HOUSE OF COMMONS
CHAMBRE DES COMMUNES
CANADA

Standing Committee on the Status of Women

FEWO • NUMBER 037 • 1st SESSION • 42nd PARLIAMENT

EVIDENCE

Monday, December 5, 2016

Chair

Ms. Marilyn Gladu

Standing Committee on the Status of Women

Monday, December 5, 2016

● (1530)

[*English*]

The Chair (Ms. Marilyn Gladu (Sarnia—Lambton, CPC)): I'll call the meeting to order.

I'm very excited to see Ruby Sahota back with us today.

Welcome.

Bryan, welcome to the committee.

This is going to be a good topic.

We have with us today, from MediaSmarts, Matthew Johnson, who is the Director of Education, and Jane Bailey, who is a professor with the Faculty of Law at the University of Ottawa. I'm going to invite them to open with their comments, beginning with Jane.

Ms. Jane Bailey (Professor, Faculty of Law, University of Ottawa, As an Individual): Thank you very much for inviting me back.

I understand that today one of the things we've been asked to focus on is this notion of algorithmic curation. I'm making these remarks as the co-leader of the eQuality Project, which is a project that in fact is focused on the big data environment and its impacts on online conflict between young people. I'm also a member of the steering committee of the National Association of Women and the Law.

Big data, or the big data environment, where each of us trade our data online for the services we get, is a mechanism for sorting all of us, including young people, into categories in an attempt to predict what we will do based on what we've done in the past and also to influence our behaviour in the future, especially around marketing, to encourage us to purchase certain goods or to consume in certain ways.

In terms of our concerns at the eQuality Project with the big data model, and with algorithmic sorting in particular, there are three that I want to touch on.

The first is this assumption that the past predicts the future. This can become a self-fulfilling prophecy, which in the context of youth is particularly concerning. The assumption is not only that what we do predicts what we will do individually in the future, but that what people who are assumed to be like us will do or have done in the past somehow predicts what we as individuals will do in the future.

We can begin with an example that will appear soon in the eQuality Project annual report, courtesy of my co-leader Valerie Steeves. Think about online advertising and targeting. If you are a racialized male online and the algorithmic sort sorts racialized males as people who are more likely to commit crimes, then the advertising targeted to those people in that category—the young racialized male—might lean more toward names of criminal lawyers and ads for searching out people's criminal records, as opposed to advertising for law schools, which might be the kind of advertising that a middle-class white young person might get. There's a study by Latanya Sweeney about this.

The shaping of our online experience, that information to which we have access, according to our algorithmic sorting into groups, then can become a bit of a self-fulfilling prophecy because it's assumed that there's certain information that's relevant to us, and that's the information that we have access to. I don't know if you have ever sat side by side with someone and done a Google search and have seen that you get different results. That's one thing. The assumption that the past predicts the future is problematic in a very conservative way. It's problematic when the groups that we're using are based on discriminatory categories as well.

The second problem obviously is the constraint that this imposes on change, the constraint that it imposes on people's equal capacity to participate and to grow. In the context of young people, our concern is around whether young people will be influenced in ways such that they internalize the stereotypes that are wallpapering their online spaces, how internalization of that stereotype may affect their self-presentation, their self-understanding, and their understanding of their possibilities for future growth and participation, and in what ways this may set youth up for conflict with one another and set youth up to judge each other according to the stereotype's marketed standards that are part of the algorithmic sort in an online environment.

The third problem that we're particularly concerned with is the lack of transparency, of course, around this algorithmic sort. We cannot question it. Most people, even people who are computer programmers, don't necessarily understand the outcomes of the algorithmic sort. When important decisions are getting made about people's lives, such as what information they have access to and what categories they're being sorted into, and we have a system that we are not allowed to question, that isn't required to be transparent, and that isn't required to provide us with an explanation of why it is we've been sorted in this particular way, there are obviously serious democratic issues.

● (1535)

Again, our concern in the eQuality Project is to focus on the impact that this has on young people, particularly young people from vulnerable communities, which includes girls.

What to do about this?

One of the important points, which came from earlier work that I did with Professor Steeves in the eGirls Project, is that more surveillance is not the solution. The big data algorithmic environment is a surveillance environment. It's a corporate surveillance environment and, of course, the corporate collection of this data spills over into the public environment, because it creates opportunities for public law enforcement access to this data.

What the girls in the eGirls Project told us about their experiences in the online environment was that surveillance was a problem and not a solution. Algorithmic sort solutions that purport to categorize young people according to surveillance of their data instill greater distrust for young people and adults, and greater distrust of young people in the systems they're using.

I think it's really important to think about refocusing and reshaping our concerns on corporate practices here, rather than on training children to accept an algorithmic model, to accept that they're going to be sorted in this particular way. We should take a step back and ask corporations to better explain their practices—the how, the why, the when—and to consider regulation if necessary, including to require that explanations be provided where decisions are being made about a young people's life chances according to algorithmic curation and sorting.

Those are my remarks for now.

The Chair: That was excellent. Thank you.

Mr. Johnson.

Mr. Matthew Johnson (Director of Education, MediaSmarts): Thank you to the committee for inviting MediaSmarts to testify on this issue.

Our research suggests that algorithms and the collection of the data that make them work are poorly understood by youth. Only one in six young Canadians feel that the companies that operate social networks should be able to access the information they post there, and just one in 20 think advertisers should be able to access that information, but almost half of youth appear to be unaware that this is how most of these businesses make money.

With support from the Office of the Privacy Commissioner, we've been creating resources to educate youth about this issue and to teach them how to take greater control of their online privacy.

Algorithmic content curation is relevant to cyber-violence and youth in a number of ways. When algorithms are used to determine what content users see, they can make it a challenge to management one's online privacy and reputation. Because algorithms are typically mostly opaque in terms of how they work, it can be hard to manage your online reputation if you don't understand why certain content appears at the top of searches for you. Algorithms can also present problems in terms of how they deliver content, because they an embody their creator's conscious or unconscious biases and prejudices.

I believe Ms. Chemaly testified before this committee about how women may be shown different want ads than men. There are other examples that are perhaps more closely related to cyber-violence. Auto-correct programs that won't complete the words “rape” or “abortion”, for example, or Internet content filters, which are often used in schools, may prevent students from accessing legitimate information about sexual health or sexual identity.

This is why it remains vital that youth learn both digital and media literacy skills. One of the core concepts of media literacy is the idea that all media texts have social and political implications, even if those weren't consciously intended by the producers. This is entirely true of algorithms as well and may be particularly relevant because we're so rarely aware of how algorithms are operating and how they influence content that we see.

Even if there is no conscious bias involved in the design of algorithms, they can be the product and embodiment of our unconscious assumptions, such as one algorithm that led to a delivery service not being offered in minority neighbourhoods in the United States. Similarly, algorithms that are designed primarily to solve a technical problem, without any consideration of the possible social implications, may lead to unequal or even harmful results entirely accidentally.

At the same time, a group that is skilled at gaming algorithms can amplify harassment by what's called “brigading”: boosting harmful content in ways that make it seem more relevant to the algorithm, which can place it higher in search results or make it more likely to be delivered to audiences as a trending topic. This was an identified problem in the recent U.S. election, where various groups successfully manipulated several social networks' content algorithms to spread fake news stories. Also, it could be easily used to greatly magnify the reach of an embarrassing or intimate photo, for example, that was shared without the subject's consent.

Manipulating algorithms in this way can also be used to essentially silence victims of cyber-violence, especially in platforms that allow for downvoting content as well as upvoting.

In terms of digital literacy, it's clear that we need to teach students how to recognize false and biased information. Our research has found that youth are least likely to take steps to authenticate information that comes to them via social media, which, of course, is where they get most of their information. We need to educate them about the role that algorithms play in deciding what information they see. We also need to promote digital citizenship, both in terms of using counter-speech to confront hate and harassment, and in terms of understanding and exercising their rights as citizens and consumers. For example, there have been a number of cases where consumer action has successfully led to modifying algorithms that were seen to embody racist or sexist attitudes.

Thank you.

● (1540)

The Chair: That was excellent. We will go to our first round of questioning.

We're going to start with you, Ms. Vandenbeld, for seven minutes.

Ms. Anita Vandenbeld (Ottawa West—Nepean, Lib.): Thank you.

Thank you very much for returning to help us delve a little deeper into some of these topics. I have a few questions for clarification, and then I'd like to talk a bit about the regulation of corporations and how we can have more transparency.

First of all, this is the first time I've heard about brigading and about manipulating the algorithms, which of course is quite alarming. Could I hear a bit more about that from Mr. Johnson?

Also, for Ms. Bailey, my understanding has always been that what you put into search terms determines the kinds of things that you see. From your testimony, it sounds like it's bigger than that. It's also more predictive. It's the group that you're in. I'm not entirely sure how that can be programmed in. Could you both clarify that?

Then I'd like to hear more about the corporate regulation.

Mr. Johnson.

Mr. Matthew Johnson: I'm not really sure what to add on the topic. We've seen a number of cases in which, either through what you might call savvy gaming of algorithms or, in some cases, just brute force, people have been able to manipulate things such as trending topics.

One study that was done of fake news leading up to the U.S. election found that there was actually a small group of writers in Macedonia producing content who managed to get it spread to a tremendous number of people, because they understood how the initial readers were going to interact with it and how that would influence how the platforms promoted that content. To my knowledge, I haven't seen that used yet on a wide scale for harassment of individuals, but the same technique certainly could easily be used.

