Transcription

Can Computers Be Racist? Spoiler Alert: Yes!

Please note: The below transcription is auto generated, and may contain some errors. Efforts have been made to best avoid any inconsistencies or inaccuracies. Sentence structure and word choice may not follow typical grammatical patterns.

Sarah: Hello and welcome everyone. My name is Sarah Garber, and I'm a junior math and adolescent education major here at Naz. I will be your host for our discussion tonight. I want to say thank you to our guests for attending this special event this evening. We are joined by our speakers tonight from the newly launched Institute for technology, artificial intelligence and society, to talk about how bias operates within the context of machine learning and artificial intelligence. We look forward to this fascinating discussion and we also welcome you to participate during the breakout session that will be part of tonight's presentation.

Before we proceed I'd like to just share some information about our event tonight. First we are recording this session tonight, and we will be sharing this video recording with you in the next couple of weeks. And as a part of our commitment to community building here at Nazareth, we ask that you keep your camera on if possible. Also please change your zoom username to your first and last name if you haven't already. Due to the number of attendees tonight we will keep all microphones muted to prevent any feedback issues. And if you have any questions for the panelists, please send them in the chat to Adonis Savidis math major and student volunteer, who's helping with moderating the questions for the evening. For any other technical questions please feel free to send a message in the chat to our technical support assistant. And lastly, we would recommend setting your zoom to Speaker view, during the presentation. Now that we have that out of the way I'd like to take a moment to briefly introduce our panelists for tonight.

So first we have Dr. Dianne Oliver. She is the Co-director of the Nazareth college Institute for technology AI and society, and she came to Naz as dean of the College of Arts and Sciences at Nazareth in 2015. Her background includes degrees in both computer science and religion and ethics, deeply connecting her experience and work to this initiative at Naz. Next we have Dr. Yousuf George. He is the Associate to the President for strategy and momentum and a co director of the college's new Institute for technology, artificial intelligence and society, he joined the college in 2008 as a faculty member in the Department of Mathematics, and later served as Associate Dean of the College of Arts and Sciences. Next up we have Dr. Wendy Norris. She is an assistant professor and founding faculty member for the ITAS. Wendy has a PhD and Information Science and brings expertise and the design of humanitarian crisis response technologies to her teaching and research. She joined the mathematics department in 2020, to help launch Nazareth’s, new ethical data science major. And finally we have Dr. Chelsea Wahl, another founding faculty member for ITAS who joined Nazareth’s sociology and anthropology department as an assistant professor of sociology in 2020. She earned her bachelor's degree in sociology from Hamilton College, and her doctorate from University of Pennsylvania. She specializes in technology and quality, work and organizations. And, here you go.

Yousuf: All right, I'm going to attempt to share my screen, Excellent. Well, thanks everybody for being with us tonight. It's really fun to have you here at one of our very first events for the kickoff of our new Institute. So tonight's discussion is about the question. Can computers be racist? We don't want to keep you waiting for the answer, too long. It's Yes. So we'll talk quite a bit about this tonight, but before we do, I just wanted to give you a brief introduction to the Institute. So this new Institute as Sarah mentioned is really devoted to the social and ethical concerns around technology broadly, and artificial intelligence very specifically. Our goal is to promote a just and equitable technology future. And we want to develop the next gen of ethical tech leaders. Dianne's going to tell us a little bit about how we got where we are.

Dianne: It is interesting since those of you who are on this on this zoom are folks who are alums, as you may be going. Why is this talking about computing and artificial intelligence and technology? This has not really been something that we would think about when we thought about Nazareth College, but we really wanted to explore starting a few years ago, how we were going to engage as an institution in technology and computing. We know that technology is transforming the ways that humans live, work and govern. What was Naz going to do, we couldn't really, you know, quit paying attention to this. I mean not really pay attention to this and so we really needed to find a way that this would become part of the experience that our students had who are here as well, given the fact that we still have a majority female student body about 75% of our students are female. We also know that this group has been highly underrepresented in technology fields and I can affirm that from my experience in the software development world. We also know that many of the discussions around technology and computing and artificial intelligence in particular, tend to be limited to engineering centric schools places that we're familiar with like RIT or Cornell, and maybe really focused on technical majors like engineering or computer science, rather than places like NAS that have a broad foundation in the Liberal Arts and Sciences, even for those professional focus majors that we have. We know that there's a lot of evidence to suggest leaving those things that are it and Cornell are not the best idea. So again, how are we going to kind of going into Naz's future how are we going to attend to technology and computing, how are we as a community, going to prepare women and men to be involved in STEM, but keep it rooted in who we are as a community rooted in the liberal arts broadly and ethics in particular. And we knew we wanted whatever we were going to do in the, in the realm of tech we wanted it to be rooted in as as values with a social justice focus connected to our mission as an institution connected to how the Sisters of St. Joseph founded us in 1924, in order to create an opportunity for women's voices to be heard and to be involved in education. So we think Naz might have a unique contribution to make in the realm of technology bringing the Liberal Arts and Sciences into intentional work in technology, and in particular the pervasive field of artificial intelligence and to use those values of social justice and attention to voices on the margins, and thus we ended up with an institute, called the institute for technology artificial intelligence in society. We want to help Naz students emerge into this pervasive world of technology, not, not so much as the tech creators though some of them will be, but also as those who are, who understand the issues around technology the ethical concerns the possibilities but also the challenges that this pervasive place of technology and artificial intelligence in particular brings to us. We want our

students grounded in the liberal arts as Nazareth students always have been, which we know prepare students for whatever is coming. We want them to do that in the realm of tech as well and the realm of technology and so that's where we find ourselves. Now we find ourselves wanting to create that for all of our students we want students who are going to go into work in tech, but also in policy or an advocacy in businesses and and not for profits to give them the tools to harness the potential of technology, but also to pay attention to the issues and so that's a lot of what we're going to be talking about tonight is that very central premise for us that we want to engage the ethical issues around technology and how technology, and artificial intelligence in particular are impacting the society in which we live. So we want to do tech, the Naz way.

