of years. This open research agenda is very much in the spirit that research has been going on for centuries, and I know that a number of countries are looking at this, including my own. But I think it's still an open discussion with no concrete conclusions yet as to how to make sure that publicly funded research is open by default.

[*Translation*]

The Chair: Mr. Ravignat, you have 30 seconds left.

Mr. Mathieu Ravignat: I won't have enough time to ask my question.

The Chair: Then we will go to Ms. Ablonczy, who has the floor for five minutes.

[*English*]

Hon. Diane Ablonczy (Calgary—Nose Hill, CPC): Thank you, all, for appearing.

I was struck by Ms. Bates' conclusion that research is beginning to show that just opening data is unlikely to generate the impact that a bright-eyed and bushy-tailed government is hoping for.

I guess, Richard, you've had the most experience in this area. What are the benefits? Open data just for the sake of open data obviously is a lot of work for very little payback. So, what's the payback? What's the goal? What's the objective?

Mr. Richard Stirling: Thank you.

I view open data as an enabler of things in the economy. The benefits of open data come from people being able to find out how to do things faster, cheaper, or better. The examples that I used in my introduction are, I think, examples that address the “so what?”. The £200 million cash savings found in one drug line in the NHS drugs budget is the basis for a sustainable business.

Proving the approach of taking a data-rich but regulation-light to a market, that's possibly the future of regulation. It gives the regulator better information, but at the same time doesn't increase the burden on business. Again, that has a sustainable model to it. I think at the moment we're just seeing the transition from thinking about getting the data out there through to thinking about the products, the services, the way they get created, and in some cases, the way you support these businesses coming through.

Just because I'm aware there's a regional thing here, we do now have bases in Leeds, Sheffield, and Manchester, as well as Brighton, so we're also thinking about how we take this, not just to the capital city, but around the U.K.

Hon. Diane Ablonczy: To some degree, we're just tiptoeing through these tulips and we're not quite sure where it's heading, but Ms. Ubaldi, from the OECD's point of view, you put a lot of time and effort into persuading various member countries to move down this path. Why is that? Is it the flavour du jour that everyone's getting excited about? What are the hard benefits that the OECD has identified?

Ms. Barbara-Chiara Ubaldi: Thank you for your question.

We try to make governments do it in a way that is right, not just because it's the flavour of the moment. So I think your question is extremely important. We think there is value and a positive impact that still needs to be fully demonstrated; that is true. In some areas it cannot possibly be quantified, like in the social area that some of your colleagues mentioned before. Instead, it's an extremely important set of values that can be targeted.

For instance, in terms of social value, there is certainly an increasing number of examples showing how open data has increased the participation and the engagement of parts of society that otherwise would not be brought into the discussion and dialogue with governments in terms of service delivery and policy-making. However, that requires that the government focuses not only on the usual actors who are interlocutors in this area, for instance, the private sector, but there are other actors in the ecosystem like journalists, civil society organizations, citizens associations, librarians, and so on, who are non-typical groups of actors who need to be reached out to.

From the perspective of the OECD, the reason we are focusing so much on this is not because many governments have pushed it up on the agenda, but because this has an impact of changing the way the government conceives a number of actions, ranging from policy-making to service delivery. The challenge is big, so I cannot tell you that there are demonstrated values. There are important estimates that my colleagues mentioned. There's no clear data yet that demonstrates the value, but there are a number of examples from all levels of jurisdictions that demonstrate there are changes in the way the government interacts with society in creating economic and social value.

Last but not least, in terms of transparency and increased trust, there is a tendency showing that the higher transparency and openness of governments in releasing key data with information on the operation—

● (1010)

Hon. Diane Ablonczy: I just have one more question, and I want to be able to get that in.

[*Translation*]

The Chair: I'm sorry, but your time is already up.

[*English*]

Hon. Diane Ablonczy: I just got started.

[*Translation*]

The Chair: There will be another round of questions.

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

Mrs. Anne-Marie Day: Thank you, Mr. Chair. You have to wait your turn.

Mr. Stirling, my question is for you.

The Open Data Institute—ODI—is an independent, non-profit, non-partisan company. According to its website, the ODI has secured 10 million pounds sterling from the UK Government and $750,000US from Omidyar Network. The ODI is working towards long-term sustainability.

How much does this kind of free market for data cost? What costs might that entail for all of the G8 countries? And what structural safeguards have been put in place to ensure the non-partisan and transparent flow of data?

[*English*]

Mr. Richard Stirling: Thank you.

Yes, we are independent, non-partisan, and not for profit.