Ms. Anita Vandenbeld: You suggested that it can also be used not just to amplify certain news items, but also to silence others. How does that happen?

Mr. Matthew Johnson: In part, it happens because any time one message is being amplified, others get lowered. Someone's message is less likely to be trending if someone else's is. There's a limited number of spots.

Some platforms also have what's called “downvoting”, in which users not only can boost a signal of one thing but can also say that another thing is less relevant. If you have a savvy group that is boosting the signal of the harassment and, when the victim and the victim's allies are trying to get a message out, they are downvoting that, they can essentially be working in both directions. We did see that to a certain extent with the “Gamergate” situation, in which women were being harassed in the games industry. Some of those techniques were being used, although they weren't quite as technically sophisticated as what we've seen recently with the fake news situation.

● (1545)

Ms. Anita Vandenbeld: Basically, those who understand how the algorithm works and anticipates human behaviour can then manipulate what gets amplified, so that if something is repeated often enough, people think it's true. Is that it? Okay.

Voices: Oh, oh!

Ms. Anita Vandenbeld: I'd like to move to Ms. Bailey and talk a bit about what you were saying in terms of the transparency. Of course, we're looking at where the federal government would be able to find remedies for these sorts of things, so are there ways in which we can force corporations to be more transparent—whether that would actually solve the problem—and how would we do that?

Ms. Jane Bailey: There are models in the EU in particular, in the EU directives around data privacy, that focus more on bringing human decision-making into the loop. Where a decision is made that affects someone's life chances, for example, there needs to be some sort of human element in the determination of the result.

Again, this adds a certain level of accountability or transparency, where neither you nor I—or maybe even some computer scientist—could actually explain what the algorithm did in terms of how it came to the conclusion that you were in a particular group, or that certain information should come to you or not. Thus, we can have some other form of explanation about what is actually being taken into account in determining what kind of information it is that we're seeing and why a particular decision is being made about us. This is becoming more and more important as we move toward machine-made decision-making in all kinds of atmospheres.

I think people or countries are beginning to think about ways in terms of how to put the “public” in public values and public discourse back into decision-making in this sphere, which, although it is largely privately controlled is really a public infrastructure, in terms of a necessity for people to have access to it increasingly for work, for social life, and for education. It's about how to think about righting the balance between the decisions being made from a private sector perspective—not for nefarious reasons, but for profit reasons, because that's what they're in business to do—and how we re-inject public conversation and public discourse around the issues in terms of what's happening, what kinds of decisions are being made, how people are being profiled, and how they're being categorized. I think this is a really important start.

Ms. Anita Vandenbeld: So essentially there's a multiplier effect. It's beyond just an echo chamber of all your friends on social media liking things and therefore you'll see only the same things. This is as opposed to the old days, when you'd flip through the newspaper and be exposed to all kinds of things. Now the algorithm is picking that up and then reinforcing it. Is that...?

Ms. Jane Bailey: It could be that some do that. In other words, they close our circle instead of opening our circle.

In some cases, people who look at this may say, well, that's an advantage, because if I don't want to see hate speech, I don't have to see hate speech. But let's take Twitter's mute button as an example. I can mute somebody so that I don't see that they are attacking me online, but the fact of the matter is that they are attacking me online and I don't know about that.

The way our worlds are being curated is that in some instances we might say that at least it relieves my pain in the initial moment. But in the long term, in terms of what violence is being done, what harassment is happening, and what issues we really need to be engaging in, it's a problem if we are closing ourselves off.

● (1550)

The Chair: I'm sorry. That's your time.

I'll go to Ms. Vecchio now, for seven minutes.

Mrs. Karen Vecchio (Elgin—Middlesex—London, CPC): Thank you.

Hi, and thank you very much for coming.

I want to start off with a personal story. Maybe you can share with me how this came about. You can say, no, these were algorithms—or maybe I had a bad past I don't know about—but what happened is this. I was on a flight the other day, and I watched two YouTube videos, parts one and two from the international advertising awards. I'll share with everybody that they were regarding men's underwear. It was a funny clip—very funny; two testicles; great.

After the first two videos, the third video, which automatically went to play, was pornography. It was a young man and a young woman. Unfortunately, I was sitting there with my 13-year-old son, and I went, “Oh, my gosh”, because the video itself that I was watching with my son wasn't too inappropriate—somewhat, but not too inappropriate—but I can tell you that the third thing absolutely should not have been there.

Would that have been an algorithm? Would that have been something from a previous search history, although I can tell you that I've never searched for pornography on YouTube? How would that have come up? Can you share with me your thoughts on how you start with something that's getting a national award for advertising and the third thing is pornography?

Ms. Jane Bailey: First of all, let me give a disclaimer. I'm a lawyer, not a computer scientist, so to say specifically what was happening, I'm not sure. It could be many things. It could be an example of the sort of thing that Matthew was talking about, that there's an algorithm that's calculating what people who look at these two videos tend to like. It could reflect other users' preferences and be associating things together. It could reflect some less sophisticated algorithm that's searching for a term or a content like underwear, or testicles, or whatever the case may be. It's aggregating like content in that way. Or it could be, although you've told us that it's not, based on your own personal search history.

Mrs. Karen Vecchio: I swear it's not.

Voices: Oh, oh!

Ms. Jane Bailey: The more sophisticated algorithms get, the better they are supposed to be at predicting what we would actually want to see. If I look at my search history on Amazon, Amazon did this fairly early. I bought a lot of books about feminism, so Amazon constantly gave me ad suggestions that had anything to do with women, especially diet and exercise books. I was like, “I think your algorithm doesn't get it. An old, rad feminist is not looking for Suzanne Somers' diet and exercise book. You guys are way off.”

Voices: Oh, oh!

Ms. Jane Bailey: What you can see is that the algorithms are getting more sophisticated, and the more data you give them, the more predictive they become in terms of thinking about emulating your behaviour to the extent that your behaviour is premised on what you did in the past.

Mrs. Karen Vecchio: Okay. I just found it extremely.... Oh, my goodness, I didn't expect it, especially from the documentary that it came from. Actually, some of the other things were Pampers commercials, so you didn't expect something to fall into this. I was really quite surprised.

When we talk about this, I think that brings us into what protective measures we could also use. For something so simple as a commercial that within two plays gives us this sort of result, what can we do to protect in the long run? Is there a way we can downvote something when it comes to pornography on the Internet? Is there anything we could do there, not necessarily legislative, to make sure we're being more cautious? Is anything like that currently being done here in Canada?

Mr. Matthew Johnson: Yes, absolutely.

One of the things that we educate young people about is, again, digital citizenship: their ability to make a difference online. We teach them, for instance, that when they see inappropriate content, particularly when it's something like cyber-bullying or hate content, there are a lot of steps they can take. Almost every platform, whether it's a video platform or a social network, has ways of reporting content. Many of them do have downvoting. That's one of the reasons downvoting exists, even though it can be misused. We teach them that they have a responsibility to do that, and that they have a right to have an online experience where they're not exposed to harassment and hate.

We also advocate and provide resources for parents to talk to their kids and for teachers to teach students about all of these different issues. We know that kids are going to be exposed to them, whether intentionally or unintentionally. We know it happens. We know that even the best filters don't block out all of this content, and often, when it comes to things like hate or cyber-bullying, filters don't do a good job.

It's important that we talk about these things, so that by the first time someone encounters pornography, they already know that it's not real, and they already know not to take it as a realistic or healthy view of sexuality.

● (1555)

Mrs. Karen Vecchio: Excellent. Thank you.

Jane, do you have anything to add?

Ms. Jane Bailey: No.

Mrs. Karen Vecchio: I'll carry on.

Those are of my some major concerns when we're looking at these algorithms: what is it that we can be doing better?

When it comes to the education of our youth, we know that there are tips if you go to MediaSmarts and things of that sort. What other efforts are we making? Can you talk of any education systems—any school boards or anything like that—that are implementing these things in their courses? What's being done by Canadians to educate their youth, other than within the family, because sometimes, as well, the family has no idea.

Mr. Matthew Johnson: There is an increased presence of digital literacy in curricula across the country. One of the things we do is keep track of provincial and territorial curricula, partly to make sure our resources match those curricula so that teachers can use them without taking time away from what they need to be teaching. We are seeing an increase. British Columbia, for instance, now has a full digital literacy curriculum. We're seeing a lot more digital literacy in the health and social science curriculum in Ontario.

It's a work in progress. It's going to be a while before digital literacy is in the place that media literacy is in, where it's formally in the curriculum of every province and territory, but we're certainly seeing an improvement. Obviously, we're working to make it easier for teachers and the school boards to get digital literacy into the classroom, and to do so from kindergarten all the way up to grade 12, because we know that in many cases we need to start teaching about these issues as soon as kids are using the Internet.

Mrs. Karen Vecchio: Absolutely.

Thank you very much.

The Chair: We'll go to Ms. Malcolmson for seven minutes.

Ms. Sheila Malcolmson (Nanaimo—Ladysmith, NDP): Thank you, Madam Chair.

Thanks to the witnesses. It's good to see you back again.

It was last June, at the very beginning of our study, when we talked with you, Ms. Bailey. I'm glad you're here again.

Just last month, the United Nations committee to end discrimination against women issued its report on Canada. The report comes out every five years, and it's a good opportunity for us to touch in. One of the items that the committee noted was this concern, and I'll quote:

> The repeal of section 13 of the Canadian Human Rights Act, which provided a civil remedy to victims of cyber violence, and the enactment of the Protecting Canadians from Online Crime Act (2015), which penalizes the non-consensual distribution of intimate images, but fails to cover all situations that were previously covered by section 13 of the...Human Rights Act.

