OFFICE OF THE
INFORMATION &
PRIVACY COMMISSIONER
FOR BRITISH COLUMBIA

CHECK AGAINST DELIVERY

SPEECH TO THE
FOCUSED EDUCATION RESOURCES IT4K12 CONFERENCE 2024
November 5, 2024

Michael Harvey
Information and Privacy Commissioner for British Columbia

Hello, my name is Michael Harvey, and I'm the Information and Privacy Commissioner for British Columbia.

Before I begin my remarks today, I would like to respectfully acknowledge that I present to you from the traditional territories of the Lekwungen people, of the Songhees and Esquimalt First Nations.

As an Officer of the British Columbia Legislature, I also acknowledge that I am privileged to work with people across many traditional Indigenous territories, covering all regions of our province.

I'm grateful to be here – to work and live on this beautiful land.

Thank you to the organizers of this conference for inviting me to speak with you about a very important topic in Educational Technology: artificial intelligence, or AI. This transformative technology has become increasingly popular in many different sectors, including, of course, education.

Educational Technology – also know as EdTech – is, of course, hardly new to educators, school districts, or students. It has been a part of the educational process for decades – from the overhead projector in 1930 to the handheld calculator in 1972.

And, of course, the pandemic accelerated EdTech into warp speed in 2020, when classrooms were forced online and educators scrambled to find remote-learning solutions. Suddenly, many previous solutions for both educators and students were now available with "enhancements," including artificial intelligence.

Which brings me to why I'm here today – to talk to you about a dimension of AI that is often overlooked – its potential impact on our personal privacy.

Part of my office's mandate is to comment on the privacy implications of new technologies.

And we've been watching the advancement of AI closely over the past decade.

Our wider mandate is centered around protecting and promoting the privacy and access rights of people living in British Columbia as they are set out in two laws:

- the *Freedom of Information and Protection of Privacy Act* (FIPPA), which applies to more than 2,900 public bodies in BC, including public schools and post-secondary institutions;
- and the *Personal Information Protection Act* (PIPA), which governs the private sector and applies to private schools.

We mediate and investigate access and privacy issues, issue legally binding orders, and conduct in-depth investigations and audits into compliance with our laws.

I need to be clear before we start that the OIPC is not anti-technology in general or AI in particular. We like to remind people that we are not the "Office of No." What we aim to do is to recognize how best to achieve the public purpose in a way that enhances the privacy of the people of British Columbia. We recognize that AI offers exciting possibilities to education, just like the rest of society. Much of what I will say will sound like cautions, but this should not be mistaken for a fundamental position against AI.

It might be useful to start by clarifying what I mean by AI. There are many examples of EdTech applications, but it is not always entirely clear what the role of AI is in them. We have come to use and accept this term as synonymous with "modern and sophisticated," but we need to understand exactly what it is and how it works to tease out the implications of using it.

While there are many ways to describe AI, its core feature is that the systems are designed to do what we might describe as "learn" in the sense that their programs can change themselves over time depending on their input. This is a key distinction between automatic processing – which would involve a computer processing data and producing outputs based on a human's instructions – and what we call "algorithmic" processing – where the computer is instructed on not just how to produce outputs but how to change what outputs it produces based on previous outputs. This means that, independent from human involvement, an output provided today might be different than an output it provides tomorrow.

Here is a simple example familiar to most people: Duolingo is an app that claims to use AI to "tailor learning so the lessons are at the level and pace of the user." Language training apps are not new, but a standard one would have human-designed lessons that every user would complete. Duolingo claims to, independent of human control, change what tomorrow's lesson looks like depending on how the student did on today's lesson.

This difference seems to cut to the heart of who is the teacher. In the standard model, the human programmer is the teacher and the app is merely an instructional tool. In the latter case, it seems to me that making a decision about what a student should be taught is a teaching decision.

I first started to become aware of the application of AI use cases in the public sector about 10 years ago – at the time I was a health executive – but at that time there was a wariness to using the term AI. People were more likely to talk about "machine learning" or "big data analysis" or even "neural networks." They generally avoided using the term AI because at the time it was used to describe a computer that was sentient and actually thought like a human. And that was basically science fiction. So what's changed?

Over the past two years, one type of AI that has hit the headlines and entered our every day vernacular is Generative AI, which is a type of machine learning where systems are trained on massive information sets – often including personal information – to generate content such as text, computer code, images, video, or audio in response to a user prompt. Pre-GenAI use cases produced simple outputs – like a number or a decision. GenAI produces outputs that seem human-like.

