Transcript for Algorithmic Bias in Education: Risks, Realities, and Responsibilities
BCcampus FLO Workshop recorded on December 9, 2025
Presenter: Britt Dzioba
Host: Helena Prins

BRITT DZIOBA:
Welcome, everyone. This is kind of part two to a workshop that I did back in April [2025], which was a little bit more of the foundations of algorithmic literacy. So I will still go over a few of those concepts, but I won't be going as deep into some of the more nuanced kind of elements of algorithms as that was covered in part one of the workshop. So I want to start off today by acknowledging that I'm coming to you from the traditional and unceded homelands of the Halkomelem speaking peoples, which include the Musqueam, Tsawwassen, Kwantlen, Tsleil-Waututh, Qayqayt, Kwikwetlem, Katzie, and Snohomish. I've included a reference here to a SHRCC funded project that aims to bring Indigenous perspectives, knowledges, and epistemologies into the development and use of AI. Somebody had put a question in the registration form around whether AI can truly be considered ethical. At this point in time, I don't want to be cynical, but I would say no. As it stands, I would say most AI models are unethical, and we'll get into that more later in the sense that they don't bring in everybody's perspectives. They don't consult with all of the people whose data is being used, and often AI comes at a huge ecological and human cost. But this project, I think, is a wonderful example of how community-led initiatives can push back against the unethical nature of AI. In total, the project's interdisciplinary team of experts includes 37 co-investigators and collaborators who come from eight universities and 12 Indigenous community-based organizations in Canada, the US, and New Zealand. And most of the team members are Indigenous, and they are motivated to expand the definition of intelligence by collaborating with Indigenous communities to integrate their knowledge systems with the AI research and development ecosystem.

So our agenda today, I'll quickly introduce myself. We're going to do a review about what are algorithms? We're going to talk about algorithm literacy. There's a short assumptions activity, so depending on how many people here, we may probably do breakout rooms just in pairs. We'll talk about bias in algorithms, and then we have a case study activity. Again, depending on how many people are here, will depend on the breakout rooms and then we'll finish with some actions to take away, and I have some reading recommendations. The slides will be shared out and everything. So, actually, they should have already been emailed out. So if you want to grab any of the recommendations that are there or any of the links, you have that available.

I wanted to introduce myself. I'm Britt Dzioba. I'm an advisor on the Learning and Teaching team here at BCcampus and you can see a photo of me there on my Dell Dimension in 2001. I've always been a little bit of a computer nerd. My dad is a computer scientist, so I always had access to the new latest tech to play around. He really encouraged my brother and I to explore technology and play with it. So that has influenced my career where I've been an educator at a tech company. I've worked with a senior’s organization in a tech project. I was also the project coordinator for the Digital Tattoo Project, and I think Paula has a link there to an article that I wrote, which was my start to getting into why I became interested in algorithms. I wrote an article about how the TikTok algorithm works. That was around the time where TikTok became popular in the pandemic. One of the things I just found fascinating about it is how complex algorithms are and how prevalent they are across all of our social media platforms and really just a huge social force to be reckoned with. And so that's always kind of been on the back of my mind through my career and just in my own life, when I engage with digital tools, I'm always kind of wondering what's behind this? How did this get created?

So we're going to start today with some Zoom polls. I'm just curious what tools you use regularly in your work? I think you can click, it's multi-select, select any of them that you use or have used. I think most people have answered. Somebody is using at least one of all of these tools at some point in their week. Okay. Thank you, Paula.

Then we have another question. There's no judgment on any of these at all, and it's all anonymous. I'd also really like to know. Have you ever tried any of these algorithmic power tools with students' work? Okay. Yeah, it's interesting to me at least, none of the above for the second poll. Okay. It's interesting that each one of these tools has been used by somebody at least once. Okay. Perfect. Well, all of these tools that I have in these polls, we're going to touch on today because all of them in some way relate to algorithms.

So I'm going to give a brief overview of what an algorithm is. And as I mentioned for those who came a bit late, this is part two to a workshop that I did in April. So if you're interested in that, we do have the recording available where you can learn a little bit more on a deeper level about the mechanics of algorithms. But at its most basic, an algorithm can be thought of as a recipe. So think of it as a recipe that provides clear, unambiguous instructions that can be followed to achieve a desired outcome. A more dictionary approved definition of an algorithm would be a step-by-step procedure or a set of rules designed to accomplish a specific task or solve a particular problem. In the digital world, algorithms are instructions that tell computers how to process information and make decisions. They take inputs, which would be data, perform specific operations on that data, and produce outputs, which would be the results. For instance, social media algorithms take information about what you've liked or viewed in the past, that would be the inputs. Process that information according to certain rules created by the social media platform and then decide what content to show you next, which would be the outputs. What makes digital algorithms powerful is that they can process vast amounts of information really quickly and consistently. However, they're limited by how they're designed and what data they're given, which is why understanding them is so important because they shape so many aspects of our digital experiences, but also reflect the priorities and biases of their creators. It's important to note that the term algorithm is not synonymous with GenAI. AI uses algorithms, but algorithms are found across our digital ecosystem, which is why algorithmic literacy is a bit different from GenAI literacy. Although they definitely have a ton of overlap, they're not the exact same thing. I'm just going to touch on a few elements of the mechanics of algorithms that will be helpful to explain their biases before we get into algorithmic literacy.