In terms of the firewall and keeping that status, despite being in receipt of public money in the U.K., our corporate structure is that we're a company limited by guarantee. We have no government representative on our board, although we do have conversations with the financial monitoring grant administration bits of government around our core metrics.

If you would like to look at the core metrics, then we have them published on the dashboard on our website. It's exactly the same metrics that I am held to account to on a weekly basis and that my board gets to see every six weeks.

On the cost of implementing a similar organization across the G-8, I don't know, which is my honest answer.

I know that in our organization we've been able to do nicely in the U.K., but we've also been very blessed by having Sir Tim and Sir Nigel as our founders and being able to attract a very good team as a result.

We also have a global network now that is operating in a number of G-8 countries, and it's possible that they will be able to address some of the same needs. It does depend on how much ambition there is, like how many start-ups you want to help, how fast you want to accelerate the economic benefits and those use cases.

● (1015)

[*Translation*]

Mrs. Anne-Marie Day: My next question is for all of you.

How important is it for a government to cooperate with other levels of government and other countries on open data practices and data set selection?

And here is my follow-up question. Should Canada and the United Kingdom be working together?

The Chair: Could you specify who the question is for?

Mrs. Anne-Marie Day: The first part, on the importance of cooperation, is for all of the witnesses.

The second part of my question is for the witnesses from the United Kingdom.

The Chair: Let's begin with Mr. Stirling. I will then give the floor to Ms. Bates.

[*English*]

Mr. Richard Stirling: There are a number of bodies being set up to foster that international collaboration. There are technical committees working at the Open Government Partnership, which I actually think has a Canadian co-chair at the moment. There are similar bodies working in places like the OECD, and the UN, etc.

Between the U.K. and Canada specifically, I understand that there are regular working-level meetings.

Dr. Joanne Bates: In terms of the collaboration between Canada and the U.K., as Richard said, there is a working-level collaboration.

I'd also say that, in terms of the research community around open government data, there is quite a lot of interaction there between researchers in Canada and the U.K.

In terms of government cooperation at different levels, that is massively important. One of the reasons Canada scored slightly lower on the open data index was because of the federal structure in Canada and local governments and institutions having access to some data but not the federal government, so that interaction is very important.

Thank you.

[*Translation*]

The Chair: Your time is up.

Mr. O'Connor, you have five minutes.

[*English*]

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

We're talking about government provided data. The way our country operates, it seems that government decides what data they're going to provide. There is also no compulsion on the government to provide data.

I asked a few witnesses recently if it would not be a good idea for business sectors to identify the sorts of information they want, then inform the government of the kind of information they want, and then the businesses might get what they need.

I'm going to ask each of you your opinion on that idea because the two witnesses we had recently said no, and they figured that the government should just be left to evolve.

I'll start with Ms. Ubaldi, then Ms. Bates, Mr. Stirling, and Ms. Da Sylva.

Ms. Barbara-Chiara Ubaldi: Thank you.

Our view is in fact that, yes, actors including businesses should voice their interests about which data sets the government should release in open format to enable the governments to prioritize the data sets to be opened up.

Hon. Gordon O'Connor: Ms. Bates.

Dr. Joanne Bates: Thank you.

In the EU, at least, the public sector information lobby is very strong and is very vocal in demanding what it wants from governments and the European Union around the opening of the right of access to public sector information.

I'd also echo what the other respondent said, that it's not just about what business demands, but about what society at large needs to be demanding as well.

Hon. Gordon O'Connor: Mr. Stirling.

Mr. Richard Stirling: I would endorse taking a sector-by-sector approach. In the U.K., we have established sector panels to be the focal point for this conversation between the departments who have the data and their agencies, and the industry working to provide services in that area. To provide the voice of business and the users of data to government as a whole, we have also established an open data user group. There are more details as to how that operates online.

● (1020)

Hon. Gordon O'Connor: Ms. Da Sylva.

[*Translation*]

Ms. Lyne Da Sylva: I would like to add something to what the others have said. The majority of sites I have accessed have a button on the lower right-hand side that says:

[*English*]

"Which data set would you like? Please ask us for your data set."

[*Translation*]

Apparently, there is already some willingness to let users choose their data set. That is obviously on an individual and voluntary basis. That should be accompanied by more collective efforts.

[*English*]

Hon. Gordon O'Connor: I have one other question.

Our government doesn't really advertise that it has these sites; you have to be in the business to know that they're there. I asked whether we should advertise, and those people responded no. I wonder what your opinions are.

Ms. Ubaldi.

