Transcript for Beyond Surveillance – The Case Against AI Detection and AI Proctoring
BCcampus EdTech Sandbox Series session hosted on September 18, 2024
Facilitator: Ian Linkletter
Host: Britt Dzioba

BRITT DZIOBA:
Good morning, everyone. We're going to get started pretty quickly here because we do have a lot to cover today. But thank you for joining us for our first Ed Tech Sandbox Series of our second round of this offering. My name is Britt Dzioba and I'm an advisor at BCcampus on the Learning and Teaching Team. I wanted to get us started off in a good way by stating that BCcampus staff are situated on the unceded territories of the Tsleil-Waututh, Squamish and Musqueam Nations as well as the W̱SÁNEĆ and the Esquimalt and Songhees Nations of the Lekwungen People. As both individuals and as an organization, we continue to learn and build relationships as we actively respond to the Truth and Reconciliation Commission's Calls to Action. Today, I'm joining you from the beautifully sunny territory of the hən̓q̓əmin̓əm̓ and Sḵwx̱wú7mesh speaking Peoples, which is known as Burnaby. And today, we have a very timely and important session on AI detection and AI proctoring led by the wonderful Ian Linkletter. So I will pass it over to Ian, who will get us started on the session. But just to note, I will also have a few announcements and a survey link at the end of the session. So if possible, please just stick around for 1 minute once Ian has finished today. I'm handing it over to you now.

IAN LINKLETTER:
Okay, thank you, Britt. It's my pleasure to be here to kick off the new series with BCcampus. My presentation today is called Beyond Surveillance: The Case Against AI Detection and AI Proctoring, and this presentation is licensed Creative Commons, by attribution, noncommercial, no AI.

I'd also like to acknowledge that I live and work on the traditional ancestral and unceded territory of the Coast Salish Nations of Squamish, Tsleil-Waututh, and Musqueam. I live in Vancouver and work in Burnaby, and I'm very grateful to be on this land.

Today's session will kick off with a presentation, the Case Against AI Surveillance, which I anticipate will last no longer than an hour. I know some of you can't be here for the full 2 hours. We'll follow that up with some facilitated discussions, questions, reflections, anecdotes, and that'll last about 40 to 60 minutes. Then we'll have an algorithmic impact assessment hands on. Algorithmic impact assessments are the new privacy impact assessments, and we'll be getting oriented to what the government of Canada has released. Just a reminder with today's session, a lot of the information that I'm going to be presenting is extremely anti-surveillance. And this is not the session for you if you came to argue that surveillance is actually a good thing or market your product that is surveillance, but in a good way. This is just not what we're here to do today. We're going to talk and work beyond surveillance and against surveillance.

A little bit about myself. I'm a graduate of the Evergreen State College in Olympia, Washington. Evergreen has a long history of social justice and activism and is regularly targeted by Fox News. I graduated from the University of Western Ontario with my master's in Library and Information Science almost 20 years ago now. I spent 15 of those years as an educational technologist at Fanshawe College, BCIT, and UBC. I worked at UBC's Faculty of Education for 10 years, and I see at least one of my former co-workers here in the participants list. That makes me very happy. So I had a long career and accomplished a lot. In 2020, and this is very important to note because I'm going to be talking about this company quite a bit today, I was sued by an academic surveillance company called Proctorio. Essentially, what happened was I shared seven YouTube links on Twitter in order to criticize them and they claimed that that was copyright infringement. You can read more about this segment of my life, which is now four years and counting at Linkletter Open. Now, I'm an emerging technology and open education librarian at BCIT, and I recently created the Canadian Privacy Library, which you can see at privacylibrary.ca, and do no harm is my fundamental principle. As an educational technologist, as a librarian, I'm always focused on protecting students.

I'll be making the case against AI proctoring and AI detection, and as has been posted in the chat several times, all of fine links from today can be accessed at bit.ly/beyondsurveillancelinks I recommend that you open those up so that you can make some bookmarks as we go.