For those that like diagrams, this is just a simple diagram of how an algorithm would work.

But this diagram is probably a bit more accurate to the algorithms that we encounter frequently, which is related to the machine learning feedback loop. Large language models like ChatGPT or Claude will learn through a repeating cycle of pattern finding, testing, and human correction. The key idea is that the model doesn't understand language the way that people do. It learns probabilities and then gets shaped by feedback so that the answers align with human expectations. The first step would be the pre-training, learning patterns from text. So that would start with being trained on very large collections of text, books, websites, articles. The model looks for statistical patterns in how words and ideas tend to appear together. So it's not learning facts or meaning, it's learning probability patterns. And then the output of this phase is a model that can predict the next likely word in a sequence and it uses that. It's not actually even using words, it turns it into what are called tokens, which are in a binary code. This is often called the pre-training model, and by doing this billions of times, it learns patterns in grammar, style, reasoning, and even domain knowledge. So then it goes into the human feedback, evaluating what good looks like. Once the model can generate text, human reviewers step in and they can look at pairs of model responses and judge which ones are clearer, safer, more accurate, more helpful to real users. This produces labelled examples of what are better versus worse responses, and these labels become the models training signals for what we value. I think you can maybe start to see how some human influence here can start to create some biases. And then it moves into the reward model, which is turning feedback into a scoring system. The human judgments are used to train a smaller companion set system called the reward model, and its job is to predict how a human would rate any given AI response. Think of this as teaching the AI what kind of answers people want. And if you ever are using ChatGPT and it asks you thumbs up if you like this, that's part of that. It's looking for feedback. Then there's the fine-tuning phase. So teaching the model to aim for high-quality answers. The large language model is trained again this time using the reward model as a guide. It generates answers. The reward model scores them. Then the LLM, sorry, learns to produce higher scoring responses. And this stage is called reinforcement learning from human feedback. And it's what makes the model feel safer or more helpful or more aligned with everyday use. Although that may not actually be true, it's how we perceive it. Then the important fact about machine learning is that the loop never stops. So as people use the model, they add more information to it. It learns off that information. It's then trained on that feedback, and then that feedback gets fed back into the model and it's constantly cyclical. And I think what's important about to know about this loop is that LLMs are not static tools, so they're dynamic and they're constantly being shaped by human choices at every stage. They're not a neutral tool. Even if it seems like an algorithm in and of itself can't really have a bias. There's so many parts in this loop where human feedback is being integrated into the system.

So another more technical term that I want to cover, but I think it's important for understanding how LLMs work is temperature. This is the reason why two people can ask an LLM the same question or you can even ask LLM the same question twice, and you'll get different answers. Temperature controls the randomness of an LLM's output. It's a crucial feature, and you might think, why would LLM want to be random? Wouldn't it want to be super accurate? But there are various reasons why you might want a random answer, and a lot of that has to do with creativity. Even though the LLM in and of itself isn't being creative, when you have a high-temperature setting, that's going to generate a much vaster net that it's pulling its answers from. So that's why if you had an LM that had very low temperature and you're always getting very precise responses, people aren't going to be using it for, like, very dynamic reasons because it's not going to work for more creative projects, or even writing projects. It's going to produce very flat bland answers. So there's a reason why temperature would be in there. And I think if you have the pro version of ChatGPT, you can actually control the temperature. If anyone has that, I would recommend trying that in your settings. I think you can control the temperature and I think it's just 0 to 10 or something, with 10 being where it's going to be very random and 0 is when it's going to be very precise. Understanding temperature just makes it easier to recognize when an output might be overly creative, less reliable, or too rigid for the task at hand.