The recommendation from the committee in paragraph 25(g) is that the federal government:

> Review and amend legislation in order to provide an adequate civil remedy to victims of cyber violence and reintroduce section 13 of the Canadian Human Rights Act.

Do either of you, in your professional experience, have any advice for the committee on a recommendation that we might reinforce in that area?

Ms. Jane Bailey: On the repeal of section 13 of the Canadian Human Rights Act under the prior government, I testified before the Senate about that. It came, I would say, at the most ironic time in history. It was a time when everyone was talking about the impact of online hate and harassment. Canada was uniquely placed in having a federal human rights provision that allowed for a tribunal rather than a court to respond to online hate and harassment as a human rights issue, hate and harassment that's identity based. We were uniquely and proudly situated in Canada to have had that remedy.

Right at the time, I would have said, when the remedy had its most meaning, when most experts were saying the way to respond to this was not going to be primarily through criminal law remedies but through a human rights approach, Canada chose to repeal section 13. I think that was an unfortunate decision. It hobbled Canada's ability to deal effectively with online hate and harassment and to offer a variety of responses. Also, it's not just that. It's the symbolic recognition that what's underlying these attacks is harassment, discrimination, and prejudice based on identity.

To me, the reinstatement of section 13 of the Canadian Human Rights Act would make a lot of sense at this time, because, with all due respect, it made no sense to repeal it at the time that it was repealed.

In terms of civil remedies, I think one of the more interesting civil remedies that I've looked at recently is in Manitoba, where they are using the body that runs Cybertip to assist those whose intimate images are posted online and to get the images taken down in a quick way. I think that's a very sort of meaningful support mechanism. It's one of the number one issues for those who are victims of non-consensual distribution: to get the image down as quickly as possible.

None of that is to negate the criminal law provision, but these things do something different. I think having a panoply of different legal responses that suit different people in different situations and their abilities and needs makes a lot of sense.

● (1600)

Ms. Sheila Malcolmson: Thank you.

Mr. Johnson.

Mr. Matthew Johnson: As an educational organization, MediaSmarts doesn't take a position on legislation unless it's directly relevant to our educational mission, but as part of our mission, we do educate adults and youth about their legal rights and the remedies available to them should they choose to use them, and the ways in which they can act positively, particularly online in relation to media. I can say that section 13, when it was still in existence, was certainly a big part of our materials on how to deal with online hate. There's certainly nothing like it in existence today that we've been able to replace in terms of those materials.

Ms. Sheila Malcolmson: If it were restored, it would be a remedy that you would return to recommending to your clients?

Mr. Matthew Johnson: We certainly would make them aware of it. We don't necessarily recommend, but as I said, we certainly would make sure that all of our resources featured it, because it would be another tool that could be used for dealing with hate material on the Internet or other media.

Ms. Sheila Malcolmson: Thanks.

There's only a minute left, so from the MediaSmarts side, I imagine that, like other organizations we've been hearing from, as we've raised awareness about the need to speak out and ask for help, your organization is increasingly being asked for help. Do you have sufficient operating funding to accommodate the requests and the needs that are coming your way? A corollary of this is that we don't want to ask people to ask for help and then have the door closed when they get there.

Mr. Matthew Johnson: We don't really have stable operating funding. We certainly get tremendous support from a lot of sources, from our sponsors, but I'm not going to list them all, because it would use up the time we have. When it comes to things like online hate, it's very difficult for us to find the time to put into it or funding for it. Our most recent project on that was in 2011. It can be very difficult for us to deal with issues that come up, even when they're urgent, when there isn't specific funding tied to those.

Ms. Sheila Malcolmson: Thank you for your work.

The Chair: We'll now go to Ms. Damoff for seven minutes.

Ms. Pam Damoff (Oakville North—Burlington, Lib.): Thank you very much.

I can't thank you enough for coming back. You gave us outstanding testimony when you were here before, and you've done so again, so I want to thank you both for the work you're doing and for the information you've shared.

The term “algorithm” kept coming up. It's still a little hard to believe. The chair and I were looking at each other as you were describing a few things and saying that it's a little hard to believe what's going on when we are on the Internet.

You've talked about the EU bringing in some data privacy legislation. I'm wondering about that. So many of these companies are global companies. For example, we had Twitter here, and their head office is in the United States but they operate globally. It's the same for Facebook. A lot of these companies are global companies. Because these companies are global, does bringing in legislation in Canada actually have any impact on what they're doing?

● (1605)

Ms. Jane Bailey: That's an interesting question. It's a point for obfuscation, I think, that corporations like to use: multiple jurisdictions, hard to keep track.... Yes, agreed, but when you are doing business in a particular jurisdiction, you have to expect to abide by the laws in that jurisdiction.

In some ways, I think, it's one of the reasons why the EU approach is so impressive. That approach has been very much to say that they have a set of values, and that they are not against development and innovation, obviously, but they have a set of values, and they have directives and legislation that you're required to comply with. You can tell them that it's expensive and complicated, and that you don't like it, but at the end of the day, that is where the buck stops.

Ms. Pam Damoff: If I can give you an example, though, you were talking about the fake news sites, with a lot of it originating in Macedonia. If you had another country messing with the algorithms and it's happening in Canada, would our legislation have any impact on that?

Ms. Jane Bailey: It's hard to say. It would depend on how you drafted the legislation.

Ms. Pam Damoff: Let's say that it was like the EU legislation.

Ms. Jane Bailey: We can look at “the right to be forgotten” in the EU. An example is the Google case in Spain, where Google was upset that EU law was applying to their situation, because they didn't think their presence in Spain was sufficient to justify the application of that particular directive.

In brief, the case was about whether they could order Google to delete a particular outcome from its search engine, so that if you searched for a particular person, this story would not surface in the Google search, on the premise that most of us get our information from Google searches, and even if it's still out there on a website somewhere, our access to it is relatively limited if it doesn't come up in the first one or two pages of a Google search. They were ordered to remove this from their search, so that when someone searched for this individual, the story about a prior proceeding wouldn't come up. The idea was that if you don't want to do this broadly, you need to figure out how to do it so that residents in Spain or in the EU don't have access to this material, because this person is entitled, under EU law, to not have other people in the EU getting access to this on their search engines.

It can be done.

Ms. Pam Damoff: Okay.

Ms. Jane Bailey: On the other question about a different kind of algorithmic sort, I think my answer would be that I'm not sure. It would depend on what was happening. That isn't to say that it's not complicated. It's just to say that sometimes it's the first thing that gets put on the table, and I think there is a corporate interest in its being the first thing that gets put on the table.

Ms. Pam Damoff: One of the things—you brought this up the last time you appeared before us—is the importance of digital literacy. You've mentioned it again today. Some of that is obviously falling within provincial jurisdiction. What can the federal government to in order to promote digital literacy? Do you have any suggestions?

Ms. Jane Bailey: Do you want to go first, Matthew?

Ms. Pam Damoff: Because that seems to be one of the things that would help. If you even knew this was going on, you'd be looking at it, whether in search engines or your news feed, with a different lens than we are now.

● (1610)

Mr. Matthew Johnson: There are a number of things that I think can be done at the federal level. One of those is supporting the development of digital literacy resources. Most of the provinces that are adopting digital literacy in the curriculum don't necessarily have resources for teachers to use, and certainly that can help them. One of the big risks in digital literacy education is “silo-ization”; that is, efforts and energy are being wasted because the same wheel is being reinvented 13 times across the country. I think federal efforts can certainly help to prevent that.

Also, certainly, there is making the public in general more aware of digital literacy as an issue and digital literacy skills as life skills that all of us need at all of our stages in life. We need also to incorporate it in early childhood education, making sure, again, that it begins as soon as young people are using digital devices.

Ms. Pam Damoff: Thank you.

Jane.

Ms. Jane Bailey: The other thing at the federal level, of course, is that we have the Office of the Privacy Commissioner of Canada. We could think more about giving them more power, about giving them real enforcement authority and the authority to deal with algorithmic curation kinds of issues as well. Compared to many jurisdictions in the world, our PIPEDA is an important piece of legislation in terms of public control over private organizations and what they do with data. From a federal perspective, thinking about strengthening the powers and the jurisdiction of that body, I think, is certainly something that should be on the table.

Ms. Pam Damoff: Thank you.

The Chair: We'll go to Ms. Harder for seven minutes. She'll be sharing her time with Ms. Vecchio.

Ms. Rachael Harder (Lethbridge, CPC): Thank you so much.

I apologize for not being here for your presentations. Unfortunately, I was called out for a moment. I am certainly one of the strongest advocates for having you at the table today, and I appreciate your time.

That said, I have a few questions for you.

My first question is for each of you, if you don't mind. What can be done to prevent pornographic images or videos from coming up on my newsfeed when I'm just searching? My nieces search innocent things all the time and come up with crazy images that pop up on the screen. I understand that my colleague Karen has asked a similar question, but can you expand on this? What can be done to prevent these things from happening to our children?

Mr. Matthew Johnson: Technically, we have a number of easy steps that can be partially effective.

If you're talking on an individual level, almost every search engine has a “SafeSearch” setting. There are also content filters that are available. Most ISPs make those available. There are commercial filter programs as well. These are never going to be 100% effective, particularly when you broaden your definition of inappropriate content beyond just nudity. There are certainly things that we recommend, especially using something free like the SafeSearch.