Let's talk about academic surveillance software. Academic surveillance software monitors student behaviour. Whether it's your screen or your body or your face, that's what it does. Educational technology, on the other hand, serves a pedagogical purpose. It supports learning, and I don't believe that surveillance technology has a place in education. It's sure not EdTech. Together, we can work against surveillance, and 2024 can be the year that we move beyond surveillance.

Surveillance demonstrates values. It demonstrates that as an instructor, you value power and control over teaching and learning. A classroom without mutual trust is not a good place to be. Students feel uneasy and can even feel creeped out by a professor that insists that surveillance is the way. It sends a message to students that they're not trusted, they're being watched, and they can be removed. Especially as we talk about AI, the fairness of removal is really a key sticking point that we'll talk about.

My presentation today is not about generative AI, but some of you may have heard me talk about the ethical issues with generative AI. I'll go ahead and say that those issues do apply to AI surveillance. The bias issues with the data that goes into these models are valid. The environmental impact of generative AI applies to all cloud technologies. Even though some of the AI proctoring tools actually do a lot of the algorithmic processing on the student's own device. The environmental impact is still significant, especially when training the models. Privacy concerns of generative AI. They're a little bit different with AI proctoring and detection, but they're valid that they exist, and the copyright concerns also exist because especially with

AI detection, the information that you feed into it, which is copyrighted by students, is then used to train the model.

My colleague, Charles Logan, a PhD candidate, has come up with a couple of really useful libraries that can support our work against surveillance. One is the Against School Surveillance Technologies Library, and one is the Against AI and its Environmental Harms Library. If you're working with people that don't understand what it means when generative AI and AI tools consume water to cool the servers or consume vast amounts of electricity that perpetuates climate change. You can refer to the Environmental Harms Library, which goes into much more depth than I am able to go into today.

Let's start by talking about AI proctoring. AI proctoring is a form of mandated spyware. It used to be very rarely used, and during the pandemic, it became commonly used during remote exams. Proctorio was bragging in the news that their business had multiplied by 10 times in March 2020. And that doesn't surprise me. I think all of us had some experience with surveillance being used when we all had to go online. It monitors bodies and behaviour. It uses AI to track a number of different things, which I think is my next slide or the slide after. It surveils what's coming through your webcam, what's coming into your microphone, and what you're doing in your web browser. Recently, in Ohio, room scans were found unconstitutional. They were found to be an unwarranted invasion of privacy. Privacy is something that we'll talk about during our conversation. It's so overarching with our discussions about AI proctoring and AI detection, that it could be a presentation all by itself. I strongly believe that both of these technologies are a form of unwarranted surveillance, an invasion of privacy. I think they go against the Charter of Canadian Rights and Freedoms, and we'll talk more about that. And AI proctoring exists on a spectrum from 100% AI with no humans involved, except for after the reports are generated to human-assisted AI where someone might be sitting in a call centre surveilling […] 16 or even more students at the same time. Human-assisted AI also usually gives the proctors control over the student's mouse and computer, which can be especially invasive. But today, I'm going to be talking about 100% AI tools.

Proctorio and other proctoring software generally works using a concept called abnormalities. This is actually what the software calls it. What it does is it takes each test taker's behaviour and compares it against everybody else. If you're an outlier, if you move your eyes more or less than other people, click your mouse more or less than other people, take longer on your exam or shorter on your exam than other people, you are then flagged for suspicion. In some cases, you actually receive a suspicion score, which is, of course, biased. The Washington Post wrote, "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, too slowly, clicking too much or not enough. These are just some of the behaviours that are tracked. Changes in audio level, whether you copy and paste, whether it detects another person in the room, which can, of course, be a person on a

poster behind you. Number of keystrokes, whether you lose your Wi Fi and have to switch to a different network, all of this stuff is tracked.