So one of the big issues with algorithms is that they're currently being used in ways that are very opaque. I think there was some senate hearing in the US and some senator was asking, I wish I wanted to try and find the clip, but I couldn't find it. But he was asking some leaders of tech companies, likely, show us the algorithm, show us what you're doing. But they said, That's nearly impossible at this point. We can't show you anything because LLMs and algorithms become so complex because they're constantly learning each other. The algorithm is almost impossible to discern. Even the engineers who created it can't really tell you can't really show you anything. And it also wouldn't really mean anything to a lay person who doesn't understand, you know, algorithms at a really high level. So, yeah, this opaqueness of algorithms has created this transparency paradox because the opaqueness limits user autonomy due to the lack of control and transparency, which is why they're often called black boxes. They're invisible, changeable, and inaccessible to users. There's three sources of this opacity. One would be corporate and state secrecy, so they don't want to share their codes because it's protecting proprietary interest. There's also technical expertise. A lot of tech companies don't want to share what they're doing because it's their IP, but also that most people aren't going to understand the complexity of the code anyway. Then there's system complexities. This is what I was explaining where the engineers don't even know the extent of their algorithms because they're interacting with so many other algorithms, it becomes a big, tangled mess. Obviously, from an ethical perspective, transparency is ideal because we want to know what these algorithms are, the datasets they're based on what they're doing. But there's a huge risk to full transparency, and this is where the paradox comes in because being completely transparent can also make these systems very vulnerable to manipulation. They can actually reduce their effectiveness and can increase bias and unfair outcomes in circumstances. So it's tricky. It's not an easy, well, just let everybody know what the algorithms are. It's much more complex than that and very difficult to parse out these algorithmic structures.

So that leads me into the topic of algorithm literacy. Algorithm literacy is the ability to understand how algorithms shape the information, opportunities, and interactions we experience in a digital system. I would argue that it's not important to actually know the code of an algorithm to be algorithmically literate. You can understand. I mean, even understanding that algorithms are so complex, we can't parse out that information is part of algorithm literacy. So in everyday terms, it means being able to answer questions like, what is this system trying to do? What data is he using? How might the data be incomplete or skewed? How is the output influencing my decisions, behaviours, and opportunities? So it's not about learning the code. It's about being able to critically interpret algorithmic decisions in a way we interpret media, statistics or research. So in education, as an example, algorithm literacy can help us understand how AI tools generate answers, identify when a system might be biased or making assumptions. Question recommendations that were being provided. And using our digital tools more safely, ethically, and effectively. So I also wanted to point out that this concept is a new concept to me, but a newish concept to me, and I really enjoy it. It's called Bildung and it comes from the German philosophy tradition, most notably the works of Heigel. If you have a literature background like me, you might be familiar with the term Bildungs Romance. Bildung is the combination of the education and knowledge necessary to thrive in your society and the moral and emotional maturity to be both a team player and have personal autonomy. Bildung is also knowing your roots and being able to imagine the future. So there's a concept called digital Bildung, which was coined by David Kurgel and it's the process of developing the critical, ethical, and creative capacities individuals need to navigate, shape, and reflect on digital technologies. So it emphasizes becoming not just a skilled user of technology, but a thoughtful, self-determined, and socially responsible participant in the digitally mediated society. So this could include the critical awareness. So it's understanding how the digital systems work, knowing what platforms are being shaped, knowing that our platforms are shaped by algorithms as part of that. Agency and autonomy, cultivating the ability to make more informed and independent choices. Ethical and civic engagement, grappling with questions of fairness, privacy, democracy, and digital rights, and cultural and creative participation. Using digital tools to create, collaborate, and contribute to our cultural and intellectual life. So integrating the concept of Bildung into a definition of algorithmic literacy emphasizes that an individual has agency over the algorithm and is further encouraged, supported, and challenged to become a contributing member of our digital society.

I just wanted to provide this graph here. Sorry if the text is a little blurry, just to continue on that critical algorithm literacy angle. This concept of digital Bildung in relation to algorithms can also be thought of under this umbrella where a knowledge and awareness of algorithms is your first step to becoming more algorithmically literate. It's a bit of a tongue twister. Through agency in an algorithmic powered world is critically reflecting on how algorithms shape your environment and knowing what you can do to influence your relationship with algorithms.

So that takes us into our first little activity here. I can't see how many participants. Okay. So we're going to do a Think-Pair-Share, a quick activity. Paula is going to pop in a Google Doc into the chat there. If you open the Google Doc, actually, maybe I can open it and show you.

There's a link at the top here. If you click on that link, it's going to open this assumptions card set. If you press the Start button, it's going to give you a random assumption. There's about 21 in there, so it should be enough for random so you can click through to find one if you don't like the one that you got. That's really fine. We're going to give you 2 minutes to think on your own about that assumption. I have some prompting questions to help you there, which are right down here, consider the following prompts. And we're going to pair you up in a breakup room so it'll just be two of you in a breakout room. You'll have 4 minutes to discuss your assumption or the assumption card that you got and your thoughts around it, and then we'll come back and have a little bit of time to share. So I'm going to do 2 minutes, starting, 2 minutes, starting now to do your assumption card.