Wendy: Hi everyone I'm so glad that you were able to join us tonight so what I thought I would do is a little foregrounding. I'm going to teach you everything you need to know about computer science in a minute. Often these terms are thrown around and they're used interchangeably in the popular press, but they have very distinct characteristics and purposes so I thought we take a moment to just kind of untangle, what is an algorithm, what is machine learning and what is artificial intelligence. So an algorithm is simply a sequence of instructions used by a computer to solve a problem and in our world, you could think of that like a recipe. If you're going to make a cake you can put all the ingredients in a bowl you know stir it all up you're going to put it in the oven. And then you're going to turn on the oven and preheat it? No, that's not the way it works. So everything needs to be an order. And this is exactly the way that computers work, just a sequence of events. So machine learning and on the other hand sort of builds on that and it detects patterns in large amounts of data and you may have heard the term, big data before. And what this is doing is it's actually training the machine to become even more efficient in the ways that it is understanding and detecting patterns. And then we build on that sequentially into artificial intelligence, and this is using both algorithms, those sequences of instructions and the machine learning to help make decisions or predictions about people's behaviors. So, what we're trying to do here is sort of understand these different terminologies because each of them require very different types of interventions to advance social justice goals. So let me talk a little bit about a topic that's very much in the news these days and that's facial recognition. So what we have here is again sort of a lot of confusion in the popular press about what facial recognition is, and some of the ways that it is sort of interfacing with us on an everyday basis. So these are just digital tools used to perform tasks on images or videos of human faces and there's three different ways that this works. One is just a face detection. Is there a face in this image? The other one is facial attributes will kind of face is in the image, what is the expression. You know what kinds of things can I predict from this face, and then there's facial verification and this is whose face is in the image, and we want to show you a very brief video to kind of give you some sense of what facial recognition technologies are, and some of the ethical dilemmas that are being kind of put forward in this technology.

YouTube VIDEO-Gender Shades
Hi, I'm Joy, and I research how computers detect, recognize and classify people's faces in my Ted featured talk I spoke about my experience with the coded gaze, my term for algorithmic

buyers. The system I was using worked well on my lighter skinned friends face, but when it came to detecting my face. It didn't do so well, until I put on a white mask. After my talk was posted I tested my speaker image profile across different facial analysis demos, two of the demos didn't detect my face. The other two. Well, they mis-gendered me the demos didn't even distinguish between gender identity and biological sex, they just provided two labels, male and female. Now I wanted to see if these results were just because of my unique facial features, or if this was something that was more of a pattern across other faces too. So I began a project that became my MIT thesis ‘gender shades’, or the long title gender shades intersectional phenotypic and demographic evaluation of face data sets and gender classifiers, or just ‘gender shades’. I wanted to see how well different gender classification systems work across different people's faces, and if the results changed based on somebody's gender, or their skin type. I created a data set of over 1000 images of parliament members ranked among the top 10 in the world, based on their representation of women and power to get out a range of skin types. I chose three African countries and three European countries, so I could see how the system performed on lighter skin and darker skin. Then I chose three companies to evaluate, IBM, Microsoft, and Face++, which has access to one of the largest data sets of Chinese faces. So now with the data set, and the companies, I decided to run a test the companies appear to have relatively high accuracy. Overall, Microsoft performed best achieving 94% accuracy, on the whole dataset, all companies perform better on males than females and all companies also perform better on lighter subjects than on darker subjects. When we analyze the results by four subgroups. We saw that all companies performed worse on darker females. IBM and Microsoft performed best on lighter males and Face++ performing best on darker males compared to the other. IBM had the largest gap in accuracy, but the difference of 34% in error rates between lighter males and darker females. I was surprised to see multiple commercial products failing on over one in three women of color. In fact, as we tested women with darker and darker skin, the chances of being correctly gendered came close to a coin toss. While more research is needed to explore the specific reasons for the accuracy differences, one general issue is the lack of diversity and training images and benchmark data sets. Failure to separate accuracy results across traits like gender and skin type, also makes it harder to identify differences, companies should do better with commercially sold products, especially since the machine learning techniques that have made gender classification possible are applied to other domains of computer vision, like facial recognition and other areas of artificial intelligence, like predictive analytics. Predictive systems can help determine who is hired, granted alone, or what information, a particular individual sees these data centric technologies are vulnerable to bias and abuse. As a result, we must demand more transparency, and more accountability. We have entered the age of automation overconfident yet underprepared. If we fail to make ethical and inclusive artificial intelligence, we risk losing gains made in civil rights, and gender equity under the guise of machine neutrality. The coded gaze reflected in the gender shades project must be faced.

