

**Transcript for Protecting Students from AI with Algorithmic Impact Assessments
ETUG Fall 2024 Workshop: Navigating Ethics and EdTech – November 1, 2024
Presenter: Ian Linkletter**

JACQUIE HARRISON:

Okay. Up next, we have Ian Linkletter. His presentation is called Proctoring Students from AI with Algorithmic Impact Assessments. Ian is an emerging technology and open education librarian at BCIT, over to Ian.

IAN LINKLETER:

Thank you. First of all, can everybody hear me? Someone give me a yes. Yep. Okay, great. So my presentation is actually called Protecting Students from AI with Algorithmic Impact Assessments. My name is Ian Linkletter, and I'm an emerging technology and open education librarian at BCIT Library Services. I'm really happy to be back at ETUG this year. I want to acknowledge that I live and work on the traditional ancestral and unceded territory of the Coast Salish nations of Squamish, Tsleil-Waututh, Musqueam, and their peoples. I'm very grateful to live and work on this land.

Today's presentation is broken up into four main parts. And before I share those with you, I want to point out that the slides with all these links are going to be shared in the ETUG Mattermost at chat.opened.ca. And I believe they'll also be posted on the website, maybe sometime next week. But all the links from this presentation are available right now at bit.ly/aialinks So let me hop into the chat and just pop that link in here. So I hope that as I go through these resources, you'll be making some bookmarks and maybe preparing to make the case for AIAs or algorithmic impact assessments at your own institutions. Part one is going to be lessons from the AI proctoring debacle. I will explain what that is. Part two is the McMaster University Privacy Complaint. Part three is the Commissioner's recommendation of algorithmic impact assessments in higher education, and Part four will be a hands on with Canada's AIA tool.

Before we get going, and I'm really going to suffer on time for this, but I want to give a full minute for people to open up this survey at bit.ly/aiasurvey2024 or scan the QR code. There are three questions. It does not ask for your name, where you work is not required. These are the three questions. And the reason that I'm doing this survey today is because my belief is that most of us haven't really gotten familiar with algorithmic impact assessments before. And my belief is that our institutions aren't doing these right now. And I'm going to be presenting on AIAs next year at the fall 2025 workshop about progress on AIAs in our province. So I would like to compare the survey results from this year, next year and see where we went. So just one last time, and I'll pop it into the chat bit.ly/aiasurvey2024. I really appreciate your participation on this, and we'll never, in any way, use it to identify you at all. There's no personal information being collected.

So as a librarian and previously as a learning technology specialist, I've always taken a "do no harm" approach to my work. We are responsible for safeguarding student privacy and protecting students from any harm that may result from faulty technology or especially algorithmic decision-making technology. As I move forward with my presentation, I want you to keep in mind and ask yourself, can an algorithm be racist? And if so, how are we protecting students from discriminatory AI? And how do we prevent the next Proctorio from causing harm? I'm going to be criticizing Proctorio, an AI proctoring company, for the next couple of minutes. And this is probably going to raise some questions about, hey, Ian, what's going on with the Proctorio lawsuit? Some of you may or may not know that I was sued by Proctorio in 2020 for criticizing them on social media by sharing public YouTube links or unlisted YouTube links. And that lawsuit continues to this day. It's been over four years. That's my only update about that.

I want to start by sharing this video. It's 1 minute and 33 seconds, and I think very valuable. So I'm going to stop sharing my screen and move over to my window that has it. And I will switch my Zoom settings to share sound and be optimized for video sharing. So here we go. Full screen this, turn on captions. Here's the video.

[VIDEO STARTS]

Back in February, I had to take a test. It was a lab quiz for my biology class. He took it remotely due to the pandemic. I woke up in my dorm room that day, sat down, and now was ready to take my test. The lab quiz required me to use software I'd never seen before. Prior to taking the actual quiz, I wanted to use the practice quiz. It did not go okay. The software we were using couldn't see me. And I had tried everything. Shades down, lights on. I had my lights off and my shades up. Half my lights on, half my lights off, Shades down, shades up. I tried everything. None of this made sense to me. It was noon, so I knew there was a lot of sunlight in my room. Eventually, what worked was me standing in the middle of my room, my lights on, my window shade down, directly under the light. I didn't really want to take my quiz standing up. Under a light. I got an idea. My dad had gifted me and my sister some LED grade construction flashlights. So I had a thought. What if I put my lights on, my shade down, facing south and using the flashlight pointed at my face directly. That's what worked. It had taken me 45 minutes to get the app to see me properly. The quiz itself was only 30 minutes. I had talked to some other friends of mine. They'd never encountered this type of problem. I was lucky enough to eventually get it to work, but it's still unfair that apps like this leave students like me in the dark.