Okay. That's the end of the 2 minutes for the Think. Perfect. So it'll be 4 minutes. Okay, I think everyone's back now. Did anybody want to share thoughts from their group? Any reflections you had based on these questions or just on the assumption? You can either pop it in the chat or unmute.

PARTICIPANT:
She put me on the spot, for sure. What we recognized is that there's just a lot of assumptions that you might make, and you really don't know for sure. So there's risks and it may not be true, so you have to be really careful to make sure that you're supporting the students the best you can in different diverse ways. Hope that captures it.

BRITT:
Great. Thank you. Did anyone else want to share?

PARTICIPANT:
I'll try to be brief. I was in a group with one of our other members, and the card that they got said that quiet students are not factored in. I'm summarizing quickly, and the one that I got said More data automatically creates better predictions. And something that I think jives with both of those points together is a larger point that there's a limit to so-called quote unquote intelligence, be it artificial or human, because there's just so much of the world that is always going to be unknowable. And, you know, an example that comes to my mind, it's just a classic one from the work of Michel Foucault, which is the idea that, you know, if you have policing in one neighbourhood and that neighbourhood is full of people that have been racialized and are out on parole, then that policing is going to treat people in that neighbourhood as if them having a beer on their front stoop is a criminal act. Whereas, if you go to a rich neighbourhood, you know, people having a glass of wine in their front stoop, they're going to be treated like a good upstanding citizen. So, you know, there's going to be more data on the people in one neighbourhood rather than another, but it certainly doesn't create better predictions.

BRITT:
Absolutely, I love a Foucault reference. And I think exactly what you're talking about is what we're going to talk about next when we get into algorithmic bias because yeah, those are some great points. Clint, did you want to contribute?

CLINT:
Yeah, I was just going to say Denise and I were in a breakout room here. I think what we were both doing was challenging the assumptions around the assumptions that we were getting. So the one I got was Early assignment submissions indicate stronger motivation. And then we started going, Well, is that necessarily true? Is that a dangerous assumption? Maybe some students are just trying to stay ahead of their work and not necessarily motivated for this particular class. Maybe they have family obligations, maybe they have. I think we were both with the assumptions that we got, we're challenging the assumptions behind the assumptions.

BRITT:
Right. Yeah, absolutely. Great. Yeah, the purpose, I think, of that activity was really to just get us primed to think about the way that various assumptions can be made and sometimes unintentionally and how they might appear in our work. So I think in an isolated activity like this, it's probably very easy to see where the bias is occurring because we're prompted to think about it. It's very, you know, it's on a card for you. But in our daily lives, bias is much harder to detect. And it's way more nuanced. I know for myself, I often really struggle with that white coat bias. So I take the opinions of people with professional designations very seriously. Of course, you were saying sometimes some students are very motivated and want to get things done soon, but sometimes that's not the case. A lot of times people with designations are very, capable, very, are correct in their opinions and are people that we should believe. But I have caught myself taking somebody's word at face value without critically thinking about it, even when it doesn't sit right with me, just because I'm afraid to disagree with somebody who I think has more authority than me. Marika, I see that you have your hand raised. Do you want to begin?

MARIKA:
Thank you. I apologize for not being able to. I'm just pulled over. I'm driving. The thing about bias that I think is really important. And especially, some of our greatest source of bias right now is the mental shortcuts that we're required to take in highly distressed environments and with work volume and pace. And mental shortcuts create a lot of bias, and we're really unaware of it. So I work in health care in a lot of education initiatives, and, you know, we see that from the moral distress residue and being a trauma organized system. And now with the use of algorithms, we've always used algorithms. Now, there's more of a leaning on algorithms as the only source of truth because it does facilitate shortcuts. But the bias, I work in mental health and substance use, and the amount of bias and perpetuation of stigma and stuff is really harmful. So I appreciate the topic. Appreciate you, Britt, for doing this. I wanted to join, even though I'm driving out to Hope, but I just wanted to be with all of you. So thank you.