Wendy:
So the intersections of race and gender in facial recognition technologies are something that is very much in the fore of research and computer science right now. But there are very real world

consequences to facial recognition technology, when it fails. And, you know, now we're going to just take a quick look at a video by the ACLU on a gentleman who was wrongly arrested, due to facial recognition technology

YouTube VIDEO-ACLU

So when I got into the interview room, the first thing they had me do was read my rights to myself. And then sign off that I read and understand my rights. The detective turns over a picture of a guy, and says I know he's like, so that's not you? I look. I said no, that's not me. He turns another paper over and he says I guess this, not you either. I picked that paper up and I hold it next to my face. And I said, this is not me. Like I hope y'all don't think all black people look alike. And then he says the computer says you. At that time, no, I didn't know that they use any type of facial recognition or anything like that. Until, talking with the detectives who showed me that. And that's what they use to apprehend me, if that's the right word. The morning everybody gets dressed, everybody gets ready. Lissa leaves and takes Rosie to my mother in law. And then I dropped JuJu off at school, it's just a boring Thursday, right? Then I'm on the phone with Lissa around four. So then my line clicks, and I click over. I'm like hello guys talking like he knows me is Robert and I'm like, Who is this? You need to come down, turn yourself in, Who is this? Officer somebody from the third precinct, and I need to turn myself in for what? And he was like I can't tell you that. And I'm like, then I can't come down. Well, if you come down to be much easier on you don't want us to come out to your job do you? At this point I think it's a prank call. So, I'm like, Look, man. If you want me to come get me I'll be at home. Bring a warrant. And I hung up on him. My mom got home with Rosie, and we were walking through and as I was walking through, I looked out and the cop car was outside and I said, Oh, so they will, it wasn't a prank call there really are people here. They came to the door. I answered it, and it kind of stuck their foot in the door and said, Send Robert out and I said he's not here, and they said, We just saw him come out of that van. He's and I said, that was my mom he's not here. I'm just like this will probably be over when I get, there's, it's got to be a mistaken identity is something. I don't know why Detroit police are at my house and finally, pull in the driveway here. Pull up in my regular spot, hop out. By the time I close the door, the cars in the driveway are blocking me in. They parked this way, across my driveway as if I'm going to back out at something and try to take off. As soon as you shut the door they were right on him, and I was in here still because I had the girls, and I tried to tell Julia to wait here, and Rosie to wait here while I was gonna go outside and of course they didn't want to do that because there. They were confused. They could tell something was not right, and they were already starting to cuff him by the time we got out there. I was completely shocked and stunned to be arrested in broad daylight in front of my daughter in front of my wife in front of my neighbors, it was as it was. I don't even I can't really put it into words. It was. It was one of the most shocking things ever had happened to me. He could have been killed if he would have resisted, you know being a black man being arrested or having any encounter with the police could go badly. And you know that is something that we worry about in the black family when someone has an encounter with the police. Detectives came to see me. They took me out of the cell and to an interview room. And then it was like would you like to make a statement and I'm like, I don't even know what's going on. They guy says, if you don't sign off saying that you want to make a statement we can't give

you any information, so I sign to say I will make a statement. He laid the papers out on the table and at that time I was still guilty in their eyes, and the pictures didn't match. And they left him on the table and they looked at each other like, oops, it says this is used only for Investigative purposes and not to arrest. You know what daddy got arrested? Because they looked up, they made a mistake and I said I'll be right back, but I wasn't right back was I? We didn't think this could happen. We didn't think it was a thing even, even with me following the facial recognition news and it'd be how it was being used. I didn't ever expect the police to show up at our doorstep and arrest my husband. So I just feel like other people should know that it can happen and it did happen and it shouldn't happen. You just brought the picture with him, he could have looked it up and down he could have left, and say, Oh, my bad. My bad. I didn't mean to bother you. To have my girls out there and seeing that that was you know their first interaction with the police and for that to be what it was to them taking their dad, like, I'm always going to wonder now. How much did this affect them how, how much do they think about it, How much did they tell their friends, their teacher like, I don't know what's happening in the five year old mind or the two year old mind what they saw and how they're processing it. I don't I don't even know what it's worth, to not be arrested in front of your children, like, what's that worth to you. The damage is already done with my family, whatever however it affected my girls. I like to think my husband's a strong person but nobody should want to go sit in a detention center for. However, 30 hours, whatever it was, they need to get their evidence, right before they come and arresting anyone, black, white, brown, red just anyone. But they really need to get their evidence right as far as the false facial recognition, because someone could be locked up or hurt or killed that's innocent. They just pull the pictures from everywhere. Matching you up from anywhere, social media to the, I guess, if you in the background of a picture. And it's a clear picture of you, they can use that. And obviously it doesn't work.

Wendy: So public advocacy has slowed the use of facial recognition, until policy ethics and regulatory bodies can catch up with the technology. And while these issues are sexy and they much, and they're much more likely to make the news. There are actually more pervasive algorithmic technologies, machine learning and algorithms that are influencing our everyday lives that we're going to focus on next. So I'm going to turn it over to Chelsea.

Chelsea: Hi everybody I'm really happy to be here and glad to be talking to you about this. Oh, as we just saw from those examples, sometimes computing programs put into place with good intentions to have disastrous impact. And one of the biggest focuses that we have here with the institute is that these algorithms tend to impact people who have the least power to resist them. One effort of the Institute and what I've heard of our presentation here tonight is to educate you all as users of technology to notice and interrupt algorithmic bias, when you encounter it. So, we're going to talk a little bit about an example from a book by Kathy O'Neill. She's a mathematician, and I have the title of the book here "Weapons of mass destruction". If you're interested in this topic. This is a great first read on it, really engaging and what Cathy and you'll argue is partially in this book is that we can disarm these harmful models, by actually changing their objective, meaning we can break down the damage and unmask the bias. By changing what the algorithm is actually designed to do. Because the algorithm itself is not bad, right