[VIDEO ENDS]

IAN:

Can everybody see my AI Proctoring Is Racially Biased slide? Great. Thanks. So as we saw in the video, this is one example of a student, Amaya Ross, using Proctorio. These stories were coming from all over the place during the pandemic. Many, many, many students of colour reported the exact same issue with Proctorio. A student actually at UBC wrote a letter to the editor at Ubyyssey, a student paper called the Ubyyssey, and the following day, UBC Senate banned Proctorio citing the racial discrimination concerns. That was back in 2021. So imagine all you

want to do is take a test, and the system your institution uses as a gateway to testing doesn't recognize you as a human being. This comes from the great Dr. Chris Gilliard, who has probably my favorite BlueSky account about academic surveillance and surveillance of all types.

Respondus, which makes the Respondus monitor AI proctoring software, has a patent for proctoring, which states that for "dark complexion persons," this is a quote, "a racial detection feature may be provided so that a downward adjustment can be made to the final risk tally." So not only is the AI proctoring software making the automated decision of whether someone can enter their exam or not enter their exam, but it's also making decisions about how suspicious that person is. And Respondus has acknowledged publicly in their patent that this is biased against people with dark skin. They can actually get a higher suspicion score as a result. And I gave a talk about this recently hosted by BCcampus. The EdTech Sandbox Series is fantastic and ongoing. It's called Beyond Surveillance: The Case against AI Proctoring and AI Detection, and you can learn a lot more about its harms at this link.

The other thing that AI proctoring software does is detect so-called abnormalities with people's behaviour. So this is a quote from the *Washington Post*, and I have a gif link on my links page. One system Proctorio uses gaze detection, face detection, and computer monitoring software to flag students for any abnormal head movement, mouse movement, eye wandering, computer window resizing, tab opening, scrolling, clicking, typing, and copies and pastes. A student can be flagged for finishing the test too quickly or too slowly, clicking too much or not enough. These are more examples of decisions being made by the technology.

Part two of my presentation is about a privacy complaint that McMaster University received. And before I get going into it, I want to point out two really great sources of information about this complaint. The first is the Information and Privacy Commissioner Ontario's Privacy Complaint Report, where they describe their investigation and the outcomes and recommendations of that investigation. And Dr. Teresa Scassa, who works at the University of Ottawa has an excellent article called "Investigation of AI-Enabled Remote Proctoring Under Public Sector Privacy Law, Leads to AI recommendations." So the tie in between what we were just talking about the AI Proctoring debacle and Proctorio, leads nicely into this because the student complained about Respondus monitor, which is pretty much identical software. So the text on this slide is going to be quite small, but I'm going to read all of it, so don't worry if you can't see it. Here's what happened. "McMaster University uses Respondus lockdown browser and monitor. A student complained to the Information and Privacy Commissioner of Ontario. They did not want to share their identity with McMaster, so the IPC initiated an investigation. And the vendor used the video surveillance being collected to "improve the services." The IPC found that those video recordings, even though they didn't have a student's name next to them, are personal information, and that vendor use of that personal information did not fall under the educational purpose for which it was collected. Students never consented and had no opportunity to opt out. And finally, the IPC recommended that McMaster protected students from the heightened risks of AI-enabled tools by conducting algorithmic impact assessments." So McMaster University is now conducting AIAs, having adapted the Canadian tool we'll be looking at shortly.

Canada's Algorithmic Assessment tool was developed in response to the Treasury Board's guidance or it's stronger than guidance. It's instruction on automated decision-making. The tool is a questionnaire. This is coming from the government website. "The tool is a questionnaire that determines the impact level of an automated decision system. It is composed of 51 risk and 34 mitigation questions. Assessment scores are based on many factors, including the system's design, algorithm, decision type, impact, and data. It helps assess and mitigate risks, ensure transparency, accountability, and fairness in automated decisions. And it's mandatory for federal departments and agencies using these tools." It's quite similar to a privacy impact assessment, which is something that all of our institutions do and are required to do when we collect personal information. And if you've never heard of or want to learn more about a privacy impact assessment and what they might look like, you can go to a website I created called the Canadian Privacy Library. It's at www.privacylibrary.ca and it has over 400 PIAs from 18 of 25 of B.C.'s public post-secondary institutions. And my question to you about Canada's Algorithmic Impact Assessment tool is what's stopping us from using this tool? If we can use it to prevent harm, why wouldn't we do it? This is a question that I think managers, senior leaders, directors need to be answering for the community. And if you report to one, I hope that you bring up this presentation and just ask, what's stopping us from preventing harm, using algorithmic impact assessments?