BRITT:
Well, we really appreciate you being here Marika, and that's such a great point about when you're in a high stress situation, the mental load you want to offload that load. So if you can rely on something like an AI tool to help you do that, it's totally I can absolutely see how that would be tempting to do and totally warranted, as well. And if anybody was in a part one of this workshop back in April, we did an activity where it was this game online and you are basically a hiring manager, and you have to make really quick hiring decisions, and the game speeds up and it gets progressively faster and faster and faster and faster. And then at the end, it shows you where your biases were because you're having to make these really quick decisions on hiring. And everyone who did the activity really reflected as soon as the pace picked up and the pressure got higher, I just was like, not even critically thinking about the resumes I was selecting. I think obviously that's a fun little game, but I really do think that can be applied to the greater work world as well, and the education world because we are in a strained environment right now, we're expected to do a lot of output and with much fewer resources. AI is offering us ways to manage our cognitive load, then totally makes sense to do that. But I think being aware of where those biases can really create a major impact is what's important.

So I think that really dovetails nicely into the topic of algorithmic bias. So just a Wikipedia, brief definition of algorithmic bias is when algorithms commit systemic errors that unfairly favour or discriminate against certain groups of people. These biases are the result of poor training data and the biases of humans who compile the data and train the algorithms. So there's a number of ways in which a bias can kind of get encoded into an algorithm. So as mentioned in the training, that would be the data input itself and also the feedback loop. If we think back to that diagram I showed, you can see how human input kind of comes in at multiple aspects of that feedback loop. So there's lots of opportunity for bias to be reinforced. So in the design of an algorithm, it could be the developers own biases. They're math, but they're created by humans who have their own lens that are applying to their work and also, algorithms are very susceptible to correlation bias because they're often lacking nuance. So they can often make correlations between two things and say that it's causation when that's not necessarily the case where a human eye could really see that that's a fallacy. There's also proxy data. These are assumptions made where certain data points are unavailable. So in the case of area codes that are associated with lower economic status when no other information is available, what Jovian was saying about how certain areas get typified as being a high crime area, but they're only looking at one data point and then also the evaluation. Even if you could find a completely neutral algorithm, the output is being interpreted by us, and we all have bias as much as we know, some of us really try, you know, to critically reflect on that. We all do have bias, and that is going to be applied to the output as well. And then that system kind of gets reinforced. So you can really see how it gets quite tangled there.

So I wanted to share just a very short clip from a great podcast episode. I mean, should be. So it's pretty short. So I'm going to just give her one answer here that I really enjoyed.

[PODCAST]

That data and technology isn't merely found, it is produced. Can you talk a little bit about why that distinction matters and why the producers matter? Absolutely. We have to think about why we ask this question and not this one? Why did we pose it in this way and not this way? That is part of the production process, even before you start thinking about where to look for the data. In the context of the new Jim code and discriminatory design, one of the main, sort of channels that's producing this inequity is the data that's used to train automated systems, the training data. If you're pulling, let's say you're trying to decide what teachers to hire in your school, and you're basing it on those teachers who've excelled previously. That's your training data. But let's say for the last 50 years you've only hired teachers that come from three elite schools, let's say, and that history is built into who is excelling now, and that means you're likely to get more of the same. Your automated system is going to learn what you have previously thought was a good teacher and give you more of that, which means if you in theory are trying to broaden that pool, you're not likely to do it unless you go back and look at that training data. Then in that way we are both reproducing this history and erasing it at the same time because we think of these automated systems as somehow removed from the past and removed from our ongoing social practices. It sounds pretty easy...

BRITT:
So I thought that that's from Rua Benjamin, who is a professor at Princeton, and she wrote a book about I linked it in one of the book resources, but she wrote a book about kind of like the new Jim code, if you're familiar with, you know, Jim Crow era and about how discriminatory design exists in our technical space. And that podcast, so the link is in the slides there, is really interesting, as well. So I'd recommend reading it. So I do have a few real world examples about algorithmic bias that I wanted to share, and then we will have a little bit of time for a case study activity.

The first example that I wanted to share was about an AI grading system in the UK. So every summer in Britain, students sit to write their A-level exams, which are kind of like the SATs and they help determine what university you can go to. They're quite high stakes. Because of the pandemic, the British government had to cancel the A levels. Instead, the government had teachers give an estimate of how they thought their students would perform on the exams. These predicted grades were then adjusted by Ofqual, which is England's regulatory agency, using an algorithm that weighted the scores based on a historic performance of individual secondary schools. Then that weighted score, sorry, the idea was that the algorithm would compensate for the tendencies of teachers to inflate the expected performance of their students and more accurately predict how test takers would have actually performed. But when the students received their predicted A level results, many discovered that they had scored much lower than expected or were reflective of their previous grades or their mock exams. For some, the algorithm weighted results meant that they were now ineligible for university programs they had expected to attend, potentially altering the course of their future. Around 40% of the predicted performances were downgraded, well, only 2% of the marks increased. But one of the most important findings from independent research done on this situation was that the biggest victims were students with high grades from less advantaged schools who were more likely to have their score downgraded. While students from richer schools were more likely to have their score upgraded. So while Ofqual reported that there was no grading bias, it was found that the algorithm favoured students from more economically privileged backgrounds, while others suffered, well, others from lower economic backgrounds suffered. This was due to each school's historic results being a significant factor in the algorithm's grade calculation. Ofqual was very quick to blame overly generous teachers, but made no comment on the algorithm itself. So I think this is a really good example that connects to that podcast episode that the algorithms are not just taken in a vacuum, right? They are building in a lot of historical knowledge there as well, going back to the Foucault reference, if students from lower economic schools are typified to students that are going to do worse, an algorithm doesn't have the nuance to critically reflect on that, so it's just going to continue to perpetrate bias.