there's no, there's no moral character to mathematics. But the impact as we saw can be extremely dense. So we're going to break into breakout rooms, and work together on an activity to disarm an algorithm that is causing problems in society. So, let me tell you about the first. First, you'll read the case study. Then as a group, you'll brainstorm a way you could solve the same problem with the same data to produce a different objective, and ideally this new objective that you will all come up with is one where bias is not operating. So again, it's important that the problem can be changed. The problem is the algorithm is set to solve, and the data cannot be changed, because these are often the same constraints that people who produce algorithms are working. Third, you'll agree on a new objective for this algorithm and you know justify your reasoning. And fourth, a selected group member or two to talk about what you discussed with the group, and one of each of the four of us will be in the breakout room with you and will be typing up your discussion in a slide. So you can just focus on having a conversation and we will be typing up your thoughts, you know, ok here's the case study that you are going to be tackling, here's the story in Washington DC, public schools are struggling. Only 50% of students go on to graduate after the ninth grade, an education reformer comes up with a plan to use an algorithm that will measure Teacher Effectiveness using a model that heavily ways energized test scores. Energize test scores. The teachers who score at the bottom 2% at the end of the school year, are fired. Some teachers begin correcting or cheating on their students exams for fear of termination. The DC school district is a majority minority district. 60% of students are black, and 20%, Hispanic, you can imagine the impact on students when teachers are removed, especially if their teachers that they trust. So, the problem that needs to be solved is struggling schools, the data that the algorithm is using is student grades and the demographics of those students, and the outcome, we want to change is firing underperforming teachers, and what's a different way, we can try to solve this problem with this data that has an impact that doesn't end up having bias outcomes. Okay, so that's the prompt and that's the plan, and the magic will occur that will send us all into our breakout rooms, and then we'll talk about it and each of the four of us will share the slide with you so you can see it again in case you want to read it more and see it more. Okay, so do you in the breakout room.

Chelsea: So I'd like to hear what you all discuss. So who is in Breakout Room 1, how did it go?

Wendy: We had such a great discussion that we failed to identify someone to talk about our ideas because we were just rapid fire coming up with all kinds of amazing ideas. So, I mean, essentially what we were trying to do is sort of think broadly about restructuring educational institutional goals, and to take the, the, the onus off of scoring teachers and looking more at providing supports within the school community, both with teachers and getting peer support but also with students and helping them sort of recalibrate, how they are evaluated. More broadly, as an objective.

Chelsea: Awesome. Yeah, and that makes a ton of sense I love these notes, how excellent y'all could do wonderful things in education, so thank you.All right, group 2.

Katie: Yes, so we ended up talking a lot about what group one talked about and what we could

do to support districts in need. In terms of funding and brought up advocates for the students. And really just, you know, recognizing that students are coming to these tests and to the table with different backgrounds. And it makes it a very uneven playing field when we're comparing them to an algorithm that determines that they should come out at a certain point, when they're taking these tests and really using this data to tell us where the funding is needed, and I think it was Dianne who mentioned that really when you're when you're looking at the outcomes of these scores that the districts that tend to receive more funding or the districts that perform better on these tests. The students perform better on the test and it should be flipped. So really, you know, looking at how we're better supporting districts where students, you know have multiple needs.

Chelsea: I love how your groups thinking, because the focus off simply features as like the only arbiters of outcomes, and focus on, you know, more systemic and ultimately more challenging problems but ones that should be tackled nonetheless. So, that's great. Thanks.

Katie: And I should add that Joe, who is in technology. Also this this last point here, mentioned that often when he's in meetings for work he's, you know, telling his executives that just because he can do something doesn't mean he should, so I think this is exactly the point that we're getting at is, you know, how do we critically analyze what we're able to do versus what we should do.

Chelsea: Absolutely. No, I think we all really love to have that plaque in our offices because that's a major. A major point, you know, just because an algorithm can be used doesn't necessarily mean it should be. And I think in a lot of institutions like education that are, you know, from between a million different polls right now a technological solution always seems appealing quick and efficient and edgy, but it isn't necessarily going to get the results that is desired so I'm really glad that you said that.Okay, group three.

Dianne: One of the key things that we kept talking about was really looking at all the other factors that that really contributed including ones that didn't seem to be paid attention to with only looking at the standardized test scores so we found ourselves looking at will put her some potential solutions. So if we rethink how we structure the teaching setting, think about you know doing team teaching it was. So how do we help this, the situation in other words, and then asking what contributed to student success besides the test scores and how would we. How can we help those things as a different way of thinking about what we're doing, not firing the bottom 2% but what if we gave them additional resources instead. In other words, so I think the other group. One of the other groups talked about this. How do you change the paradigm of what's going on? It’s also intending to support learning across abilities. A lot of our conversation focused on what additional support could be put in place, versus a focus on how do we fire underperforming underperforming teachers and all of this and then we also questioned whether whether we should be using the test at all because of the biases that it likely held, or if it should be used differently if at all.