One other thing that I want to point out. So I just talked about how the Canadian Privacy Library is publishing PIAs. This is actually coming from the federal government's guidance. Departments are responsible for releasing the final results of the AIA in an accessible format and in both official languages on the open government portal. And this is a screenshot from the AIA tools homepage, which has Start your assessment on the left, and Search completed assessments on the right.

Are 21 algorithm impact assessments publicly available from organizations including Immigration, Refugees, and Citizenship, Canada and the Public Health Agency of Canada. Obviously, these two agencies work with a lot of very confidential information that impacts the livelihoods and futures of many people. And I just want to point out, so do we. Student data is a crucial bit of their future, and accusations of AI use or cheating during an AI proctored exam are some of the strongest accusations you can make and may lead to tragic outcomes.

So let's get hands on with the AIA tool with the time that we have left. I'm going to go ahead and open it up. So this is the AIA tool on open.canada.ca/ia-ela-is

So here we go. I've gone ahead and entered the project title, which is required, so I called it AI Decision-Maker, and this allows us to now navigate through the sections.

Section two is called Reasons for Automation. And this is a really important question to ask. Why are you using an automated decision-making system? Why is AI being used? This is similar to research that we do at our institutions, where prior to asking questions and implementing a

study, we ask what the purpose is. So are we using the systems to eliminate a backlog of work or cases or improve the quality of decisions or lower costs, or what?

Next up, we have the risk profile. So think about AI proctoring as we look at these questions. Is the project within an area of intense public scrutiny because of privacy concerns or frequent litigation? I can tell you. Yes, AI proctoring applies to that. Our clients in this line of business, students, are particularly vulnerable. I've selected yes. Are the stakes of the decisions very high? I've selected yes. Will this project have major impacts on staff, either in terms of their numbers or their roles? Maybe, yes. Will the use of the system create or exacerbate barriers for persons with disabilities? As we've talked about, the answer for AI proctoring is yes. It goes on to ask about the project authority and whether new authority would be required to implement it. It asks questions about the system. And let's look at what it asks about the algorithm. Just two questions. Will your algorithm have any of the following characteristics? The algorithm used will be a trade secret. The algorithm process will be difficult to interpret or to explain. If you're answering no on these questions, and I think right now a lot of the AI companies out there would say you have to answer no because they're not going to give you this information. It really begs the question, what are we even doing here? We're educational institutions, and we need to understand the technology that we're putting in front of students. So if you have a company that's making you sign an NDA about how the software works, or coming up with a black box solution, that students cannot understand and instructors cannot interpret, why? And the Algorithmic Impact Assessment tool flags that as high risk.

Moving on, there are a few other sections like about the decision, impact assessment, about the data, consultations. But I'm just going to leave this to you to explore. It's one of the links on my page.

As I wrap up my presentation, I want to do a quick demo of a tool that just came out this week. This is the emerging librarian emerging technology librarian side of me coming out. So yesterday, I got an invite to use ChatGPT's new search enabled feature called ChatGPT Search. I tried asking it, how do I get started doing an algorithmic impact assessment in B.C.? I'll paste this link into the chat. Well, let's just check it out. So I'm clicking on this link, and these are the results from this search-enabled feature, which you can access even if you don't pay for ChatGPT, which is currently required to use the feature. So it actually comes up. I read through all of this. I'm not just doing a query and being like, Look, it's awesome. I read through this. It makes sense to me. I didn't see any red flags or things that were inaccurate about it. So you can actually use this link, if you wish to get going at your institution. It breaks it down into steps, and it cites real sources from the government of Canada about where you can get started with those steps and learn more about them. So, you know, I think the ChatGPT search tool has some concerns. I asked it, who's a librarian at my library? And it picked one of the librarians and gave their full name, title, and email address, and phone number, which is on our website. But notably, I did not ask for contact information. And it begs the question of what would happen to somebody that has their information doctored, and then ChatGPT finds it and provides it to people without even asking. So this is the rest of my presentation.

We're going to have question time now for a Q&A or a chat. I'll be turning off my slide here and reposting the links in the chat. And I hope that you have some good questions for me. I'm happy to open it up for those now. Thank you, everybody, for watching my presentation.

JACQUIE:

Thanks, so I'll just ask you to raise your hand if you want me to call on you or put a question in the chat. Hi, Keith, you have your hand up?