This is another study which was published out the University of Washington. They looked at three advanced large language models used for resume screening and discovered substantial bias across race and gender discrimination. They tested 500 real world resumes with varied names associated with different racial and gender groups, and it found that it favoured white associated names 85% of the time compared to just 9% for Black associated names, and it showed a strong bias towards male associated names, preferring them 52% of the time as opposed to 11% of the time. Research revealed unique intersectional harms, particularly against Black men as the systems never preferred Black male names over white male names. Yet favoured Black female associated names, 60% of the time compared to 15% of the time for Black male associated names. So these findings raise major concerns for the rapid adoption of AI hiring tools, which are now used by an estimated 99% of Fortune 500 companies. So yeah, very interesting and worrying statistics.

So this was a piece by the Brookings Institute about enrollment algorithms. So these are being used by over 700 institutions across the US now, and they employ a two-step process to optimize student recruitment. So first, predicting which accepted applicants are likely to enroll and then using that data to strategically allocate scholarships to maximize either enrollment yield or net tuition revenue. And while these algorithms can help colleges with financial planning and can increase enrollment efficiency, especially when resources are decreased, the evidence suggests that they typically reduce the average amount of scholarship funding offered to students, and that often impacts low-income students the most. I tried to find whether there was any data on how many Canadian universities use this. I couldn't find exact data on Canadian universities using AI in the enrollment process, but I did find this post from Queens University saying that they do allow AI in their student recruitment process. To me, that suggests that this is something that is being thought of and happening. We just don't really have the data yet to see how widespread it is. But I also did find an article in the National Post about Border Pass and Apply Board, which are two Canadian companies that help international students access education abroad. Recently they have introduced AI technology that ranks university applications to determine their likelihood of receiving a provincial attestation letter. Canada's latest requirement for international students. So an applicant's chances of approval are based on financial stability, visa history, and other risk factors. You can see again, where certain students are being privileged by these algorithms versus others.

I just quickly because we do have the case study activity. These next two examples, I think really highlight something that I'm going to talk about in a second, which is the flattening effect. So this was an article about students using chat bots and how when the students would get their reaction from the chat bot, it was really changing their behaviour. So they were really trying to manipulate their responses in order to get a good grade from the chat bot. You're not really getting these authentic, creative, innovative responses from students anymore if they're only trying to match the expectations of a chat bot. And if we're not auditing these chat bots to see what they're privileging or what kind of answers they're interested in, then you're just going to get a whole bunch of answers that are essentially the same.

This is a comment from another article that's very similar, which is when an LLM is rewarding, when it's being rewarded by finding faults in things, it's going to change the way that people provide answers because nobody likes to do poorly. People want to do well. So they're going to try to mould into algorithmic perfectionism. And there's quite a bit of research that's showing that this is creating a flattening effect. AI grading and feedback tools, they disincentivize students from trying innovative answers and doing creative responses because it doesn't fit the mould of what the AI has determined to be the model answer. Students are going to learn that very quickly and start to gamify the system. And I also like this quote from this paper. “If AI grading systems are trained on limited or biased data set, it may inadvertently perpetrate or amplify existing inequalities disadvantaging certain groups of students. An AI tool trained primarily on business plans for male-led start-ups in certain industries might inadvertently penalize business plans that address specific needs or challenges for women, non-binary, and other unrepresented gender identities." I think we've seen that in these past examples that we don't really know the history of these datasets, and in business, particularly if you want to take that example, obviously, the business world has been dominated by, you know, upper, middle class white cis straight men. So if their history of, you know, business plans and business research is being used to train datasets are being used that are being used now, it's not reflective of the current business landscape.

That was quite a bit to chat about. We don't have a ton of time, but I did want to do this case study activity. So I think we could probably. They're not super long. There's case study one and case study two, so maybe we can pop into two breakout rooms, Paula, just split evenly. If you're in group one, you'll do case study one. If you're in group two, you'll do case study two. They are in the Google Doc that was shared out. If you scroll down, there is the activity two. I'm just also reading some comments here as well. We'll go into our groups and then you'll have a few minutes and then bring everybody together and wrap up at the end.