Chelsea: Awesome. Wonderful. I am so like group four have really similar thoughts so I think we should just go to group four and we can kind of. So whoever did the typing for group four, that was me, really sorry. We're having a very vibrant conversation. There are many wonderful alumni educators in the room so I'm going to, you know, explain my bad job but if anyone from the group four wants to jump in. I know I talk a lot, but jump right in. So we also had a problem with this premise of like struggling like what does that really mean it feels like it's covering up a lot of deeper issues and like a struggle seems to really know reduce that and I'm not helpful way, and kind of like group one, one of the things we thought of is, maybe what this data is really pointing us to is where there are teachers who need more resources more support. Maybe we're really seeing is an overburden, and a few groups talk about that as well as how you could use use the same data to come to a different conclusion, instead of the conclusion being firing the bottom 2% of teachers to conclusion to see, I see where there is more need, and I will send resources to those places, that's that's one way you could change the outcome of. We also talked about AI, like this is really AI?, which I think is a wonderful question because. Very little is really AI. Most AI that you will actually encounter or see advertised is an approximation of it but not a secure AI that we often hear about in the media. So the short answer is, this is a version of AI but it is not pure AI is not a computer thinking, and there's actually a significant amount of person/ human labor involved in AI so that you want it to come take me where the classes you're welcome to, we talk more about stuff like this. And something that didn't make it to the slide that our group also discussed was similar concerns about the standardized testing really capturing real learning differences or is it just capturing differences of socioeconomic background. And, you know, what do we really care for the value of standardized testing anyway and obviously that's a question that public school educators genuinely have to contend with. But it didn't have a lot to unpack here but I love how everybody came up with a different way to use the same data that would not result in an outcome that makes things worse. Instead we came up with ideas of how we can make it better. So, that's wonderful.

Yousuf: So before we move on to Q & A, I wanted to just wrap everything up and let you all know that as sort of the heart of our institute. We're going to have three academic programs at our core so two programs are already out there and available to students. We have a BS and ethical data science, and a BS in business artificial intelligence and innovation, the ethical data science major is all about preparing students to go out there work with data, but do so in an ethical fashion, make sure that they're doing good in the world. The BS in business artificial intelligence and innovation is really training the next generation of business leaders to go out there, use artificial intelligence, with sort of a lens of corporate social responsibility. Also again making sure that we're using this technology for good. And we're currently in the middle of developing a BA degree in technical technology, artificial intelligence and society. Again, I think we see these folks as being the leaders of a just and equitable technology future for us all. So we're really excited to have Wendy and Chelsea here, teaching these courses, giving amazing experiences to our students. And we'll keep looking for all sorts of cool ways to engage alumni as well in all this work. Okay, I think, the floor is open for questions.

Sarah: So the first question we have is can you talk about bias and electronic health documentation and how this bias affects health care for people of color.

Chelsea: I can try to jump in, I don't know a ton about bias in healthcare data in specifically in how emerging technology and AI is being used, but I do know about existing racial and socioeconomic disparities and healthcare and a general insight, you can apply to most areas when it comes to AI when it comes to emerging technologies, is it the bias already exists as it does in health. You would have to be very cautious to implement any kind of automated or AI, machine learning based system in a setting where bias already exists as it does in health. So, Wendy is nodding. And maybe he has more specific info about this, you know these overlaps. But I only know from a medical sociology perspective how these kinds of disparities did so. But those health care disparities are very real and not solved in my opinion, by automating anything. And, yeah, so maybe I'll pass the torch but that is a great question and an area of real concern.

Wendy: And I think I will just add to Chelsea's remarks, is a lot of what happens here that we have a lot of difficulty, particularly in my field and information science is making the data representative. And, you know, we oftentimes have a lot of difficulty understanding what we need to know from a service provision perspective, versus what we can actually add to the computing system, and that one of the really great challenges with what we're trying to do at Naz is to educate people who may not necessarily ever want to touch computer programming, ever, and I get that, but they will have a literacy about technology so that when they're at the table, they can start raising these kinds of issues, what are we thinking about in terms of how do we collect information from this particular demographic or people or, why are we so insistent on having apps and websites for registration for healthcare. When we have a fair number of people in our community who can't access that. So those are the kinds of things that we're actually going to be talking about and, but it's a fabulous question and it's something that if you can solve it. You are going to be a hero, tricky problem to solve.

Dianne: And just to add to when these point there for a second. If any of you have tried to register to get the COVID vaccine. The amount of knowledge, one needs and the amount of access I had like my spouse and I had like five devices each open, trying to get into the system to actually register to get a vaccination, and you think about what that meant I mean we are exceptionally privileged that we had these devices that we could do this. But Wendy's points about access I think are really, really important and the choice to do it in the way that we have done it. While we can look at the efficiency and the way to do it. It also means that you're closing doors to certain people, the same way with a medical record so we have electronic medical records where people who have access and the ability to develop the knowledge base to access the records that way, can benefit from it, but then everyone starts assuming that everyone can access it that way, and all the sudden you're cutting out health communication from people who, who don't have that level of access to technology and so you created another level of bias not so much in the algorithm itself, but in the assumption about the access to the data, which is another level so not the algorithm, but the access that becomes a part of the issue.

Yousuf: One last thing I, one of the common themes that you see whenever you're looking at AI and machine learning in particular is that as Wendy sort of described machine learning is all about finding patterns in big data. And in general, the data that we feed into our computers to find these patterns is historical and all sorts of issues around race, class, gender, are encoded in the data that we feed into these models. And then when we let them run amok with those they tend to amplify any sort of disparity that is already there, you know it's like putting racism or sexism on steroids. And so we already know there's robust data that shows that doctors treat patients quite differently depending on race, gender, class, etc. And so when we remove doctors and start automating it with AI. You know those problems, only get worse. And, you know, these are the sorts of things that we have to be careful of. I'm going to keep going back to Joe's point, just because we can do it doesn't mean that we should.

Sarah: Alrighty, I think we're set to move on to another question. And here it is, for those working in digital digital marketing today, how can we minimize the black box effects and be more equitable in our advertising.