This is one of the reasons why we approach digital literary in a holistic way. This is why things like authentication and search skills address content issues as well. One of the best ways to avoid finding this is having sufficient search skills so that you're looking for only the one thing that you're looking for, so that you're able to craft a successful search string that will narrow out things you don't want.

There are certainly also steps that you can take to avoid having a profile built. If you watch a video that may, for whatever reason, have inappropriate content algorithmically connected with it, if you're not having a profile of you built, it's going to have less of an effect. There are measures like using search engines that don't collect data on you, possibly using an IP proxy, or using, in some cases, the incognito modes of browsers, or activating the do-not-track function in browsers.

All of those, again, are incomplete on their own. Again, that's why we say that you can never entirely shield young people. That's why we have to talk about these issues. Those are all effective steps that you can take to reduce the odds of those things happening.

● (1615)

Ms. Rachael Harder: Jane, your brief comments, please.

Ms. Jane Bailey: The sorts of things we're doing as damage control, which Matthew gave us a really good rundown on, have other implications, but aside from that, I think my answer to everything, really, as you may remember—I said this the last time too—is to end patriarchy, and then we won't have to see unwanted pornography. There are root causes here that are the issue. We can and we have to do damage control, because we want to educate our kids and we want to protect our kids. Kids need to know how to deal with this content, and they need to be able to think critically about this content as well.

At the end of the day, if violent pornography is an issue, that's a systemic issue. We have to take care of misogyny.

Ms. Rachael Harder: In your estimation, is it an issue?

Ms. Jane Bailey: Is violent pornography an issue?

Ms. Rachael Harder: Yes.

Ms. Jane Bailey: Certainly.

Ms. Rachael Harder: What exactly is the issue, if you were to get to the heart of it? You have 30 seconds.

Ms. Jane Bailey: The heart of the issue is misogyny. The heart of the issue is representation of rape or sexual assault as sex. We shouldn't be confused about that. That should not be confused. That's the heart of it. It's overlain by all kinds of other intersections, such as racism, classism, and ableism. It's overlaid by all of those things, but if the heart of an industry is to make money from enacting sexual violence against women, then we have some hard questions to ask ourselves about what society we're living in and what kinds of industries we're supporting.

The Chair: All right.

We're going to Ms. Sahota for five minutes.

Ms. Ruby Sahota (Brampton North, Lib.): I want to get a bit more information about upvoting content and downvoting content. Those are terms I'm not familiar with. Can you break it down for me in terms of popular sites that we would search on and how we would do such things?

Mr. Matthew Johnson: I'm using upvoting and downvoting in a generic sense. I'm using “upvoting” to mean taking any action that boosts content, that spreads it, and particularly that makes it seem more relevant to the algorithm, and “downvoting” to mean anything that does the opposite, that limits the reach or makes it seem less relevant.

Each platform does that in a different way. An easy example would be “liking” something on Facebook, which is a way of upvoting it, because in future,things that you “like” will be seen as more relevant to you. Facebook is more likely to show you that if you've selected “most relevant” rather than “most recent”. You do have the option on Facebook of toggling to just a straight timeline, but the default normally is to be shown what the algorithm feels is relevant to you.

The reddit platform, for example, has pure upvoting and downvoting. In reddit, each user can literally boost something by making it more popular or drop it by making it less popular. In the case of reddit, that's also a big issue in terms of what appears on the front page of the site, which is to say what you see when you just go to reddit.com, rather than one of the many sub-reddits. That is something that we know hate groups have manipulated. They have made an effort to get certain hateful messages to the home page by getting enough people to upvote them, and again, when they've decided to target particular critics, they have tactically downvoted them in the same way.

Ms. Ruby Sahota: I've heard about the delivery issue that you were talking about in certain areas in the U.S., that being a case that was brought up. What disturbed me a bit in the presentation was the effect on certain outcomes of a person's life, not just with regard to the content that you're viewing currently, but in regard to the long-term effects that this could possibly have.

Can you shed more light on how somebody could consciously upvote or downvote something to perhaps get rid of something like this, when it is so subconscious and maybe no one is doing that themselves on the computer intentionally...? It's very difficult for me to understand how these people are getting targeted, especially if they're so young. You're talking about criminal record checks versus law school advertisements. How would that ever end up happening? Would it just be the demographics of where the person lives?

● (1620)

Ms. Jane Bailey: Yes, that's the thing: it's a kind of a tragedy of the commons problem, where, when you and I make individual decisions in individual situations, and we think we're fine because we've agreed to what we've done, the implications of what we've done, our choices, can be part of what aggregates. It's the algorithmic sort of aggregation.

I'll give you an example from Latanya Sweeney's research. She did research in the United States which showed that black-sounding names were more likely to have pop-up advertising for services that allowed you to get a criminal record check than white-sounding names were. The advertising itself reflected embedded prejudice.

Then the question became, how did that happen? The search engine said, “Well, it's not us, we didn't program in a prejudice.” They said that it must be that the algorithm was reflecting societal prejudice. They said that it was more likely in the databases that we're searching that more people are searching for a criminal record check on a black-sounding name than on a white-sounding name, so they put it back as a reflection of consumers.

Part of the answer is that we won't necessarily know, but it's a powerful indicator of how it can happen, whether or not.... The algorithm curates our aggregate bias and our aggregate discrimination and feeds it back to us in ways that obviously have disparate impacts on members of marginalized communities, impacts that are not felt by members of the majority. It's complicated.

The Chair: All right. That's your time.

We're going to go to Ms. Harder for five minutes.

Ms. Rachael Harder: I'm wondering if it's possible to change algorithms to pick up on buzzwords. For example, Twitter once picked up the buzzword “slut”, and said that it was a negative word, that anytime it was used, it was negative. Then they began to realize that most of the time when people were making use of the word “slut”, it wasn't necessarily negative, so they changed their algorithm.

My question, then, is that if algorithms can be changed in that way in order to guide us as users, would it be possible to change algorithms in such a way as to be able to prevent underage Internet users from having access to pornography?

Mr. Matthew Johnson: I think that's a question that's a little more technical, if I can speak for Jane, than either of us is qualified to answer.

What I will say is that I'm not certain that there is any power in the world that will prevent teenagers from accessing pornography. I don't mean to be flippant with that. Certainly, there are tools that can be used, and I know there are tools that are being used. This is being discussed in the U.K. right now. I believe legislation has just been passed on that very topic. There's a lot of discussion going on, in terms of Internet safety in the U.K., about how to actually make this work and whether it is worth doing.

Certainly, there are ways of identifying or guessing people's ages. That's a big part of what algorithms do, because part of your profile is how old you are, but there are a lot of technical challenges to something like that. Like all other blocking or filtering tools, it's never going to be 100% effective, and there's a very good chance that it will result in a lot of false positives, like most blocking and filtering tools do. At the very best, it would be a complement to a digital and media literacy approach to pornography.

● (1625)

Ms. Jane Bailey: I'll jump in there. The other thing we have to be conscious of is that I don't think we actually want to be keeping kids from sexually explicit material. I think there's a lot of information that's necessary for kids to know about sexual activity and sexual health, which I distinguish from violent pornography. The idea of surveilling kids to prevent them from access to content about sexuality I think would be a real problem, whether or not you algorithmically distinguish between violent pornography—which in my view isn't just a problem for kids but a problem for adults too—and sexually explicit material, which is important for people to have access to. That's another problem.

Filters often over-filter, so that you don't have access to material that's important for sexual health, for example, or for developmentally appropriate sexual curiosity and self-definition. Again, going back to eGirls, the girls told us that surveillance is a problem, not a solution. I'm not sure that mechanisms that are surveilling kids or blocking kids are necessarily the approach we want to take, even if scientifically we actually could design the algorithms to do that fairly well.

Mr. Matthew Johnson: I would add as well that most of the well-documented negative effects of pornography are also found to be caused by other forms of sexualized media that aren't explicit. For most of the things that we see in youth that we are fairly confident are caused or influenced by pornography are also caused by sexualized advertising, sexualized music videos, and so on. In some ways, pornography really is just the most extreme end, but blocking that is only going to have a very limited effect on those issues. We really need to take a broader media literacy look at gender and sexuality and those related issues to be effective.

The Chair: This has been a wonderful session. I want to thank both of our witnesses for being here and for illuminating us. Now we have even more questions, I'm sure, but we're out of time. Thanks for coming. We hope to see you again in the future.

I'm going to suspend while we switch panels.

● _____ (Pause) _____ ●

● (1630)

The Chair: I will call the meeting back to order. We are going to start our second panel discussion. I have a couple of announcements before we get to that.

I want to remind members that tomorrow is the National Day of Remembrance and Action on Violence against Women. You will remember that years ago the most savage, violent attack in Canada happened at École Polytechnique, and women engineers—I have to say that they were my sisters—were killed in an act of horrific gender violence. Please remember tomorrow. I know that we're not meeting because of votes in the evening, but I'm sure there will be other activities going on to remember that by.

The other thing I want to let you know is that when we were discussing our next study at committee and how we were going to move forward, we were going to have a bunch of the economic development area networks come and speak first. They've all declined to appear—amazing—so we have an opportunity instead to have one panel discussion with ISED, ESDC, and StatsCan, along with Status of Women. We could have that whole bunch come and talk to us in the first hour. For the second hour, the analysts have agreed to get our work plan ready by Friday and sent out to us, so that we can start talking about the work plan and at least agree on some of the initial meetings in the new year. Unless there's an objection, I'm going to suggest that we do that.