So if there was, so we have some time now to debrief the case studies. And then I think hopefully, depending on how long it goes, we'll have some time to if you want me to go over anything again or if you have any questions, we should have some time for a more general debrief as well. And in the main room, Marika and I are having a great conversation about AI and health care right now and some of the concerns there, especially if bias is getting integrated into a health care system that already has a lot of biases. So yeah, we had a nice little chat. So group one for case study one, who would like to chat about that? About automatic feedback software and how that might impact students? I don't even remember who was in group one. So, Maria, Clint, Judy, Monica. Would anybody like to chat about what their thoughts were?

JUDY:
I'll start. I'm expecting the rest of the crew to add. We had a lovely conversation, wide ranging for sure. There's some issues and some concerns about this. I mean, it's not good learning. So we talked about, you know, different things that we could do and good old fashioned, you know, conversation was one of the options. Some creative alternatives were kicked around. I'm going to count on the rest of the group to fill in the blanks. There's a start. Great.

PARTICIPANT:
Thanks, Judy, I can add to that. Looking at the case, formative comments on clarity, structure, and argument quality, putting ourselves in the shoes of a student, getting all that feedback in one go, from what I see is overwhelming and the students wouldn't really know how to reflect on the feedback that they've received and how to fix things. That's one thing. Also when we're talking about argumentative writing, higher levels of thinking, we're not really encouraging that. We're just providing the automatic feedback. I often find with the students a little bit of a back and forth conversation with them is a lot more helpful than just giving them a standardized automated feedback from a software. And then we further talked about the responsibilities as educators that we may have using automated feedback systems, especially when it comes to diverse student populations. We all I think agree that this is definitely not going to address the different needs that we come across in our classrooms.

Britt:
One thing I'm thinking about with this case study is also, you know, students whose English might not be their first language are going to be really impacted by a software like this. We have research that shows that AI detection tools often unfairly target students who are English isn't their first language. Or English as an additional language, so I think I could see that bias being perpetrated in a software like this. And also, I think, in this case, thinking about what is the purpose of the assignment? I had a good conversation at an AI workshop at Douglas College with an instructor there in their Creative Writing department, and he said that his biggest fear is students losing their voice. He's like, my role is to help a student develop their voice, develop their flavour in writing. And if they're using AI, it's going to be flattening, like what we talked about, flattening, flattening their voice because it doesn't understand nuance, it doesn't prioritize or incentivize creativity or innovation. So I think, if the assignment is just hand in an academically structured paper, then maybe these feedback tools are good because it's going to stick to that rigour. But if you're wanting to help students develop their voice, to take risks, to try new things, then I think taking a critical look at tools like this is really important and auditing

these tools and really keeping track of the feedback it's giving to see where some of those biases might be coming into play. So thank you so much for group one. What about group two? This is helping students recognize algorithmic bias using a library search tool. Couldn't think of the word there. The group two is Helena and Denise.

HELENA:
I'm just starting us off. I'm going to count on the team. We didn't get through all three, but I did make notes. I think some of the responsibilities of the instructor as a subject matter expert to provide diverse resources from the beginning too and to really highlight those that are not within the typical Western English language authorship. That's one thing because we also acknowledge that students don't know what they don't know. So maybe there's some guidance from instructors required. Also, the suggestion was to talk to students about what you observe, and I think that's what we say often even around AI too, right? They just have those conversations in the classroom, make time for that. There was also the suggestion of how important it is and the action of students to ask for more than one resource to specifically point to historical as well as current and contemporary. That was kind of the first bullet. I think we got a little bit stuck on how to redesign. Because I just want to acknowledge, like, you know, the examples that were shared are what you just said about the voice of students being lost. Like very sad when there's a lot of effort being put into inviting students to bring personal experience in as part of the redesign, and then students still choosing to use AI, and then it results in a very inauthentic response rather than the authentic one, which in the past was not necessarily a problem. So that's kind of where we got stuck. There's a shift to in-class assignments, which I just want to say is also could be problematic if there's not still a choice for those who have difficulty with that and processing and then maybe creating podcasts, but we've heard also that there's a flip side to that. So yeah, anyone on the team that wants to add to those comments.