Yousuf: That's a great question. So, I don't know that I have an answer yet but I'm going to relate a little story recently of a class I visited so I was in one of our digital marketing classes here at Naz, and the students were looking at some of the data collected from the Nazareth website you know students coming to visit prospective students, etc. And we were breaking down the data to see what patterns we could find in it before we, you know, automated things. And at one point the students recognized that there was gender information from visitors to the website and their question was, how do we know the gender of somebody that visits the Nazareth website. And the answer is in, you know, you probably have seen as you visit new websites little releasing that you accept the cookies. And when you do that, you're transmitting all the data that your browser has on you to the company or whatever organization, it is that you're visiting the website of. So I think one thing just to note is that we as consumers and citizens need to know what is actually out there where our data is going and what we can do to make sure that it's being used in ways that we're, you know okay with.

Dianne: I want to go right ahead. No, you go Wendy.

Wendy: I just was going to say I love this question because I had a student in the fall, who did a study on Facebook ads and maybe she might want to talk about that a little bit, Sarah.

Sarah: Okay. Um, yeah. So, one of the big takeaways that, I had Wendy, as my professor for. I don't remember what the actual, it was special topics in math but it was really about data and AI and machine learning. So, one of the big takeaways for me was like targeting ads like ads on Facebook or Instagram or social media, where you're like, when I was like talking about that specific brand earlier today or gi was like just shopping in that store last week and now there's an ad for that store on my Facebook feed so I decided to do a little experiment. And it was kind of silly but I would talk into my phone for five or 10 minutes every day, yelling basically like cats,

cat food Petco over and over like everything that I can think of related to cats. See if I could get ads that we just have cats on my Facebook feed. And then I would write down like how long it took for the ads to show up wish I did a lot of the time, there were ads related to what I was talking to it really creepy and I knew is going to happen because, like I had allowed Facebook access to my microphone and all these other things and that was another thing that we talked about one day in class like if you look through all the apps in your phone and what they have access to it's crazy you'll be like, why does my calendar out need access to my microphone or my photos or contacts or. It's weird like all the things that these apps have access to and you might not realize it, but yeah that was my, my final project was just a little experiment on Facebook to see how targeted the ads were and the results are kind of crazy, for sure.

Dianne: And so, Sarah has just recounted for us the kind of digital marketing that is the norm for all of us today sometimes in the ways that we don't even know that it's happening, Sarah clearly does now because she's got the whole cat thing going on. But, we know that this is happening so when you think about the black box when we talk about the black box it is that the algorithm itself. The set of instructions as Wendy talked about before, are learning from the data, information and so that becomes a black box that we're no longer we meaning the people who are programming, a computer, or no longer kind of in control because it learns new things from all the data by putting the data together in a certain way and saying oh I see these patterns, and then it starts looking for a different set of patterns we didn't tell it to look for the new set it figured out it needed to look for the new set because of the patterns at found. And so the hard part about that is that if you're going to allow the patterns to help you learn about how to find new patterns, then you lose complete control so a lot of the ethical work in an artificial intelligence is trying to look at, well how do we do this how do we each step of implementing artificial intelligence, ask ethical questions How do we create guidelines, how do we create norms, how do we create the computer equivalent of the UN Human Rights set of criteria, so that at each point we're asking what is the impact that this might have, depending on how its developed and so there's no great answer or agreement on how you help the black box but there's a lot of insight and work being done on figuring out how do we help create some guidelines in this, and how do we keep one set of patterns from Indiana, where we've seen several eyes go which is that they end up with the chat bot that makes racist comments, or you end up with a hiring profile that doesn't hire women. And so we want to stop those things from happening so while Sarah is working with cats, the cats translate into a whole bunch of other things that end up with three inscribed in the bias once again you don't have making it happen over and over and over again. So it's a real challenge and that's why it's a great question that you've asked.

Sarah: All right, I think we're ready to move on to the next question, which is, can iris reading identification be racist?

Chelsea: I'm googling this because I don't know what it is…

Sarah: Iris, like the part of your, you know how like some laptops have the recognition thing. I

think that's what she means.

Chelsea: Okay, thank you. I was like, Iris? If this is the name of a computer program that I'm not, I'm not familiar with? I'm not sure of the impact of that but I am aware of the problems like we saw with the video showing how face facial recognition technology is unable to detect her face because of the color of her skin. This is an enormous problem. And it shows up in a lot of different programs. One of the most troubling in my opinion is automated test proctoring services that require students to have their cameras on, and students of color routinely report that the cameras cannot detect their face, they have to shine. Like many lights upon their face in order for them to be visible to these programs that is being equitable, not only does it take them more time to take the assessment, some students find that the software is unable to recognize them there, they have to reschedule their assessment or their tests, all of this is unequitable it creates barriers to learning that do not exist for white students. So, I'm not, I haven't heard about eye color and I haven't haven't seen data on that but I have some data on problems with facial recognition technology, especially dark skin tones. And it's extremely troubling and really problematic.

Wendy: And I can just jump in for a second, again, a lot of this is about context. So if you are using iris identification and our detection software. Are you using it to identify and verify the identity of an individual because you are in a highly secure situation where someone needs to be able to provide, you know, an identification card maybe they need to have a you know a code to implement into a keypad to, again, identify themselves as someone who has access to a particular place. It's a very different context when you're saying that you want to do Iris scanning identification for a humanitarian context to allow someone to get food or to allow someone to get shelter or to, I don't know, you know, verify someone's eligibility for public assistance or even things like that so context matters a lot in this, and, and again those contexts can be racist, the technology itself generally is not. It's just the context in which it's being used.

Sarah: Next question is are people programming these computers to discriminate?