Does that mean that it thinks like a human? Well, maybe how we think sometimes. The reason that GenAI needs these massive information sets to produce its human-like outputs is that it is basically faking it. When a human asks it a question it starts using immense computing power to generate responses and compares those to similar responses that it has learned from its dataset. It does this over and over again, at hyper speed, amending its response each time and testing the probability that it is accurate against the real thing. Eventually it will decide whether its output looks realistic enough and spit it out.

If you've ever used GenAI applications then you will have seen some of the implications: if you ask one to produce a bibliography, it will often produce a mix of actual citations that it has scraped off the internet as well as some other citations that are entirely fictional but look very much like citations. If you ask a GenAI image app to produce, let's say, an image of a person wearing a hockey jersey, you might get a picture that looks very much like a human drew it but often the lettering on the jersey is all weird – GenAI models certainly know how to deal with letters if they are dealing with words, but if they are dealing with pictures they don't know that the letters are letters rather designs on a sweater. It will spit out designs that kind of look like letters and are placed generally where letters will be. Does it know what they mean? Certainly not.

This is probably not what you associate with the idea of "the spark of human ingenuity." People have described these models as basically parrots that are really good at probability.

But let's be honest – I think this is actually how a lot of humans learn a lot of things a lot of the time. It reminds me of Daniel Kahnemann's very popular behavioural psychology book "Thinking Fast and Slow" in which he described two types of thinking – System 1 thinking, what we do most of the time, and System 2 thinking, which is methodical, step-by step and leads to insight and understanding.

In System 1 thinking, we interpret information and base our decisions on what we have seen before and patterns that we recognize without actually thinking about them methodically or even necessarily understanding what we are seeing. We may have thought similar things through and understood them at some previous time… or maybe not. But we need to act now and so we have to think fast.

Kahnemann says "when we think of ourselves we identify with System 2 – the conscious reasoning self that has beliefs, makes choices and decides what to think about and what to do." But the reality is we use System 1 most of the time.

GenAI is comparable to System 1 and so it can be very handy. It can use this basic form of thinking to produce outputs for us efficiently.

But it can be dangerous too. I'm sure that each of you have found yourselves in a conversation with someone who has been expressing an idea or conveying information to you and you can tell that they don't actually understand what they are saying. Depending on how good their memory and mimicry skills are, what they're saying might indeed be correct. And depending on how convincing they are, you might believe them. But they don't know if they're right. They don't actually understand what they're saying.

It's become a favourite internet exercise to ask certain well-known large language models how many times the letter r appears in strawberry. If you were to take a look, it will probably tell you that there are two r's in strawberry. If you challenge it, it will recognize its mistake, claim to have re-counted and found that indeed there are three r's. If you ask it what the problem is, it will tell you that it was an honest mistake, no big deal, anyone can mistake the number of r's in strawberry and that it does not represent a fundamental flaw in how it thinks or counts. Then ask it how many c's are in ketchup.

The reason I've gone on this diversion is to highlight something really important about GenAI and its application in education and that is to highlight that we should never mistake its appealing outputs for actual understanding.

So why is this something that as privacy commissioner I have a mandate to worry about and talk about. The answer is that these AI systems need a vast amount of information to be trained, and much of it is personal information. They do things with that personal information that we don't understand and crucially, that they don't understand. And they can produce outputs, often which include personal information, that are of dubious accuracy.

Privacy is often narrowly linked to security – keeping our personal information safe from access by unauthorized users. Of course, security is an essential component of privacy. But privacy today is about control. Our ability to control our information, in today's Information Society, is how we control our autonomy and dignity. The question of how AI systems collect, use and disclose the personal information of our children and youth is fundamentally a privacy question.

And, as you can imagine, children are at a particularly high risk of negative impact by AI technologies, including generative AI.

They are less able than adults to identify or challenge biased or inaccurate information.

Young people are also generally less able to understand and appreciate the long-term implications of data collection, use and disclosure.

This is why, as a vulnerable group, children need even greater privacy safeguards.

Unfortunately, privacy legislation lags behind these advanced technologies.

There are a few fundamental incompatibilities between the way that AI systems work and the foundations of our privacy laws.

British Columbia's privacy legislation, like elsewhere in Canada, is founded on principles that emerged from the Organization for Economic Cooperation and Development in 1980. There are 10 principles that have been recognized by the Canadian Standards Agency and are features of our law. I'll focus on a few that are particularly relevant here:

The first two are the principles of necessity and proportionality. Our laws recognize that personal information has to be collected, used, stored and disclosed for important purposes, but it should only be done if necessary. If there are viable options to achieve the purpose otherwise that are less privacy invasive, then these should be the ones chosen.