DENISE:
I would just add that this is something that I definitely struggle with and thinking about how to help students with it because I see that a lot of my students, many of them are fairly new to studies at a university. They have this misconception that if they're getting something from Google's AI, it's somehow objective and they're really not aware of the biases that are built into AI platforms and helping them think through that and really recognize those biases. The sorts of conversations we're having here about what biases are there, how do you detect them? How do you deal with that? Seem to be an important part of a course that's not really part of the course content per se, but it is something students are going to have to talk about in order to understand these tools and when it's appropriate to use them and how it's appropriate to use them and how to get beyond them into more meaningful work that the students are actually producing.

BRITT:
Absolutely. I think that's the hard thing Denise, what you mentioned about, it's not part of the course content, but it's so important and it's hard because it adds more work onto instructors

and educators to have to take on this work, but it is so important because we're seeing it crop up time and time again. So it's like, how do you integrate that conversation around bias and AI into course content? I think it's a message that needs to be told time and time again. If I had a magic wand, we had to ask this question actually in a podcast Helena and I did about if we had a magic wand, what would we do for higher ed that we could change? One of the things for me has always been there should be a Digital Literacy 101 course that is required for all students to take and how English 101 is required, learning about these research, research in an AI age and about filter bubbles and algorithmic bias and things like that Google isn't neutral. I really dislike the new Google AI thing, that summary thing that pops up and you can't turn it off and it is so frequently wrong. I've seen so many of examples of it being wrong. I've also fallen prey to just quickly checking it, too. So I'm not saying that I'm, you know, doing an amazing job because it is tempting when you're just trying to quickly look up something. I find that often it really falls into the I don't know what you call it, it really tries to agree with you on things. So depending on how you Google something, they actually did some interesting research about how people google a question will give up certain answers. So, if, you know, if you get Google, Is smoking that bad for you question mark, you would get results that would kind of downplay the harms of smoking versus Googling the negative health effects of smoking, which would produce much stronger cases against smoking. That's a very rudimentary example, but even how we're Googling things changes the answers that we're getting and I think the AI feature of that really is susceptible to that. I'm just reading through some comments here because Jovian had said, even before the rise of AI, there has been an intense conformist pressure through neoliberal capitalism. When we mark papers, we are testing high conformity to the standards of the elite. I’m stressed that this was already a problem, so I'm worried AI will make it much worse. Yeah. That's a very, very real worry right now. And these biases and standards already existed, and so we heard the history of our societal structures are being reinforced in AI. So we have to be very aware of that. Yeah, there are also different browser options that are not Google. You know, I've tried Duck Duck Go before. It's not as good, it's not as user friendly, but it is a much more ethical platform for sure. And yeah, so I think kind of one of the important things here is reminding students that these things like Google searches are not without bias. They're not neutral. It's not pulling just, you know, the consensus of what the best outcome is. It's being influenced by a lot of other factors, not just the algorithm, but also people paying for things to be on the top. The sponsor thing is a lot more sneaky now, I've noticed than it used to be. It's not as clear now which results are the sponsored ones. So yeah, always keeping an eye out for that. So we have about 5 minutes left.

I wanted to leave you with a few action items and then open the floor for any questions. As I know, there was quite a bit of content covered here and apologies if I kind of sped through at the end there. I was worried about time for no actual reason. So here are just a few things that I kind of thought of, for me, one of the things is these topics can be very depressing and it can feel like, what can I even do? Because the machine seems so much bigger than me as an individual. But one of the things that I learned from the climate justice movement is that small actions are the antidote to despair. So whenever you can take a small action, this is what keeps the momentum going for yourself and for your community, and I think not to think too big a picture about it. What small actions can you take? I think that sentiment applies across our lives actually. One of the things is to believe students when they report concerns about being treated unfairly. If they think that their grades are if AI is in the process and they're being treated unfairly, I'd say believe them and hear them out. Document patterns. If you are using AI tools for feedback or evaluation, document what are the patterns that you're seeing. If certain students are always getting similar grades, why is that? I would say provide alternative pathways if you can. Offer non-algorithmic options like an oral exam instead of a proctored online exam. We did go more into proctoring in part one of this session. We didn't really cover that here, but that's a tool that has been shown to have a wide bias. And communicating. So we just talked about this, letting students know if their data is being collected or how it's being used, but also communicating to them about algorithms and how that might interact with the materials that they're being asked to use or how that might appear in their work. And ask students to use diverse sources. So, create that into the assignment that, you know, it's not just pulling from the top results from the library search from Google search, ask them to actually think outside the box and go for some sources that are non-academic or from marginalized voices.

I also have some book recommendations here, so these are in the slides. You know, Christmas is coming up, so here's some stocking stuff or ideas for you if you want to learn more about algorithm literacy. And then I have my references at the end here. But all the slides were shared out, feel free to grab them in there. Thank you all. I know it was a small group, but everyone contributed, and I think we had some great conversations.