Ms. Barbara-Chiara Ubaldi: Well, we believe that for all actors, including the private sector, it is very important to know what's going on before advertising, and awareness-raising is essential.

I would like to echo what another witness said. It's about businesses as actors, but it's also about other groups in society, so taking active steps to advertise and let people know and engage is, I think, essential.

Hon. Gordon O'Connor: Ms. Bates.

Dr. Joanne Bates: I'd agree. I can't see a reason that you wouldn't let people know about data that is available. If you're wanting to prevent some sort of reuse or something, then the policy and regulation would be at a different level.

Thank you.

Hon. Gordon O'Connor: Mr. Stirling.

Mr. Richard Stirling: I'm going to say that it depends on who you think the users are. It depends upon the sector whether advertising is the best way of reaching them. It could be that the most effective way to increase the awareness of the data is to send it to whatever the representative body of the charities and civil society is, in a letter from a minister or a senior official, asking them to spread the word to their members, and the same would be the case with the energy representative body—rather than paying for advertising.

But in principle, I favour drawing people's attention to it, absolutely.

Hon. Gordon O'Connor: Ms. Da Sylva.

Ms. Lyne Da Sylva: Yes, you should advertise. I can't understand why you wouldn't advertise, unless you're afraid that access to the website might be too tremendous, but I don't think we're there yet.

I think advertise, yes, but also make sure that you give sufficient explanations as to what the people will find there, because when you sift through the sites, there are so many different things and they may be difficult to sift through. The Canadian site is just so huge—there are so many things—that to figure out what might be of use to you might take a while. So make sure the advertising educates the user.

Hon. Gordon O'Connor: Thank you.

The Chair: We'll go to Mr. Byrne now for five minutes.

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

I'll continue with the line of questioning I ran out of time for with Ms. Bates and Ms. Chiara Ubaldi.

Ms. Chiara Ubaldi, I asked, of political context within open data, whether or not there might be an incentive to prevent data collection to ensure that it is not disseminated by governments. Have you had any experience or any thoughts or concerns about that?

After you, I'll follow up with Ms. Bates, please.

Ms. Barbara-Chiara Ubaldi: What we have come across is that some governments have had some hesitation in opening up data sets, for privacy and security reasons, but also for quality-related reasons, meaning that data sets that have been collected for a long time are not always of the quality, in terms of timeliness and of accuracy, that is required, for instance, by the community to reuse the data.

There are concerns in that sense, but I think the more we move ahead in the implementation of open data policies, the more those concerns can be tackled.

I hope this answers your question.

Hon. Gerry Byrne: Thank you.

Ms. Bates

● (1025)

Dr. Joanne Bates: Thank you.

This is coming back to the question of whether transparency can impact upon behaviour of the public sector institutions. This is a key argument that we see around transparency and the Freedom of Information Act—issues that we see in the media of politicians deciding to use private email accounts rather than their work email accounts to avoid the Freedom of information Act.

For open government data, I have not seen that behaviour modification in the way that we've seen it for the Freedom of Information Act. Then again, I haven't looked for it, and as far as I know there's been no research done to explore this as yet. So I can't give a concrete answer with regard to open government data.

Hon. Gerry Byrne: Thank you very much.

My next question is to Mr. Sterling.

One of the values of reporting on open data advancement would be.... In Canada, each department provides a report on plans and priorities, and we have the main estimates—a very similar system, I understand, to the U.K. system.

In the U.K. or in other jurisdictions, have reports on open data initiatives and progress been included, instead of from across government or at a very high level, from individual departments and agencies that are required to report their initiatives and progress on open data initiatives?

Mr. Richard Stirling: Each department in central government in the U.K. reports on its progress through to the Cabinet Office, which then publishes reports on departmental progress.

The other thing we're doing is to publish as open data the reporting on other measures; for example, office utilization, compliance with procurement frameworks, etc. This means using open data to underpin the behaviour change and the compliance with some of the more stringent cost-saving measures that have been put in place around austerity. You use open data to drive people and drive the behaviour.

Hon. Gerry Byrne: Ms. Chiara Ubaldi, would you have anything further to that to offer concerning other jurisdictions?

Ms. Barbara-Chiara Ubaldi: I believe that France and the United States have a similar approach, and I think Spain has as well. Increasingly, governments are moving towards that road.

Individual agencies are a big part of the ecosystem within the public sector in opening up data—more and more, the idea is to have a regular reporting on the implementation of the open data initiative—and also across levels of government.

Hon. Gerry Byrne: Thanks, Mr. Chair.

[*Translation*]

The Chair: I will now give the floor to Ms. Brown for five minutes.