I'm noticing some issues in the chat, so I hope that those get worked out. We'll correct that link. AI makes proctoring especially harmful. I think many of us have had experiences of taking tests in a crowded room with somebody pacing around, and this is of course a form of surveillance, but it's not the same type of surveillance and it's not as harmful. AI proctoring discriminates against disabled students through biased algorithms. It discriminates against students of colour, as I will prove to you shortly. It's an unwarranted invasion of privacy, as we'll discuss, and it causes emotional harm. This image comes from a Twitter account called Procteario. It's a play on the word Proctorio. What it really emphasizes as you go to the Procteario Twitter account is that many students actually experience traumatizing experiences when they take exams through AI proctoring. It is a very common theme on this page, which retweets real student voices for students to say, This made me cry, and I couldn't even do anything about that because I knew I was being tracked.

I mentioned that algorithms of AI proctoring are discriminatory, and these are just some examples. If you're a blind or visually impaired student, the room scan process of actually moving your laptop in a 360 way is inaccessible. So are screen reader extensions because extensions are typically blocked by these softwares. It kicks you out if you step away from your computer. If you need a bathroom break or need to take some medication and it can't see your face on camera, you're kicked out. The eye and head movement tracking that it does targets neurodiverse students. There are a number of different reasons that your eyes may move differently from other people or your body may move differently from other people, and no one deserves to be called abnormal for normal behaviour. Having to stay on screen, looking at the screen on camera and not moving too much causes a lot of stress. It's unnatural. If you're taking care of another person or like I mentioned before, have a poster behind you, it can kick you out for detecting another face. The facial detection software that's used in AI proctoring is biased. In fact, I'll even say that it's racist, and we'll talk about whether algorithms can be racist later. Many people experience face not found errors where the facial detection software will actually not see your face and block you, flag you, or remove you from exams as a result. Dr. Chris Gilliard says, "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." You can find Chris on Bluesky. His handle is hyper visible.

I'm about to go through some examples of the racism that's inherent in these facial detection algorithms. Facial detection is, of course, a subset of facial recognition technology, which is known to be problematic to say the least. Here's a student who says, "Okay, exam support told me to sit directly in front of a lighting source, such as a lamp. I'm receiving the same issue preventing me from completing the NY UBE mock exam. Facial recognition technology is racist. Do you all think I have adequate lighting?" Here you can see what his face looks like on screen and the light directly in front of him. I can see him just fine. Can you?

Robin Pocornie is a student from the Netherlands who also experienced the face not found error, this time with Proctorio. She actually took her school to the Netherlands Institute of Human Rights arguing that this was a form of illegal racism. Let's take just a moment, about a minute and a half to learn from Robin about her experience. Okay. Okay. So, Kelsey, can you see my screen? Yes, I can, looks good. Great. So, here we go. It's about a minute and a half.

[VIDEO STARTS] I sit behind my desk for my first exam of my new studies. It's a pandemic, so we're using online exam software that surveils students to check if they are cheating or not. We call this proctoring. I sit behind my desk. I log in. My name, Robin Aisha Pocornie and my password. Welcome 123 exclamation. And it has a facial detection step. I sit in front of the camera and I wait for it to count and it counts one, two, three smile, takes a picture and it tells me, face not found, room too dark, which is interesting because it's broad daylight out and I'm in a well-lit room. I try again because we only get 15 minutes to log in. After these 15 minutes, we are barred from the exam. We have to log in on time and I sit there again and I sit one, two, three smile. Again, it tells me, face not found, room too dark. And in the darkest of times, we must create light. That's what I do. I grab a lamp and I just shine a bright in my face. I one, two, three, smile and bam. It works. I get into my exam. After the exam, I go on the Student Discord Channel. [VIDEO ENDS]

IAN:
Okay. Robin continues with her excellent TEDx talk to talk about how all of her white colleagues, her white classmates did not experience this issue. She was the only one. I just have one more video that I want to show you from Amaya Ross. This was produced by Mozilla.

[VIDEO STARTS] Back in February, I had to take a test. It was a lab quiz for my biology class. I 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 new, 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:

Some powerful videos. I always suggest listening to real human real student experiences if you want to know what's going on. If you hear even one example of an experience like this, it's happening to many more people. Okay. Okay. Kelsey, can you confirm that you see my screen again? Yes, we can see it. Great. Thank you.