Without any further ado, we want to welcome our witnesses for this panel discussion. We have with us Sandra Robinson, who is an instructor at Carleton University. I will just let you know that Sandra wants to be sure she can hear your questions, so if you would ask them loudly and enunciate, that would be very good. We also have with us, from the Department of Industry, Corinne Charette, Senior Assistant Deputy Minister, Spectrum, Information Technologies, and Telecommunications Sector.

Welcome, ladies. We are going to give each of you seven minutes for your remarks.

We'll start with you, Sandra.

Dr. Sandra Robinson (Instructor, Carleton University, As an Individual): Thanks to the committee for the invitation today. It's a pleasure and a privilege to appear before you.

I am a full-time faculty member at Carleton University in communication and media studies. I teach in the areas of media and gender, law communication and culture, and algorithmic culture and data analytics on the more technical side. I'd like to share some concerns and considerations about the role of algorithms in the context of networked communications, such as those for social media, search, and, in particular, what is broadly conceived as automatic content curation by algorithms.

There's been some discussion of this already, obviously, so I'll focus on three things: defining algorithms and their operations; the trade-off between user interfaces and the increasing complexity of software; and, the impact of algorithmic content curation.

I want to be clear at the start about what I mean when I refer to an “algorithm”. In very simple terms and in the context of information systems and networked communication, it can be thought of as a series of computational steps or procedures that are carried out on information as an input to produce a particular output. For example, a search term typed in as input to “Google Search” produces an output in terms of search results.

Also, they don't operate in isolation. Algorithms are part of a complex network of digital devices, people, and processes constantly at work in our contemporary communication environment.

Embedded in any algorithmic system is a capacity for control over the information it analyzes, in that it curates or shapes the output, based on multiple factors or capacities the algorithm uses to generate the outputs. Again, in the case of Google Search, their suite of algorithms takes in the search term, personal search history, similar aggregated history, location, popularity, and many other factors to generate a particular set of filtered results for us.

The rather amazing thing about any of the algorithms incorporated into our contemporary communication is that these computational systems know much more about us than we know about them. They're often mysterious and non-transparent, as has been mentioned: a black box that governs our information landscape, persistently at work to shape information flows, determining what information we see and in what order we see it, and then nudging us towards certain actions by organizing our choices.

Algorithms do govern content automatically, but they do so because they have been designed that way. The capacity of algorithms to curate or sort information has been designed to sit behind the user interface of our popular search and social media applications, so we don't directly interact with the algorithm. Curation and filtering of information is sometimes something that we can see happening, but it's not entirely clear how it is happening. For example, the simplification includes things like swiping and tapping, and clicking icons in our mobile apps—highly simplified behaviour.

The extraordinary complexity of algorithms in automated curation is thus deeply hidden in the software and digital infrastructure necessary for networked communication, and this leads to a sort of distancing effect between us as human users and the complexity in the systems we are interacting with, such as Google Search, for example. It becomes difficult for us to connect our simple button choices or search queries to any wider effect. We don't necessarily think that our own individual actions are contributing to the ranking and sorting of other information searches or the popularity of a particular newsfeed post.

Social media companies tell us that reaction buttons like “Like” and “Don't Like”, or love or angry icons, are a way to give feedback to other users, stories, and posts, and to connect with the issues, ideas, and people we care about, but this effectively trains us to input information that feeds the algorithm so that it can generate its output, including ranking posts and shares based on these measures.

I was recently reminded of the powerful ways algorithmic curation happens. In the context of a group of Facebook users making a few original and offensive posts, the situation quickly escalated over a week, and hundreds of reactions or clicks on all those “like”, “angry”, or “haha” buttons continually moved up that cyber-bullying incident in people's newsfeeds. As Facebook itself notes on the relevancy score of a newsfeed algorithm, “we will use any Reaction similar to a Like to infer that you want to see more of that type of content”. These simple actions literally feed the algorithm and drive up the issue.

I also find Google's auto-complete algorithm even more troubling. While Google likes to make grand public assurances that their auto-complete algorithm—the drop-down of suggestions you see when you're searching—is completely objective and won't link personal names with offensive auto-completes, it still drives users to problematic content via its complex and comprehensive knowledge graph.

● (1635)

Google's knowledge graph combines search results in one page with images, site links, stories, and so on, but it still combines information that is problematic. For example, the Google auto-complete algorithm still points us to details of the late Ms. Rehtaeh Parsons' horrific case that were propagated by Internet trolls and continue to feature in Google's “searches related to” suggestions that appear at the bottom of the search page, pointing to images and other problematic content.

Recent changes to automated curation techniques point to our need for sustained efforts to build digital literacy skills, as discussed earlier, that steer young people into thinking more critically and being ethically minded in terms of what's going on. I would argue that we also need, then, a specific effort to educate young people about what algorithms are, not in their mathematical complexity, but generally how it is that they're operating with these simplified user actions that young people are so eager to participate in.

Visibility and publicity, and shares and various Snapchat scores are part of the new social accounting that young people value, and it's driven by an increasingly subtle yet complex infrastructure: an algorithmic milieu of communication and control that leaves very little in the hands of users.

Algorithmic sorting, ranking, and archiving is persistent and ceaseless. It churns away continuously as social media and search users navigate, click, view, search, post, share, retweet, @mention, hashtag, and react. As users, these actions and immediate results feel dynamic and vital. At its best, it affords us efficiencies in information retrieval and communication, and at its worst, it amplifies some of our most problematic and prejudicial expression, action, and representation online.

Thank you. I look forward to your questions.

● (1640)

The Chair: That was excellent.

Corinne, you have seven minutes.

Ms. Corinne Charette (Senior Assistant Deputy Minister, Spectrum, Information Technologies and Telecommunications, Department of Industry): Thank you very much, Chair. Thank you for inviting Innovation, Science and Economic Development to address the issue of big data analytics and its applications to algorithm-based content creation, to the detriment, in some cases, of young girls and women.

[*Translation*]

This is an important issue for me not only because of its impact on my work, but also because I am a woman engineer who was in Montreal during the events at the Polytechnique, which were devastating for me.

[*English*]

Following graduation as an electrical engineer, I was very fortunate to have many great roles in technology with a lot of leading organizations, including IBM, KPMG, and FINTRAC, our money-laundering detection agency. I was the government CIO, until my current post as the SADM of SITT. For 30 years, I've been working in technology, I've seen the adoption of many great technology trends, including the Internet and big data analytics.

[*Translation*]

Now, as senior assistant deputy minister, my job is to use key tools —policies, programs, regulations, and research, to advance Canada's digital economy for all Canadians.

[*English*]

Briefly, my sector is responsible for a wide range of programs, including the radio frequency spectrum, helping to maintain the security of our critical telecommunications infrastructure, and building trust and confidence in the digital economy. We safeguard the privacy of Canadians through two key pieces of legislation: the Personal Information Protection and Electronic Documents Act, or PIPEDA, Canada's private sector privacy legislation, and Canada's anti-spam legislation. In my capacity, I can affirm that the Government of Canada is committed to seizing the benefits of big data analytics through the discovery, interpretation, and communication of meaningful patterns in data, while protecting the privacy of Canadians.

[*Translation*]

Today, I would like to share with the committee two linked ideas about predictive analytics and algorithm based content curation.

[*English*]

The first relates to the personal stewardship of our digital information and the second is about the Government of Canada's commitment to building trust and confidence in the economy.

[*Translation*]

To begin, I would like to note that what citizens, business, and government do online generates a massive amount of data about our world and about us as individuals.

[*English*]

Every day, businesses and consumers generate trillions of gigabytes of data, structured and unstructured, in texts, videos, and images. Data is collected every time someone uses their mobile device, checks their GPS, makes a purchase electronically, and so on. This data can provide beneficial insights on developing new products and services, predicting preferences of individuals, and guiding individualized marketing.

This is a tremendous opportunity for Canadian innovation. According to International Data Corporation, the big data analytics market is expected to be worth more than $187 billion in 2019. The amount of data available to analyze will double very quickly and progressively; however, there are growing concerns about whether the benefits of big data analytics could be overshadowed by the accompanying pitfalls and risks.

Studies demonstrating biased results and decisions that impact whether people can access, for example, higher education or employment opportunities, are increasing. We do need to better understand how biases towards individuals are generated and how we can guard against this. This phenomenon may be explained in part by algorithms that are poorly designed—poorly from a user's perspective—or data that is poorly selected, incorrect, or not truly representative of a population.

[*Translation*]

It is easy to see that spotty data and mediocre algorithms could lead to poor predictive analysis which can be very detrimental to individuals.

[*English*]

I share my colleague's comment that step one in terms of risk mitigation involves better digital literacy for all Canadians as an increasingly important tool to ensure that we know what bread crumbs we are leaving behind online. It can give Canadians the knowledge and tools to understand how to use the Internet and technology effectively, critically, and responsibly.

Personal stewardship of our online information can help all Canadians, especially young women and girls, but it needs to also be supported by my second point, which is about the frameworks that preserve our privacy, and now I will talk about PIPEDA. Canada's federal sector privacy law, PIPEDA, sets out a flexible principles-based regulatory framework for the protection of individual privacy.

The principles set out in PIPEDA are technologically neutral and are based on the idea that individuals should have a degree of control over what information businesses collect about them and what they use it for, regardless of the circumstances.