Joe: Do I get to answer this, can everyone hear me?

Chelsea: Yea I'd like to hear your thoughts.

Joe: I think I guess it depends on the context. Um, so for those that don't know I work in cyber security specifically. So I do a lot of different interesting work. The most people aren't normally familiar with but I can think in some sense know probably like some things like advertising, you're probably just taking it and looking at the data and saying hey we're going to build a campaign maybe for like a film release or something like that and and go towards it and seems pretty ethical. On the other hand, you could have some, like a defense contractor or these other I know Pell and Tier, got a lot of fire for this when they reselling equipment to CIA and border control and things like that to target people. And so I guess it really depends on the product and the output, that's going in some people will know that they are actually contributing to these

things. Same as Google was running a pretty big program, I think a lot of staff that were kind of raised contentions with I can't remember the name some being on the spot but I think there are some projects in, in, in activities people go into knowing what they're contributing to and the other hand, and sometimes it can be misconstrued versus other times where it's more of a direct outcome like I'm doing advertising for a film and whatnot but my take, I don't know if that's the right answer so someone else can definitely answer it on that.

Chelsea: Thank you, I appreciate your perspective.

Yousuf: I'll say, I'll say a couple things so one of the things to note is that in lots of ways the algorithms that we're talking here are not explicitly programmed by a human, again it's based on machine learning in which the computer itself has sort of taken in the data and found patterns in it that we're discriminatory in the first place. So I think a lot of the time, the intent is not to be racist or sexist or what have you. I'm sure that there are cases but even when there are not. I think it comes down to how the algorithm is generated, what data has been put into it and how we have been careful about that. And then secondly the context is so important. Quite often we're using these technologies to try to make our lives easier. But the question is what does it actually do in practice so I don't know if any of you have gone through. Do you remember in the olden days when we used to travel. So, just before the pandemic and I haven't actually been on a plane in more than a year. Airport started instituting some facial recognition tech to get you through security faster. And it sounds like a great idea. And again, as you see, you know, these, these facial recognition algorithms do really well with white male faces and do really poorly with black female faces. So when we run one of these sorts of tech, as we run one of these things at an airport, it's going to work really well and make my life super convenient, which by the way I look a lot like most of the people that program these algorithms. But then instead of, you know, not making life better necessarily have a black woman that runs through the same checkpoint. She may actually be detained, held up, made her trip so much longer and do real harm. So I think this is where the context is key. There are situations in which these technologies are making the lives of some people much better. But those same technologies are making the lives of other people much worse.

Dianne: And just add one more comment to that back to the issue of intent. If you think about the Joy video (Gender Shades), that's related to what Yousuf is just saying it is likely not the case that the people working on the facial recognition went, Oh, I'd like to make this work on white males and not work on black females, is because the lack of black females or women just period or people of color in general meant that they didn't think to ask the question, and therefore by not asking the question they created a system that as it continued to learn more from a set of assumptions that weren't accurate. It got worse and worse. I mean, you keep hearing that say that but that's, that's the biggest problem is that what starts out as a small seed of something, you know, maybe not accurate becomes a context for bias, whether intended initially or not, as Yousuf was saying.

Wendy: Right and that's the real paradigm here is that we like to protect our data like to keep

our own lives, our purchases online, our social media, the things that we do, the things that we read, we like to keep that private. But in order for the machine to learn, it has to have a variety of data. So how do we then make that trade off that machine learning is going to get data that is more representative of the community at large, while also being able to protect our privacy.

Dianne: There's the rub.

Sarah: Okay, ready for the next one? Great. Here's the next question. Could you develop a study which asks participants to see where they could change the algorithm to machine Intel to AI path, and how to change it so that different results would exist? Do you want me to say it again? It's a long one. Could you develop a study which asks participants to see where they could change the algorithm to machine Intel to AI path, and how to change it so that different results would exist.

Dianne: It sounds like it's talking about almost some kind of practice algorithm like could there be ways that if you tweaked it this way you could you could make changes to it to see what the outcome would, how to make the outcome be different and there are actually algorithm testers that exists that allow you to, you know, put in this thing and see what the outcome is and and kind of keep testing that one of the challenges with machine learning is that the input of the data and the identification of the patterns happens over time with more and more and more data so that running it through to test all the possible outcomes. It is an infinite kind of process. Right, so you can't do some tests on that exactly like you're, like, like it's been described here and that's one of the things that are important, that in the process of thinking about the ethics of things looking at ethics that every spot so you can look at potential outcomes. But at some point you no longer have the control of it and that's when it starts running a muck, as it were, I'm sure Wendy can say more about some of that.

Wendy: Yeah, I think, you know, the other thing is this is a fascinating question, and what's difficult is that we as consumers as citizens of people of the earth, don't always know that we're in the midst of an algorithm or machine learning, or AI and and Chelsea's comments earlier were spot on. A lot of AI is what we call Wizard of Oz, it's not really AI, it's just sort of the man behind the curtain kind of moving some bells and whistles and doing some things and they look kind of cool but it's not really a I at all. It's just fairly advanced programming to guess what you might be thinking or wanting. So this is called general artificial intelligence and there's other sort of strains of AI, but it's what I kind of pull out of this question which I think is maybe taking us down a little bit of a different road but i think it's it's an interesting thing to think about is how you opt out as a human being to say, I don't want to be a part of this. I don't want my information being stored, I don't want my data being brokered to be sold for things for products. I don't want to be part of the surveillance world. How do I click the box that says no cookies and no tracking. And that's a really hard thing to do right now because a lot of what we do online these days has to do with trying to provide you convenience, and that convenience comes at a cost, and that cost is our data and our privacy.