And so this is not a rare experience. I've shown you three examples so far. Here is Amaya Ross's mother posting "Daughter 1 was taking an exam today being proctored by some type of software that apparently was not tested on dark skin. She had to open her window, turn on the lights, and then shine a flashlight over her head to be detectable." Someone writes. "This happens to me using Proctorio." Someone else writes this. "There's no reason I should have to collect all the light God has to offer just for Proctorio to pretend my face is still undetectable." Gavin Gordon, who's a student at UBC, had the same experience writing this in a letter to the Ubyssey on March 16, 2021. "During the pre-test tech check, the software was unable to recognize my face. Being fairly experienced with technology, I tried the usual troubleshooting procedures: using a solid background, making sure I wasn't backlit, increasing the lighting on my face, trying different angles, etc., yet it still wasn't recognizing my face. To ensure this was not a problem with my webcam. I had my Caucasian roommate try the tech check, which instantly recognized his face without any issues. I know that if it happened to me, it's likely happening to other students. UBC banned Proctorio the next day. It was on the agenda for senate to discuss and citing racial discrimination concerns, it was banned with the slim exception of accredited exams. And Femi Yemi-Ese also experienced this face not found error. Nora Kaplan Bricker writes for the New Yorker. "Like many test takers of colour, Yemi Ese, who is Black, has spent the past three semesters using software that reliably struggles to locate his face. Now, whenever he sits down to take an exam using Proctorio, he turns on every light in his bedroom and positions a ring light behind his computer so that it shines directly into his eyes. Despite these preparations, he says, 'I know that I'm going to have to try a couple times before the camera recognizes me. Adding sources of light seems to help, but it comes with consequences. I have a light beaming into my eyes for the entire exam,' he said. 'That's hard when you're actively trying not to look away, which could make it look like you're cheating.'

And so there's clearly a problem here. This issue is not isolated to Proctorio, but all facial detection software and facial recognition software struggles to detect faces of all backgrounds and colours. In this case, Lucy Satheesan, who is a student, discovered that Proctorio she examined the extension and found that it was using open source facial recognition software called Open CV. When she tested Open CV against another open source database of faces called FairFace, she found that Proctorio could detect Black faces less than half of the time. I recommend reading her blog to learn more about the methodology, which was later replicated by RTL News in the Netherlands. This obviously became huge news. "Students of colour are getting flagged to their teachers because testing software can't see them," writes The Verge, and "Proctorio is using racist algorithms to detect faces," writes Todd Feathers of Motherboard by Vice.

This problem is known within the industry. Looking at patents on Google patents, I found that Respondus, which also sells AI proctoring software, actually was granted a patent for systems and methods for assessing data collected by automated proctoring. In the patent, they actually add that part of the patent is an adjustment to final tally function, which, if it detects dark skin, will actually use a racial detection feature so that a downward adjustment can be made to the final risk tally. This highlights that in addition to not being able to access an exam, the inability to accurately detect a face can actually lead to increased so-called abnormalities and a higher suspicion score. In this way, I believe that this software perpetuates racist biases. Because if you well, not you, but for instructors that already have racial biases, the fact that dark skin students appear at the top of their suspicion score tally will perpetuate their beliefs.

For our group discussion, we'll be talking about this. Can an algorithm be racist? How many face not found errors are acceptable? What does it mean to say AI proctoring is racist? How are institutions protecting students from discriminatory AI? And what does accountability look like? What are we spending for this privilege?

I did a freedom of information request of Ohio State University, which recently switched away from Proctorio and found that since 2018, over $465,000 had been spent. University of Colorado Boulder also recently switched from Proctorio. From 2015 on, they spent $722,000 on Proctorio. The way that this was licensed was based on the number of full-time equivalent students that they had. $3.68 per user. In their report, justifying their decision to get rid of Proctorio's campus-wide license, they pointed out that only 5% of instructors actually used Proctorio. Do the math, 368 times 20 is what they were paying. This was a huge win for the over 1,400 students that had signed a petition asking CU Boulder to stop the use of Proctorio. Here's their notification that it was retired in August 2024, except for certain programs.