● (1645)

Of course, some information, such as demographics, geographic location, etc., can be determinants to targeted advertising. This is the data that these algorithms use to produce these recommendations, but this can have significant implications for the privacy of individuals, especially given the lack of transparency of the privacy policy of many online sites, and the lack of awareness amongst young people—and also older Canadians—about the data that they are freely sharing.

[*Translation*]

We need to strike the right balance between privacy and the economic opportunities resulting from the collection of personal information.

[*English*]

In conclusion, Innovation, Science and Economic Development Canada

[*Translation*]

works with a number of other government departments to promote the use of big data analytics and other digital technologies by the government and the private sector.

[*English*]

We need to promote an increased understanding of both the opportunity and the risks of our digital world, of how our data can be used, and of the privacy obligations of prediction analytics users, so that the benefits can be enjoyed by all, especially young women and girls.

I want to thank you for making this issue a part of your important work.

[*Translation*]

Thank you.

The Chair: Thank you very much.

We will now begin the question and answer period.

You have the floor, Ms. Ludwig.

[*English*]

You have seven minutes.

Ms. Karen Ludwig (New Brunswick Southwest, Lib.): Thank you.

Thank you for your presentations.

I want to share a bit of my personal background. When I was working on my Ph.D. in education, one of the areas that I focused on was technology, and when I was teaching curriculum development to teachers what I found was that it was very hard at times to get the message across—this is for Dr. Robinson—about the use of technology in the classroom and its impact of that when we have so many teachers in the K-to-12 system whose original introduction to education never included technology—it just kind of lands in the classroom.

What recommendation would you give to our committee about how we could help teachers, for example, to understand the implications of technology in the K-to-12 system? Also, where should that be incorporated? Should it be in a class of sex education or any course in general about the implications of algorithms?

Dr. Sandra Robinson: That's a great question. Thank you.

From my perspective, it's interesting because I catch up with youth in their first year of university. I think I get a sense, then, of that lack of digital literacy. They can Snapchat the heck out of the world, but they are struggling to understand how some of the pieces make that technology happen.

For teachers, perhaps, in those years prior to their students bursting out onto the world, I think we need to make a concerted effort. One recommendation is, organize appropriate training for teachers. I even think that teachers who feel that they have a facility with technology maybe should be marked out as champions within their schools or within their program to help be leaders and to encourage their colleagues. People fear technology and, in most studies, women more so than men. I think there has to be a very safe and encouraging environment to consider what kind of participation can happen.

I think we need to tackle not just the surface level of software applications, the how do we use this.... We've come full circle now with the Internet and now it's time to come back and say, “Hang on a second.” These simple user interfaces are masking a very complex ecosystem of software, and we can't escape trying to make an effort to understand and then to share that understanding with youth and among ourselves in pulling each other into the 21st century. I think it absolutely has to happen before students reach the upper level of their school training, in high school and whatnot. I think that in some ways it's never too early, given where you see young people and kids with cellphones.

Ms. Karen Ludwig: Thank you.

One of the things I want to follow up on, Dr. Robinson, is that you mentioned “champions”, as in the champions of technology. You're probably well aware of adoption theories for taking and incorporating technology. The early adopters are the ones who are always rewarded, particularly in education. Those who hang back and ask why or what's in it for them are seen as the resisters. Thank you for that.

Marshall McLuhan is well-known for coining “the medium is the message”. Listening to the presentations today, I was certainly reminded of that. For example, how a communication is conveyed can be more important than its content. In so much of what we've heard regarding cyber-bullying, we have anonymous abusers. I think the system of the algorithms themselves, from what I've learned today, sets up the system for, as Ms. Charette mentioned, targeted marketing. In many respects, it almost creates a environment for targeted victims. Would either one of you agree?

● (1650)

Ms. Corinne Charette: I would say that it's the amount of information freely available on the Internet that can be so easily aggregated by using the most basic of tools, such as search engines of any kind, that makes it easy for malicious actors to aggregate information on individuals or groups of individuals and to exploit that. It's not always algorithmically based.

Fundamentally, the algorithms will not work unless the data is available online. Really, the amount of data that most people deposit online wittingly—but mostly unwittingly—every single day is staggering. Over time, there are very many bread crumbs, and the search engines and other tools seek them out and aggregate them.

Ms. Karen Ludwig: If I may ask this, then, in looking at the bread crumbs, let's say there's a family of five. They all share the same computer. There is one log-in, or potentially no log-in, and maybe there are varying ages within a family. Is it quite likely, then, if a child goes off to do homework, that child unwittingly, to the parents or anyone else in the home, in terms of whatever interests there are in that home, could be confronted with images and messages that were never intended? Because of the bread crumbs that are left, could that be presented on the screen to whoever the user is of that shared unit?

Ms. Corinne Charette: There's no doubt that malicious actors find ways to infiltrate networks, home networks, and home computers and are going to find children, as well as adults, using those computers. Today, most households have more than one Internet-enabled device. They have at least one cellphone, probably some kind of tablet or computer for the kids to do their homework on—a lot of homework is online now—and they may have other Internet-enabled devices, including their home thermostat, their smart TV, and so on. All of those devices basically will stream out of that home at multiple times during the day, based on who's using what.

Ms. Karen Ludwig: On that, if I may add, here's my next question. As a committee that has been studying cyber-bullying and violence against young women and girls, I'm sure we've all been doing our own Google searches. Are we also leaving bread crumbs about violence against women and violence against girls, because of sites that possibly we may have been researching? Are we sending those bread crumbs out there as well to be added to our profiles?

Ms. Corinne Charette: I'll let Sandra do that one.

Dr. Sandra Robinson: I think that's something that I'm glad you've raised, because I—

The Chair: You're out of time. I'm sorry.

I'm going to have to go to Ms. Harder for seven minutes.

Ms. Rachael Harder: If you can do so in under a minute, you're welcome to finish that thought.

Dr. Sandra Robinson: Sure. Even when we search for things and we have the best of intentions, our searches are ramping up and feeding the algorithms. That's the answer to that, unfortunately.

Ms. Rachael Harder: Okay.

I've asked witnesses this question before, and I'll ask it again, because I'm interested in your thoughts. With regard to the use of algorithms, you're using the word “curate”. I actually really like that word; I think it's a good one. Or there's “steer traffic”, or whatever you want to say. I asked the other witnesses this question: could algorithms be used, then, in order to prevent access to pornography for those who are under age in the same way that algorithms are used on, say, Twitter? Algorithms are used to detect the word “slut” or the word “bitch”, etc., right? In the same way, could algorithms be used to positively steer young viewers?

Ms. Corinne Charette: The problem with that is that the Internet doesn't distinguish the age of the user online unless you have some form of search engine that requires your age to be disclosed. You might have parental filters on your home Internet connection that would prevent your youngster from getting to these sites from your home PC, just like there are filters in business and government that prevent users from going to any malicious sites and so on. Unfortunately, a search is generated by an anonymous user.

● (1655)

Ms. Rachael Harder: Okay.

Do you have any additional thoughts?

Dr. Sandra Robinson: I would agree. I think the other thing is that because there are so many mobile devices now, the difficulty with that sort of “one fix” is that the platforms work quite differently—the mobile technology platform versus the browser-based desktop or laptop—so it's difficult for that to be an extensive utility to prevent access.

Ms. Rachael Harder: Thank you.

Here's another question, then, for each of you to answer. I guess the basic question is, for algorithms, could legislation be brought in to build some parameters around how algorithms are used in an effort to essentially safeguard our young people in particular? I guess that would be my main interest. Could that be done?

Dr. Sandra Robinson: I don't think so, because we've seen regulation fail us, I think, in many ways when we try to contain or constrain information flows. They have a way of slipping out of the leaky boundaries we put around communication. I think the education angle is a much more powerful one.

Actually, legislating private organizations and private companies is much trickier to do, given that we are talking about many different kinds of organizations and the many different ways that they approach collating and curating information. Each algorithm is proprietary, so it's a trade secret. It's like Kentucky Fried Chicken: there's no way they're letting that out into the open.

Voices: Oh, oh!

Dr. Sandra Robinson: You might want to answer that, Corinne.

Ms. Corinne Charette: The other thing I would add to that is the notion of a free and open Internet. As for the notion of censoring, it's very difficult in terms of what can be done when you don't know who the user is. I'm not sure that's the best route to the end goal.

Ms. Rachael Harder: Thank you.

For both of you, once again, I'd like your thoughts on this question. How could we go about changing the public narrative, then, in order to make online users more aware of the impact that algorithms have on their day-to-day lives? How do we do that? How do we change that public narrative?

Ms. Corinne Charette: I think we do that by explaining what happens online to everyday and all Canadians, who are not necessarily all in the technology sector and therefore are less aware of what a search is really doing and of the ins and outs of different algorithms or capabilities. I think we need to educate everyday Canadians, teachers, parents, high school students, and even grade school students.

I would start in grade school with the consequences of what happens when you use a browser and what you do online, putting it in "everyday speak", such as asking them if they would leave an unbecoming photo of themselves in the middle of their classroom. Probably not, but then why they would post it online, you'd ask, because that's even bigger than a classroom, and so on.

As a society, I think we have a big awareness challenge to take an increasingly complex digital world and to help make those concepts "real life" to people, with real-life examples for children, teenagers, and adults alike about what they're doing and what the potential impacts are going forward. I don't think we spend at enough time doing that, not at all.

Dr. Sandra Robinson: I would agree. I think the U.S. election actually did escalate the conversation with the public narrative and discourse around the implications of online speech, hate speech, complex arguments around what people are saying and the implications of that, and then the impact of what people are thinking when they go online. The other side of that is that it tends to show also how hard it is to regulate that speech, right?