Sarah: Alrighty, so this is the last question that we have submitted so far and after the panelists answer this one. I'll ask and if anyone has any last minute questions so here's the last question we have, how in depth will students learn about topics like how to handle large data sets and privacy in an ethical manner? So this relates to like the new classes and majors here, at Naz.

Dianne: This seems like a faculty question to me.

Wendy: I can give a very concrete answer to that if you like. So one of the things that we're doing in the beginning programming class is. I have students using pop culture data to analyze large datasets. So we are looking at a data set of 23,420 characters from the Marvel and DC universes. And that equates to about 440,000 data points, and these are our students who have never programmed before in their lives and they are terrified when they walk into the class and I lay this out for them. And by the end of the semester, every one of them has produced outstanding data analysis on the representation of these characters, their back stories information about them doing light statistical analysis on this large data set, and they're super excited about it. And this semester, you know, teaching and beginning AI and machine learning courses, and we are actually pulling data from Twitter. So we're pulling live streams of data about movies, this time. And looking at how people are talking about movies, what kinds of things are popular are there correlations between the number of hashtags that are on a tweet and how many people are sharing that hashtag and talking about the ethical aspects of how would you feel if you were on Twitter, talking about your favorite WandaVision episode, and now all of a sudden you find out that there's this classroom full of people who are looking at your comments and looking at your conversation. And how are we thinking about who's on Twitter and who's represented on Twitter and is this really how people in the world, think about these particular movies, or is this a subset of people and we need to be careful about how we're presenting this data in this analysis into the world. So there's a couple of concrete examples of how we're doing that right now. And the students are just having a blast and I'm really excited to be teaching these courses.

Sarah: All right, if anyone has a last minute question you should be able to unmute yourself.

Margaret: Okay, can I may I ask. I've been writing computer code since 1956. That's all. When the very early stages. And I think there's something, you know, before we get as complicated as identifying a face, which requires lighting, angles. It has a 5% error rate, at least I don't think it is ready for big time and I think it's terrible to use it for police work, but I want to get back to something very simple, which is what we will use where Microsoft Office products.
They can't even tell you you got a misspelling or grammar yet today, you know, I mean, I talked to some of my young enough nephews you know they don't, they've learned, we call diagramming sentences so we know parts of speech, we look at texts today from people. It's not, it's not well formed senses. So where are we on that part of it. The other thing is, which I think should be a big improvement is you get on the telephone and you want to get an answer to a technological problem, and you're frustrated. I want to talk to a real person even though they're not the brightest person in the world, you get your answer. So I find these are AI

problems that should be on the forefront. I'll let you guys discuss it. Okay.

Wendy: Margaret I'm so glad you asked that question because we were just having this conversation in class today about Clippy, the little paperclip icon in Microsoft Word that would give you suggestions. And the students remembered this actually they were like, oh yeah I remember I was in elementary school and I'm like no I feel so old. And I had my professor at Stanford was actually on the team at Microsoft who invented Clippy and he said he would never be able to redeem himself in his life for foisting Clippy on all this was a very early chat bot. I mean essentially this is what this little icon on your screen was intended to do is to have a dialogue with you about what you were trying to do, and it was very unsophisticated and oftentimes gave you terrible advice about what you were actually trying to accomplish. And we're not that much better about it and actually that's one of the projects that we're going to be doing in my AI class is building a little movie chat bot. So we're going to digest tons of information from the IMDb database that's free and open source and access, accessible to anyone. And we're going to teach the students ethically, how would you build a chat bot. That would be able to engage with you and your friends about what movie you should go watch.

Margaret: That's a great idea. I mean, it's so basic it's been worked on for decades and haven't we haven't still can't get the right spelling for words.

Dianne: Thanks. Margaret.

Chelsea: I think also related to the question Margaret is that we don't always perceive the technology products we want. And the technology products that are most well distributed through, you know, the phone and computer using society is the most profitable one not necessarily the best one, or the one I wanted it or the one I really hope for. So it's not a question of, like, whether that can be accomplished because there's no doubt it could be, it's a question of whether there is a will. And much of. I mean people find themselves in small groups or making all kinds of unbelievably cool useful things, but they don't have the platform that Google has. So I think as long as our technologies are controlled by corporate interests, we're going to receive very nice things and things that work well but maybe it's not exactly what we asked for. So something I talk about a lot in my classes with my students is, what do you want from technology with the new one from AI, and we should push for it. Because, you know, if corporate interest is left to its own devices, I'm not sure I don't know that we'll ever get those, you know, effective texting tools that we deserve.

Wendy: What I teach, and I think Chelsea also talks about quite a bit, is this notion of computational thinking, which is just breaking things down into manageable parts, and using that as you know it's just one more tool in the toolbox with critical thinking feminist thinking, you know, other ways of looking at the world. Again, Sarah's laughing because I made her do this as well. But it's actually, I think quite helpful. I know Sarah could probably tell you more about the student perspective whether she found it helpful, but I do find that students are very interested in understanding the world in different ways and that's one way to do that.

Sarah: Alrighty, unfortunately that's all we have time for today. Thank you again to our panelists for the wonderful look into the world of AI and machine learning. Thank you so much to all of you for joining us this Thursday evening. Please look forward to a recording of this discussion in the coming weeks and I hope you all have a wonderful rest of your night.

Yousuf: Thank you all for a wonderful discussion.
Dianne: Thank you all for joining us.
Chelsea: Thanks everybody.