UBC was also a very famous user of Proctorio. That's where I worked. I was there for this. I opposed this. And in the fiscal year of 2021, which ended in March 2021. March 2020 to March 2021, almost $300,000 was spent. And in fiscal year 2022, $0 were spent thanks to the ban.

This is the message that was sent out to the community, saying "Last evening, the UBC Vancouver Senate voted in favour of a motion to direct UBC V faculties to stop using remote and invigilation tools that involve automated recording and algorithmic analysis of data." I think the point that I want to make showing all these examples of people deciding to stop spending so much money on remote proctoring is that you're in good company not using it. Your institution is in good company not using it. You're not taking some anti-surveillance stand to not use software that discriminates.

It doesn't actually work. This million-dollar software doesn't actually stop academic dishonesty. You can spend $722,000 on academic surveillance software that can be defeated by a two cent sticky note on your screen or a $2 10-foot HDMI cable that goes to a monitor you can see in the room behind you. Students are exchanging techniques for how to do this all over the place. You

just have to Google how to cheat and you'll find out how. Many, many videos on YouTube, TikTok, Tweets, everything. It's all over the place. It's easy to do.

A study out of the Netherlands found that on the efficacy of online proctoring using Proctorio, they had an experiment with 30 students, 6 of which were asked to cheat in various ways, while 5 were asked to behave nervously but take the test honestly. They found that with human proctoring, 1 out of 6 were caught and with Proctorio proctoring 0 were caught. The efficacy should be put very close to 0 and is best compared to a placebo, where it doesn't actually work, but if people believe that it does, then maybe it does.

As I mentioned earlier, listening to students is a vital part of our practice as faculty, as learning designers, as educational technologists. Please go to Twitter, read Procteario, read ProcterrorU, and read the thousands of Google reviews that are online from people forced to use the software against their will. Share these voices with people who might otherwise know what's going on because they contain evidence of harm.

A question that I get sometimes is, what about lockdown browsers? They're not watching students as they take an exam. They're a special software that students install that prevents them from using other windows and tabs and extensions while they take a test. And so in my years of experience supporting lockdown browsers, I can say that technical issues abound with them. When I was at Fanshawe College, I would get support requests saying, I'm locked out of my computer. I reset my computer and I can't get back in and I don't know what to do. At that time, it was actually creating a new user on your windows machine with no data and locking you into that user account. When I was at BCIT, part of my job involved answering phone calls from students. And I don't remember how many students called crying, saying that they couldn't access their test, saying that they couldn't access their computer. And it's stuck with me my entire life. I'll never be able to forget these students. And the big, this is a formative experience that turned me against surveillance for life. And speaking about how it's not even effective. It doesn't block you from using your phone or a second device. I'll also point out that when we got those phone calls saying, Help, I can't get into my computer, we would read the students a password that worked for every single respondent's lockdown browser on campus to unlock their computer. It was just impossible for us to troubleshoot being locked out. And this is happening at institutions all over the world, we had a password to help them get out. That password could easily be shared. This also has compatibility issues with accessibility software and extensions, and the potential for data breach is real anytime you're installing a proprietary software on your computer. That's AI proctoring.

Next, I want to talk about the other type of academic surveillance software today, which is plagiarism detection, specifically AI detection. This is a type of software that scans assignments for similarity with other student work that's in their database. Turneditin has been around about 20 years, and every time a student work is submitted, it's then added to their database, whether students agree to that or not. How Turnitin used to work is that it would scan the internet and look for matches with phrases and sentences and quotes that a student put into

their work. And then an instructor could actually review those matches, see where the wording was claimed to originally come from, and talk to the student. But the business model was in danger because of ChatGPT, which could generate unique text that didn't exist on the internet. And so Turnitin added AI to the product in April 2023 without warning anyone. Since then, over 200 million papers have been scanned.