This is a very tricky one. I think education is key, but also key is making a link between the grossly oversimplified mechanism for uploading, posting content, and hitting the heart button, etc. There's a huge distance here between that and realizing what the material consequence of it is. Each one of those iterations drives popularity. Every single "like", "dislike", or smiley face pushes the ranking of a story. No matter how ugly that story is or how beautiful it is, up it goes.

● (1700)

The Chair: Thank you.

We're going to Ms. Malcolmson for seven minutes.

Ms. Sheila Malcolmson: Thank you, Chair.

Thank you to both witnesses.

Do you have any inspiring examples for us of where algorithms were used for good to identify, call out, or to actually prevent and address cyber-violence?

Dr. Sandra Robinson: Sadly, no. Nothing leaps out at me. I think there's potential there for an increase in natural language processing and machine learning to actually help us identify some of the speech that goes on in the public Twittersphere and on Facebook. That might over time improve our ability to catch these sorts of hateful events as they happen, but it's an extremely difficult thing to do. Algorithms are only as smart as as the people who train them and the data that's fed into them. It's pretty tricky.

Ms. Corinne Charette: Yes.

Ms. Sheila Malcolmson: Is that a "no" from you also?

Ms. Corinne Charette: I know that there are algorithms online that present a lot of good, but not necessarily in combatting cybercrime. For instance, in the field of medical technology, they are doing a lot of big data work and even finding specific individualized potential cures and so on. There are a lot of great applications of big data in a positive sense, but I can't think of any of them in a cybercrime concept.

Ms. Sheila Malcolmson: In looking for more good news in this conversation, have you seen some good examples of tools where the Internet community itself, using the human approach, identifies, calls out, and combats cyber-violence?

Dr. Sandra Robinson: Yes. There are some key examples of it that have occurred on Facebook. Groups of women did come together to organize a response to Facebook's refusal to remove gender-based violent images and content. It took quite some time. It was started by a group in the United Kingdom. They were very successful at pushing back on Facebook.

I think we've seen some of this happen more recently in the last four months on Twitter as well. I think an organized speaking-back campaign can make a big difference. It has to marshal not just individuals but often advocacy groups, which can get a little more attention because they can cross over many different kinds of social media platforms.

There are some stories out there, even those that don't directly intersect with cyber-harassment and cyber-bullying, such as women who spoke back against Instagram and Facebook for taking down their family photos of themselves nursing, but women who are nursing are being picked up by algorithms that are searching for nudity.

There are ways in which people do organize to push back on violence online and on violent speech and hate speech. They have been successful.

Ms. Sheila Malcolmson: How do you separate those actions from feeding the trolls, which we are not supposed to do?

Dr. Sandra Robinson: One of the things they did, for example, was that they didn't speak back strictly by using the same hashtags in campaigns. They went directly to Facebook executives. They went directly to the leadership of Twitter. They went directly to the leadership at Instagram. I think that's really what we need to do in terms of being organized: to speak back to those who have the power to change the way information is organized on their particular social media.

Ms. Sheila Malcolmson: So far as our final report goes, which is coming up pretty soon, do you have any recommendations that would make your heart sing if you read them in our final report? What would you be delighted to see as our recommendations to the federal government?

Dr. Sandra Robinson: Personally, I'd love to see a commitment to advancing this issue through education, right from primary school through to high school and well beyond. I use MediaSmarts materials in my courses from the first year of university through to the fourth year, and they're outstanding. They reach students. They speak to students. Students go after those materials themselves. I'm an educator, and I'm a firm believer that this would be appropriate.

● (1705)

Ms. Corinne Charette: I would certainly agree with Sandra. I think whatever the federal government can do to sponsor, champion, and encourage education with young girls and young boys, starting in grade school, in methods, a vocabulary and a vernacular that they can relate to, is so important, including for parents. I think the opaqueness of this issue is that it's so easy to click online and not understand the ramifications of any click, selection, or search that you put in.

People need to know this as much as they need to know how to drive or do banking. I think we have to put a big emphasis on this in our society. As computing capabilities continue to grow and as data continues to grow, the awareness of the users of this technology needs to keep pace with it. We need to become much more sophisticated so that we can be critical about what we do online, how we do it, what tools we use, and what tools we don't use.

Ms. Sheila Malcolmson: Thanks, Chair.

The Chair: We're going now to Mr. Fraser for seven minutes.

[*Translation*]

Mr. Sean Fraser (Central Nova, Lib.): Thank you, Madam Chair.

Before I start with the questions, I would like to thank the witnesses for being here today.

I have a question for Ms. Charette.

[*English*]

You spoke a bit about the innovative practices and all the great applications that exist, which seemingly nothing to do with violence against women specifically. I'm very excited about the further development in Canada of a personalized user-based web experience. I think there are a lot of great applications, both from the consumer's perspective and from a business perspective.

Of course, I share the same concern the committee has, which is that the Internet seems to have gotten very good at showing us the good things that we want, but also at showing us bad things, whether we want to see them or not.

As we're seeking to make recommendations to the government, I've heard you loud and clear: education has to be at the fore when we're making these recommendations. Are there certain areas that you'd like to see us avoid to make sure we don't stifle the positive innovative practices that are developing in the private sector?

Ms. Corinne Charette: I think that in any dialogue it's important to cover all sides of the story, so we wouldn't want to fall into strictly negative dialogue on big data, big data analytics, and the negative potential. I think we need to keep stressing all the good things about big data, including predicting when there are weather events happening, and also in terms of how, based on social media, communities can come together quickly in an emergency to solve a problem and to help find missing people and so on.

I think we need a balanced dialogue that highlights the great examples of the positive use of big data analytics and the great innovation potential for our economy from firms that are developing these tools, collecting big data, and making open data available. The federal government is a big sponsor of open data, through the open data portal. We have many open datasets, a lot of which are geospatial, which help Canadians, as well as many people internationally, understand the different constructs of the country,

Mr. Sean Fraser: That's excellent.

Ms. Robinson, do you have anything to add? I want to make sure that we cover off the potential pitfalls here if we're going down this path.

Dr. Sandra Robinson: On the issue of violence and cyber-bullying online, it seems to me that it would be tough, I think, for any corporate citizen in Canada not to get onside there, so I don't think the kinds of ways in which we would address that are necessarily going to constrain the apparatus of data collection that private corporations require to do business in this day and age. By no means would I want to see any kind of constraint in that regard. I think it's not realistic in this contemporary era.

But I think the issue is well beyond just the “where” and the “how” of data being collected, because we're not actually going to roll that back. It is about the consequences in that networked, social, communicative environment that people, and young people in particular, are so engaged in. I think there's enough difference there between the objectives of a corporation versus that level of communication such that we should be able to find a path on which we can agree., I would think.

● (1710)

Mr. Sean Fraser: On these social spaces that you're talking about, you spent a bit of time in your testimony in talking about the massive distances between the simplicity of the user interface and the algorithm on the back end. If we're able to pull the mask off the user interface and everything is laid bare, I'm at a complete loss as to what we would do with that new information. If we see the mechanics behind the machine, what as a federal government could we actually do with that information to help ensure that the social space is safe for young women and girls?

Dr. Sandra Robinson: I think that opening up that mysterious black box that runs those things can only be done through reverse engineering, by studying how they work, because those companies that have those proprietary platforms of algorithms are not going to divulge them, for trade secret reasons and for competition reasons. I wouldn't think that we could actually ask them to do that, but what we can do is try to publicize what is that step between clicking that heart button and what that action is actually doing. What does that do in principle on Twitter to drive up particular kinds of stories? What kinds of choices do people make, and how do we connect their choice to click on one of those reaction buttons on Facebook and the material consequence of the action?

To me, that's where this sort of cloud gets lifted: in exposing what feels like a very intangible click to the fact that it has a material consequence, particularly in cyber-bullying cases, where there's a real effect for young women who are targeted in those practices online.

Mr. Sean Fraser: Is it really about ensuring that the product's end user is making an informed decision when they click that “like” button or whatever it might be?

Dr. Sandra Robinson: In my view, yes.

Mr. Sean Fraser: Okay. Thanks very much.

We heard from the previous panel about the concept of brigading and in particular about the ability of groups with certain nefarious agendas that are able to commandeer a platform and take it over. To your knowledge, is there any sort of technical product or strategy that exists to prevent that kind of taking over of an otherwise useful social platform?

Dr. Sandra Robinson: Not really. There's a huge effort going on right now at Google and elsewhere to try to pre-empt that kind of activity. Google has actually done a pretty good job of tweaking their algorithm to stay one step ahead of what is essentially, instead of marketing-based search engine optimization, more malicious search-engine optimization by hackers and by groups who would choose to manipulate them in that way.

I actually think that this will be solvable over time as machine learning becomes more powerful and as they become better at tracking those kinds of nefarious methods. That one, I think, is more likely to be solved—not soon enough, perhaps—to block this kind of influence.

Mr. Sean Fraser: Let me jump in there. I only have about 30 seconds left. A while ago, we were talking about education. Should we be targeting grade school right through to the senior population and the corporate community as well? Are there other target groups we might be missing for an education initiative?

Dr. Sandra Robinson: Digital literacy education can never hurt. It's particularly crucial among young people.

Mr. Sean Fraser: Thank you very much. That's very interesting.

The Chair: We'll go to Ms. Vecchio for five minutes.

Mrs. Karen Vecchio: Thank you.