This is connected to a second principle of proportionality – that the privacy invasiveness should be proportional to the benefit that is to be achieved. AI systems are selected often on the promise that they are much more effective. Is the extent of their effectiveness enough to offset the increased privacy concerns of an AI system? This is very hard to tell.

Another principle that I'd like to draw attention to is minimum use. The use of personal information should be the minimum necessary to achieve the purpose. AI systems, however, are based on the opposite principle – maximal use. AI systems need red meat… the more the better. They not only need the information they collect to produce the direct output sought, but also to train their system and the more information used for the latter the better.

And this flows into the final two principles I would like to highlight – consent and consistent use. Generally speaking, information that is collected should be based upon the consent of that from whom it is collected. Now there are many circumstances in which consent is not appropriate – such as when it is involved in the provision of a public service for which there is no alternative. Health care and education are two great examples. There is often no sense seeking consent when the alternative is not to receive a critical public service. I could speak for an hour about the issues with consent, but for the moment let it suffice to say that consent should be obtained where appropriate and necessary.

Moreover, that information, once obtained, should be used for purposes consistent with that for which it was collected. Otherwise, if it is used for something else, what we might call a secondary purpose, then express consent should be obtained for that. As mentioned, AI systems inherently use information for two purposes – to produce the output and to train the systems. Again, a challenge for our privacy principles.

There have been positive movements in Canada when it comes to privacy reform and AI.

For example, in 2018, Quebec adopted a statement of principles for the responsible use of artificial intelligence by public bodies.

In March of this year, Ontario became the first province to pass a law requiring employers to disclose the use of AI in the hiring process.

And there are movements in this space on the federal side. In 2022, Canada proposed the Artificial Intelligence and Data Act or AIDA to attempt to regulate and promote the responsible use of artificial intelligence within the private sector. It remains in the Committee stage of the federal legislative process.

Meanwhile, the Government of Canada has created a voluntary code of conduct identifying measures that organizations are encouraged to adopt when they are developing generative AI systems.

Here in British Columbia, the Ministry of Education and Child Care has introduced a framework to guide decision-making for the integration of AI tools in BC's' K-12 school. This framework includes seven distinct categories and I encourage you to find it on its website and review it. Many of these overlap with our own considerations.

And that is where I will go now: I think there are things that we can do to enhance safety of AI systems in the education sector, so I'll finish my remarks with some advice:

In late 2023, my office, along with other Canadian Information and Privacy Commissioners, issued a resolution detailing proposed principles for responsible, trustworthy and privacy-protective generative AI technologies.

I recommend that you look at this statement yourself, which you can find on our website under "collaboration." But I'll briefly walk you through some of the principles now, starting with legal consent and authority:

- When building a generative AI tool, ask developers if they know and document their legal authority for collection, use, disclosure and deletion of personal information.
- Ensure that where consent is the legal authority, it is valid and meaningful. Consent should be as specific as possible, and deceptive design patterns should be avoided.
- Ensure that where personal information is sourced from third parties, the third parties have collected it lawfully and have authority to disclose it.
- Since AI is only as good as the data, an important question parents, teachers, and administrators should ask when choosing a product is where is the data from that was used to train the model? Are you sure it is accurate? Are you sure it is up to date?

Openness is a second principle I'd like to highlight.

If you use an AI system you should:

- Inform students and families what, how, when, and why personal information is collected, used or disclosed throughout any stage of the generative AI system's lifecycle.
- This includes stating the appropriate purposes for these collections, uses and disclosures. Ensure that system outputs that could have a significant impact on an individual or group are meaningfully identified as being created by a generative AI tool.
- Ensure that all information communicated is designed to be understandable by the intended

audience, and made readily available before, during and after use of the system.

In closing, I think we can all agree that we have come a long way since the handheld calculator.

But algorithmic systems are different than other EdTech products because, if improperly used, they may act as more than just an educational tool and may make actual teaching decisions themselves. Moreover, because of the sheer scale, speed and power at which they operate, they can risk misuse of the personal information of our children and youth, to whom we owe the highest duty of care.

So, while we wait for legislation, it’s important to put your own guardrails in place when adding AI platforms to your educational curriculum. Carefully assess apps and vendor PIAs and be sensitive to any aspects of AI that could harm students.

Children should be able to benefit from technology, safely and free from fear that they may be targeted, manipulated, or harmed.

That’s why we all need to be involved in this discussion: children and their parents, legislators, the education sector and the private organizations that market to young people.

Thank you for inviting me to speak to you today.