How do AI detectors work? Well, they use natural language processing to evaluate what's called text perplexity. Perplexity means how surprising the word choice is in an essay. The more common your words or phrasing, the lower the perplexity. Because large language models like ChatGPT 3 and 3.5 are designed to emulate human writing, low perplexity text is the default. But you can instruct them prompts not to use low perplexity text and later models like GPT 4.0, which is the model that everyone uses, whether they're logged in or not, is increasing and less detectable. Turnitin's AI detector has been opted out of by many institutions, including UBC and including BCIT where I work. Their promise was a 1% rate of false positives. The truth is that false positives are higher in the real world, and they were forced to admit several months later that the sentence positive rate is actually 4%. The Washington Post, when they tested Turnitin software, found that in a test of 16 student human written writing examples, one of them was falsely flagged. That's a lot higher than 4%. Even the 4% rate is far too high because if you think about the 200 million exams that have been looked at, that's 800,000 false positives. If you think about 10% of those false positives, leading to student discipline that they didn't deserve or 1% of those false positives leading to a student dropping out or being kicked out of their education, even 0.1% of those 800,000 having a tragic outcome from a student who can't cope with the false accusation. It's just I can't emphasize enough that you can't do something like this to 200 million people and expect it to be okay.

If you ask, Turnitin or go to their website and ask, how does it work? This is their explanation, which I think is severely lacking. They say, "When a paper is submitted to Turnitin, the submission is broken down into segments of text that are roughly a few hundred words (about five to ten sentences). These segments are then overlapped with each other to capture the sentence in context." This part is especially weasley. […] I want you to think about what they're writing here. "The segments are run against our AI run against our AI detection model and we give each sentence a score 0–1 to determine whether it is written by a human or by AI. If our model determines that a sentence was generated by AI, it will get a 1, and if it is determined to not be written by AI, it will get a 0." This actually isn't an explanation at all. The algorithm is a total black box. It's extremely proprietary. They're just saying, if our model says, yes, it gets a 1, and if it says, no, it gets a 0 as if this is a satisfying explanation.

They even have this diagram showing, Hey, look, the stuff that's not written get 0s and the stuff that is written by AI gets 1s. That's how it works. It's totally unacceptable. We can actually take a look at the Turnitin interface. This is not a secret website or anything. This is part of their marketing materials.

Here I'm opening a new demo and the link was just posted in the chat. This is what Turnitin looks like. All this red text is stuff that was not cited or quoted according to Turnitin. It's transparent in the sense that as an instructor, you can go look at it and decide for yourself. But this part right here is new AI writing score, and it says 18% detected as AI. It highlights in blue what it believes was AI. It puts an asterisk here to say that actually, because this is less than 20%, it has a higher likelihood of false positives. But it doesn't indicate what that likelihood actually is. They never have admitted what the likelihood actually is for writing that falls beneath the 20% threshold. This is not enough. This is not evidence. This is not proof. This would not survive an academic appeal. This would not survive a lawsuit, and this will cause a lot of stress to students that are falsely accused.

In the markup, it was reported that writing by non-native English speakers is more frequently confused with AI, which actually makes sense because native English speakers have usually a higher vocabulary, and that means more perplexity. 5.1% in this study, of native English speakers were flagged as AI written, and 61.3% of human written texts by non-native English speakers was flagged as AI written and doing simple paraphrasing changed those likelihoods pretty dramatically. So this is a study in an article worth looking into. Because again, it's an example of bias being perpetuated by AI. Because what if your student speaks with an accent, this AI score may actually convince you that the student doesn't know English and is cheating.