[*English*]

Ms. Lois Brown (Newmarket—Aurora, CPC): Thank you, Mr. Chair, and thank you for having me back to the committee.

To all the witnesses, I'm just subbing in for one of my colleagues on this committee for a couple of meetings. So I haven't been privy to all the discussion that has taken place prior. I'm sorry that I won't be able to come back to all of the meetings because I am finding this discussion really interesting.

My particular responsibility in our government is that I'm parliamentary secretary for international cooperation. I've had the opportunity to visit a lot of emerging economies. I know that many of these economies are going to leapfrog over where we have been in our development process simply because they're going to have access to technologies that we've had to develop. I was 200 miles northwest of Juba, in South Sudan, two years ago and everybody has a cell phone. I'm amazed by the technology that is out there and the access that people have.

I know in a study that we did in our foreign affairs committee a year ago, we had Scotiabank in, and they were doing a particular project in the Caribbean with people who are doing telebanking now because in so many areas there just aren't the facilities, the bricks-and-mortar facilities. So Scotiabank is developing a process in the Caribbean for many of those countries to have access to telebanking services.

I look at this and think with the leapfrog ability that's going to happen for many of these emerging economies—and maybe this is a question to you, Ms. Ubaldi, because the OECD, the European countries, are working in many of the emerging economies, along with us as partners—are there developing programs that are looking at emerging economies to help them from the get-go, to start with open data and to be able to communicate this information? I think it would be very instrumental in helping them as governments trying to build communication with their citizens—an open government communication strategy.

Do any of you know of anything that is being promoted?

● (1030)

Ms. Barbara-Chiara Ubaldi: Very quickly, the answer is yes. From the OECD perspective, we have, in particular, three programs. One is for Latin America, which covers all Latin American countries, not only Chile and Mexico, which are member countries.

The purpose is to transfer the practices and experiences from OECD countries to non-OECD countries and emerging economies. The impact that we're seeing is extremely important because some of these emerging economies are extremely active in terms of open data. We are then working in the MENA region, specifically on open data as part of the open government project. We are working in Southeast Asia.

So the answer to you is yes, and then globally—and Richard may join on this—there is the Open Government Partnership. I think one of the main values of the Open Government Partnership is really to help emerging economies and developing countries through the international collaboration to leapfrog and utilize open data...and even open government and the value creation we were talking about earlier.

Ms. Lois Brown: Thank you.

Mr. Stirling, you looked like you had some comments.

Mr. Richard Stirling: I endorse everything that was just said. Our work in this area is focused around Africa, in particular. We are looking at supporting them in building their open data strategies and supporting their local economies because as you identified, they don't have the legacy. They have a huge amount of talent and drive. In my opening remarks I identified that this is a global market and a global opportunity. This could be a great way for them to build local businesses that serve customers around the world and grow their economy to catch up.

Ms. Lois Brown: You're right; they have a great deal of ability. My son-in-law is from Ghana. He is currently the guest professor at the University of Mines and Technology and is supervising the master's classes for electrical engineering, so the talent pool is there. With this kind of access to information from their own governments, I think we would see some of these economies grow like wildfire, I would think.

I have one other question—

The Chair: Thank you. The time is up. I'm sorry.

Now we'll go to Mr. Ravignat.

Mr. Mathieu Ravignat: My question is directed to Madam Bates.

I want to follow up on something my colleague Madam Day was asking about in regard to how Canada scored on some of the indicators. You mentioned that it has to do somewhat with the relationship with other levels of government. I wondered if you could unpack that. Where does the problem exist, precisely?

Dr. Joanne Bates: For the open data index study, which is basically counting the number of data sets that have been opened in a number of different categories, there was a note on that study that because of Canada's federal system, the public transport details, for example, were not held at a federal level. They were held at a more localized level, which caused issues in terms of being able to rank Canada relative to other countries in that field, which led to a zero in that case. There will be other data sets, I imagine, similar to the timetable data, which will lead to a zero score, even though the kind of governance situation is different in other countries.

Thank you.

● (1035)

[*Translation*]

Mr. Mathieu Ravignat: My next question is for Ms. Da Sylva, whom I thank for being here with us.

It worries me a bit that a data expert like yourself should find the data made accessible by the government difficult to use. If an expert like yourself has trouble, I guess every other Canadian has trouble too.

If you could only make three or four suggestions to further democratize the use of Government of Canada data, what would they be?

Ms. Lyne Da Sylva: Actually, I was referring to trouble with automated data use. This comes back to something Ms. Bates mentioned earlier. The expression “open data“ does not just mean public and freely accessible data, it also means data in a format that computer applications can use.