KEITH:

Hi, Ian. Good and quick presentation there. And I know, so I have to confess, when you submitted your proposal, I was suddenly all over those links. So but I am wondering, do you think this is something that we're going to see B.C. post-secondary institutions adopting and perhaps even collecting and sharing the way we do PIAs today? You know, I am probably going to be releasing the AI companion in Zoom to the staff and students at Royal Roads in a little while, default off, but enabled to be turned on. And looking at doing the same process with it. Thinking through it, I originally thought there probably aren't very high risks. But going through it, you can see there is still the risk of privacy invasion, personal information being collected, etc. The purpose of it is innocuous compared to proctoring or, you know, those other kinds of things. Do you think this was something that you'd end up having to look at or that perhaps you should be looking at when we have tools that we are already using. They're getting bits of AI jammed into it?

IAN:

Yeah. Yeah, so there are a couple of questions there. The first one is, do I see this happening in our sector in B.C.? And I don't know. I'm going to check the survey results. I'm going to do this presentation in a year. Here's my plan. Over the next year, one of my primary goals is going to be for B.C. public post-secondaries to implement AIAs. I'm going to be writing to the Privacy Commissioner about this. I'm going to be talking to my MLA about this. Check my stats. I think I'm going to be able to succeed on this. And I think that sharing these publicly will help organizations have what the B.C. Digital Literacy Framework calls sector collaboration, where we're using the best knowledge that we have about how these systems work, which is hard knowledge to get and developing iteratively and collaboratively for those systems that we use in common. I feel quite strongly that we need to be doing something, and Canada has this tool already. So I'm hoping that it's something that our institutions will look at. And did I miss anything about your second question?

KEITH:

As a side thing, we're in discussions for the actual campus. We already have a teaching and learning side AI guidance document. But looking at the campus side, we already have Microsoft Office Suite. We have all these other tools that are starting to get little AI plugins or extensions or gee, it seems a lot more wordy about the way it's suggesting I finish my sentence nowadays. That looks like more than just a scan against 100 options.

IAN:

Yeah. Yeah, I'm seeing this with Microsoft Teams in particular. They have the new speaker coach feature, and even a speaker teacher assignment where you can have students do a presentation, and then it uses AI to track their pitch, their pace, their pronunciation, for real. It's doing this. And we were never assigned to that. We never did a PIA for that. We don't know necessarily where the data is going to do that analysis. And one thing that I found a little bit disappointing with the 2021 amendment to FIPA, which required institutions to do PIAs if they hadn't done them before is that people that hadn't ever done a PIA weren't required to go back and do them for tools that they had already been using. However, that amendment did say that if the data collection differs, if the implementation is changed, then a PIA is necessary. So I'm hopeful that language like that will be interpreted by institutions as meaning, yeah, you need to do an AIA on these tools, you need to do a PIA if the tools are different now. But it remains to be seen. I've seen about five Microsoft 365 PIAs now, and none of them include the more recent features. UBC did one for the AI transcription, I should say. But other than that, I haven't seen PIAs include the new features that are out.

JACQUIE:

Thank Ian. Maybe I'll read one of the questions from the chat and leave the rest for you to answer. Sure. Question I'll ask you is from Sally Goldberg Powell. She says, "Amazing. How would you approach instructors who are very excited to add GenAI in their courses? What are the key questions they should be asking?"

IAN:

Sure. So I mean, this is a little bit of a broader question from the AIAs, but I think it's something that's applicable to us today as instructors are learning and collaborating and teaching each other how to use GenAI in their courses. So I think that it's important for those of us that intersect with privacy and policy to keep an eye out for instructors that are super happy to use ChatGPT. And just emphasize to them that there's two legislative statutory requirements for us before we upload student data into ChatGPT or another system. So if you're copy pasting student work into an AI detector or asking ChatGPT to give you feedback on student work, keep these two things in mind and maybe caution people about it. The first is that the Canadian Copyright Act assigns copyright immediately, automatically to people when they write or create something. That includes students. So if you are uploading student work into a system like ChatGPT, which will then most likely reuse those prompts to calibrate its own system in the future, remember, improve the services. That is potentially a violation of the Canadian Copyright Act. And the other one is our provincial FIPA, the Freedom of Information and Protection of Privacy Act, which requires us to do PIAs before collecting personal information and definitely before sharing it with a third party. So personal information is in the metadata of documents that you may be uploading to ChatGPT 4.0. It's also possible to determine people's identities from stories that they may be telling in their assignments, and that may actually be a violation of FIPA as well. So there's so much enthusiasm right now. As an emerging technology librarian. My job is not to crush that enthusiasm, but to find ethical ways forward. And I think

that raising awareness of those two things is very important to making sure that we're not walking into basically a terrible lawsuit.

JACQUIE:

Thank you. So there's a couple other questions, but we need to switch over to our next presenter. Maybe I'll leave you to answer those questions.

IAN: Okay. I will. Thank you, everybody.

JACQUIE: Thanks, Ian.