You mentioned hackers, or the discussion of hackers came up here. I'm very simple when it comes to using the computer; I do not understand the scientific background of it. We were talking about the hackers. How is it that they can also work with these bread crumbs that we were talking about when you were talking about the algorithms? What is it that they can do that is going to send us to some of these malicious sites?

Ms. Corinne Charette: Hackers have a variety of tools. The most basic one used is the concept of phishing. Everybody here probably routinely gets a number of phishing emails. Now they're starting to do text-message phishing that says “click on this link”, and as soon as you click, that's it: your device has just been downloaded with an unwelcome visitor that resides there, often gathering information about your daily online activities without your knowing, until such time as that information is of value, for whatever purpose.

Phishing and the ability to infect devices with malware is very prevalent, and hackers are really key at wanting to do that to capture personal information as well as information such as your bank account number, your password, and so on, so that they can log in as you from another place and time and transfer all of your funds to some other destination.

Phishing is number one, but you might have heard a couple of weeks ago about quite a successful denial-of-service attack generated by the Internet of things in things as basic as your home thermostat. All of these devices that connect to the Internet come with often pre-set standard passwords. The user may not always know what the standard password is or even know how to change the standard password, but hackers will know how to do it, so they'll be able to penetrate your home networks and lurk until such time as they need to do what they'd like to do.

Unfortunately, given the complexity of technology in our homes and businesses.... At least in our businesses we have LAN administrators and technologists who are looking out for us, but at home we're all obliged to become at least basic technology defenders in some way.

● (1715)

Mrs. Karen Vecchio: I appreciate that. Thank you very much.

Sheila asked earlier whether there was anything you saw that was good. Are there examples of responsible leadership by service providers or of any platforms being used? You mentioned Google. What is it that they're doing specifically that is going to help us when it comes to our children and the cyber-bullying and the various things that we see going on today?

Dr. Sandra Robinson: Google has the good and the bad at times, right, in terms of the leadership they show. There have been some things this year that have pushed companies like Twitter to be more proactive in combatting hate speech.

When you think about this, I think we're in a funny situation here, because these are American companies whose servers reside in the U.S. and have the first amendment governing things like aggressive hate propaganda and hate speech. They don't have the same visceral reaction we do. They simply don't. Even at the corporate level, I think, it is hard to galvanize action, but we do see them taking action when there's a lot of publicity. It may take an extreme event to push that over the top, such as the targeting of the actor Leslie Jones with racist, sexist, and misogynist hate speech on Twitter.

I think some of those companies are getting the message. In particular, a company like Twitter has to get its act in order because, for all their billions, they are actually struggling to broaden their base and their platform. They are motivated. It is capitalism. It is a private sector company, so I think we can vote with our thumbs. We can push them—

Mrs. Karen Vecchio: That's perfect.

The one thing we've talked about also is the legislation and what it is that our government can do. A lot of times when we talk about the government, we talk about legislation and regulations, but in cases like this, I just don't see legislation working, because this is so much greater than just what the government can do. What are some of the techniques? Or what can we do as Canadian citizens?

The Chair: That's your time.

I'll go to Ms. Damoff for five minutes.

Ms. Pam Damoff: Thanks to both of you for your testimony and for shedding light on this.

One of the things that has come up, from both our previous witnesses and from you, is that the best thing we can do is digital literacy and education. Ms. Malcolmson was asking earlier about funding. Status of Women has very little funding to fund the programs that it has, and it seems to me that it would fall within more departments than just Status of Women.

I have a two-part question. One, where is the funding coming from? If this is a multi-billion dollar industry with big data, should it be falling only on government to be funding this education, or should it also be falling to the industry to provide some funding to groups like MediaSmarts for them to provide their programming?

Ms. Corinne Charette: Funding is always a big issue: who should be the funder? Certainly, education is a provincial jurisdiction, but the federal government can be a champion in helping to establish the interest level in this kind of debate and in establishing the need to raise digital literacy. Public Safety Canada works hard on keeping “cyber safe”. There's a variety of different programs across government departments federally that contribute to that and to consumer protection and so on.

Fundamentally, if we are going to get things into the curriculum, it's certainly a private sector issue, but it's also for the non-profit sector. I think corporations can do their part. Right now, for instance, we're seeing a huge amount of interest in teaching women how to code and getting women into STEM fields, which I think is fantastic, because we are seeing a declining enrolment of women in STEM education. Some of these programs are going to young girls in grade school and then right through to high school and university.

I think the non-profit sector is able to do a whole lot with really limited resources. I also think that the provinces are obviously the keys in education curriculum, but the federal government can be a champion by pointing out the national nature of this need and making sure that all the players across Canada do their part in helping to make a dent in this issue.

● (1720)

Ms. Pam Damoff: One of our witnesses talked about the need for more women to be involved in the coding and creating the algorithms. Obviously, we can't regulate that, but if there were more women developing the algorithms, I'm wondering if that would help in terms of harassment, and in particular harassment on social media.

Dr. Sandra Robinson: I think it does make a difference. I came from the private sector in the software industry back into academia, and I think it absolutely makes a difference. We can't be what we can't see. If we go into those software development places and participate fully in designing and developing software, why wouldn't it make a difference? Parity makes a difference everywhere in culture when we have women contributing in all kinds of ways. I think it brings a level of consciousness to the table that is based particularly on women's experience.

If you follow things like Hollaback!, you know that young women are being harassed in considerable numbers. These things are being tracked, thanks as well to big data and mapping and things. If you think about that, I think it will make a difference to bring women into that experience at that early stage of developing, just in the sense of consciousness.

If you Google “CEO” in the images on the Google search function, what do you get back? You get a sea of white male faces, and we know that our world is more diverse than that. That's the training data that was used to train that algorithm: Google Images.

Ms. Pam Damoff: I only have about 30 seconds. In terms of specific actions that you could see the federal government taking, if there were one, what would it be?

Dr. Sandra Robinson: I would invite the Microsofts, Googles, and Twitters of the world to help us combat cyber-bullying, because they have the power and the funds to do so.

Ms. Pam Damoff: Do you have anything to add, Ms. Charette?

Ms. Corinne Charette: I would agree. I think the corporations we deal with are doing their best to be good corporate citizens, and they work hard to meet the government whenever they can on issues of importance. I think that perhaps this is something that we need to leverage more and put an emphasis on in going forward.

Ms. Pam Damoff: Thank you.

The Chair: That was excellent.

I'm going to do something rare. I'm going to ask some questions.

I haven't given up yet on being able to regulate something. I was listening to our previous panel, and people were testifying that in Spain and the EU they have local regulations, and even global platforms like Google have to conform to them.

I would love to have some kind of regulation that wouldn't allow these platforms to upload content that hypersexualizes women or that is reflective of many of the rape culture items we've heard about, such as slut-shaming and catcalling and different things like that, and that there would actually be some requirement for diversity on the teams that are designing them.

I'd like to hear both of you talk about the regulation and the design.

Ms. Corinne Charette: In the EU, certainly, the regulation on the right to privacy is very, very strong. For instance, in the EU they have the right to be forgotten, or they're proposing the right to be forgotten. In Canada, our privacy legislation is strong, certainly, but it does not go to that extent, although in the EU it's still a proposal.

Regulation is a two-edged sword. The reality is that regulation is very hard to craft in an equitable way and comes at a cost to both government and business. To work out that fine balance is not easy. Also, compliance requires an effort. Given that no effort is free or without cost, I think that with the capacity we have and the resources available to us, there are probably more fruitful investments that would improve literacy, more than regulation would keep back. It's a constant balance.

● (1725)

Dr. Sandra Robinson: My view is similar. I don't know that it's legislation, but regulation is certainly possible by code. If we're working with big corporate providers that have the power to write the code and control the code, as they currently do, they already can filter code as it's uploaded. It's entirely possible.

Of course, in the U.S. it's called “censorship” and you can't have that, as it's against free speech, so there are issues there. Let's think about Canada. We could easily have a filtration system that parses or reads the code—image, text, video—as it's uploaded. It's coded based on a scale, so the scale has to be set up. It can interface with an age-based system much like they do in video gaming when you go on those sites. Also, it can be verified if there's access to the age of the individual who's doing the work, which means disclosing something other than having the person put in their fake age. You would actually have access to some kind of metadata that confirms the age of the user. These are all possible to do currently.

The Chair: I still have time, so my next question has to do with the binary nature of some of the algorithms that are being used. When we talk about gender and we see that we're getting away from binary definitions of gender, what do you think we could be or should be doing in that area to prevent discrimination against gender?

Dr. Sandra Robinson: I'm thinking here of a previous witness, my colleague Dr. Rena Bivens, who probably addressed some of that. Again, there are all kinds of things that are possible through the code. Even a company as established and wealthy as Facebook chose to, at the surface—where the user interface exists for user interaction with Facebook—give people a range of 50 choices to self-identify, based around gender identification, or non-gender identification if they chose that, but the way in which the back-end system works at Facebook is that they still read data as binary.

What they're doing is monitoring people's actions as they leave the site. Their algorithms look for patterns that suggest this person is either male or female, that they're queer or heterosexual. There are so many places in which that can be determined and inferred from other actions. It's very tricky to actually build in a system that is truly open and truly encourages that kind of self-identity. People can still be discriminated against if other people begin to make the same inferences.

The Chair: That's our time for today.

Thank you very much to both of our witnesses for doing an excellent job.

Thank you to the committee for very great questions today. I really enjoyed the session.

I'll see you on Wednesday. The meeting is adjourned.