Some data from the Canadian government are perfectly accessible to people. I had no problem viewing a certain number of maps and accessing certain data. It is just that it so happens that some data is in zipped Word files. As a human being, I have no problem dezipping and reading a Word file. However, this unstructured data is much more difficult to analyze directly with a computer, unless you use natural language processing technologies to extract unstructured text and structure it so it can be used.

One solution would be to take suitable data sets and put them into formats that are already structured. CSV files are an example of structured data. It is also possible to go all the way to true RDF format, the champion of “reusability“.

I am not an expert in all formats, for example, those for visual geographic data or maps. I am not an expert on the part of those files that is digital data and the part that enables them to be viewed. There is certainly a data subset that could be stored in tabular format.

The idea is to structure the information. That is what will make the data more open in the intended sense. The fact that the data is accessible to Canadians poses no problem whatsoever, but that is not what actually makes for open data. You may have 190,000 data sets, go through them and try to find something of interest. However, the principal of open data is about having more easily reusable formats.

The Chair: Thank you, Ms. Da Sylva.

To conclude, I will invite Ms. Ablonczy to ask the last question.

[*English*]

Hon. Diane Ablonczy: Thank you, Mr. Chair.

I appreciated the responses to my question about the practical benefits of this initiative. They were very helpful.

To follow up, I wonder whether countries that are committed to this course of action are confining themselves to just putting as much data as possible out into the public domain, or whether they are linking, prioritizing, or triaging information based on the results that are anticipated and desired.

How is this being approached? Is it in a logical, prioritizing fashion, or is data just being put out with the hope that at some point somebody will do something practical with it?

Maybe Mr. Stirling should begin.

● (1040)

Mr. Richard Stirling: The U.K. government is taking an approach of prioritizing and triaging data sets around impacts. That's partly for the practical reason that government is huge and you need to start somewhere, so you might as well start with the high-value stuff.

This linked back into the discussions that we've had with the sector panels and the open data user group and the feedback mechanisms, hearing what the existing industry would find useful, but also hearing what the innovative, disruptive start-ups also would find useful, and using that to then set the backlog and prioritize the data sets that they'll focus on day to day.

Now that is something that we also play a part in, in that we try to bring our start-up and our membership community...and act as their voice to government, to say this is what would be incredibly helpful.

Hon. Diane Ablonczy: Ms. Ubaldi , in other jurisdictions, how is that working?

Ms. Barbara-Chiara Ubaldi: The tendency is to prioritize the release of open data sets based on the needs and the values of the community of users. That implies a big effort made by governments right now to better understand the demand, meaning to engage with the various groups of actors, which is not an easy task, but they're moving towards that direction.

We are seeing a new level of sophistication. The governments may have started with opening as much as possible, and now they are becoming more targeted in prioritizing which data sets to release.

Hon. Diane Ablonczy: Ms. Bates, are you finding that this sort of refining of the initiative is bearing some fruit? Is that starting to draw in more end users in a practical way, or does more research need to be done at this point?

Dr. Joanne Bates: I think more research would need to be done to ascertain whether that approach was drawing in more users.

I think in the U.K. the approach seems to have been to develop a business case and the higher economic impact scores that you've seen in the open data barometer compared to the example of social impact scores perhaps indicate that the approach is leading to impacting the economic sector.

I would also just point out that in the U.K. the approach has been towards people developing business cases for opening up government data, so it's very economically driven.

Although there is a data unlocking service, civil society accessors have found it quite difficult to access the data sets they need through that service, so the triaging that is going on seems to be directed towards certain ends at the minute.

Hon. Diane Ablonczy: How much time do I have?

The Chair: You have 35 seconds.

Hon. Diane Ablonczy: Thirty-five seconds....

Ms. Da Sylva, from the point of view of an end user, then, do you see greater focus starting to occur, and is it something that is useful to you?

Mme Lyne Da Sylva: To focus on...?

Hon. Diane Ablonczy: To focus on what types of information is spread out on a priority basis.

Ms. Lyne Da Sylva: I think we're still at a discovery stage. A lot of users are just finding out that this is available and are having fun sorting through it, but I'm not sure we're at that point.

[*Translation*]

The Chair: Thank you.

I would like to thank the witnesses and all of the committee members.

This brings our meeting to an end. I am sure that the expertise of the witnesses will assist our committee in its study on open data. It will certainly assist in the drafting of its report.

Once again, thank you for appearing before us.

A reminder to committee members that we will meet again in two weeks. Thank you.

The meeting is adjourned.