Turnitin claims that they mitigate bias. They say, "Our model is trained on a representative sample of data spread over a period of time that includes both AI generated and authentic academic writing across geographies and subject areas. We took into account underrepresented groups." This is total BS. The truth is that their models are biased. They've been proven to be biased, and no amount of jargon or technical speak is going to change it. If they believed that Turnitin actually didn't have bias, they would open it up to researchers to do their own research. If Turnitin had a model that wasn't biased, it would be their number one selling point because this is something that is impossible to do. On decision making, they write that "We won't say much about the parts where it may be transitioning between human writing and AI writing. It's a fuzzy boundary and you don't want to do any harm by saying the wrong thing." I have a YouTube video that we won't open of their AI scientist actually saying this quote. This is important to point out because this language that "Turnitin doesn't make decisions. You do," is very key. As you consider whether AI detection is something that you can depend on, that the company will back you up for making an accusation that it told you to. That's actually not true. The company claims that it makes no decisions, and same with AI proctoring, it makes no decisions. Teachers do.

And students know that AI detectors don't work. This was a Reddit thread where a student was falsely accused of having a 62% AI score. They say it was 0. And some of the comments include, "Complain to the dean." "Your professor is using an unreliable tool to avoid fairly grading your work." "Run some of your professor's works through an AI detector." "Make sure you complain in writing and use terms like 'hostile environment' and 'creating an unnecessarily adversarial relationship between student and professor. Of course, I no longer have an expectation of my

work being evaluated fairly." This last student points out, truly, that Open AI shut down their own AI detection tool because of low accuracy. They're the ones that invented ChatGPT, and they can't come up with a detector that's reliable.

UBC hasn't enabled the AI detection feature and they have a really good LT Hub website talking about this decision. These are quotes from the website. UBC has not been able to review and validate the feature. Testing for accuracy is in early stages. Instructors cannot double-check the future results. Testing for potential bias is also in early stages. Results from the feature are not available to students. This is a one-sided form of surveillance, and ability of the feature to keep up with rapidly evolving AI is unknown.

Students can easily circumvent AI detectors if they have the resources, and this is a real equity consideration to consider. If you look up the word "AI humanizer," or "word spinner" or "rewriter," there are many services that charge money to give students access to their own AI detector to check, re-spin, re-humanize, rewrite their work, and then recheck until AI is not detected. Cutting edge models, which generally have to be paid for like Claude or ChatGPT plus are less likely to be detected. Students with money and digital literacy have an advantage. False positives are obviously a big concern, but false negatives are also a big concern. The only way that Turnitin is able to keep the rate of false positives down is to increase the rate of false negatives, which is when AI detected text is not detected, and that falls into these categories of ways that you can evade detection.

And so to wrap up my presentation, I just want to emphasize my own belief that AI detectors have no place in education. A 1% rate of false positives is unacceptable. False negatives exacerbate inequity. Student assignments are confidential. What they give to you cannot be uploaded to some third party service, especially one that your university or college or institution doesn't support. Students hold copyright on their work. You can't be copy pasting it into some website form that then uses it for their own algorithms and models. It's just unethical to do and opens you up to liability. As I mentioned, BCIT banned Turnitin's AI detector last year, makes me very happy. I was part of that decision, and LLM development, large language model development far exceeds Turnitin development. Billions and billions of dollars are going into making this text more perplex, have higher perplexity, and basically seems smarter than a human instead of like a human. Don't buy into the arms race. You don't want to be buying into this and putting students in harm's way.

If I have not convinced you, if you're still going to use AI detectors, if you think that they work really well or you're running student work through five AI detectors, thinking that that is better than one, please read what Dr. Sara Elaine Eaton has written about ethical principles for detecting AI. She talks about how to talk to students, how to make sure that it's in your course syllabus on day one, how to make sure it's okay with your department to begin with, and how to make sure that you're not violating any policy. Again, I feel like this is the best advice that I've ever seen for how to use them ethically. I don't believe you can use them ethically, neither does Sarah, but she wrote this, and it's very helpful.

Thank you very much. This concludes my presentation. We have time for some questions. You can email me at ian_Linkletter@bcit.ca. My handle on LinkedIn, BlueSky, Mastodon, and Twitter is Linkletter. And after a period of questions, we'll be talking about privacy, AI, integrity, and more. Thank you.