Center for Humane Technology | Your Undivided Attention Podcast
Bonus: Coded Bias

Tristan Harris: Welcome to Your Undivided Attention. Today our guest is Shalini Kantayya and she is the director of the new film Coded Bias, coming out on Netflix on April 5th. We actually originally saw Shalini's film, Coded Bias, at the same Sundance Film Festival that The Social Dilemma premiered at and we're just excited to have her on to talk about her incredibly important film. Shalini, what is Coded Bias about and what had you decide to make this film?

Shalini Kantayya: Oh, well, first of all, thanks so much for having me. It's such an honor to be in conversation around these issues. Coded Bias follows the work of Joy Buolamwini, who's an MIT researcher, and she stumbles upon the fact that facial recognition doesn't see dark faces or women accurately and stumbles down the rabbit hole of the ways in which algorithms, machine learning, AI, is increasingly becoming a gatekeeper of opportunity, deciding such important things as who gets a job, who gets what quality of health care, what communities get undue police scrutiny, sometimes even how long a prison sentence someone may serve. These same systems that we're trusting so implicitly have not been vetted for racial bias or for gender bias or that they won't discriminate or have unintended consequences.

Shalini Kantayya: And, these systems are black boxes that we can't question. Oftentimes, we don't even know when we've been denied an opportunity because of this kind of automated gatekeeping. And, that's when I really realized that we could essentially roll back 50 years of civil rights advances in the name of these machines being neutral when they're not. Everything that we love in our democracy is being transformed by AI. Fair housing, fair employment, access to information, so many things. And, I think the urgency of that is kind of what inspired me to make the film and I really believe this is where civil rights gets fought in the 21st century.

Aza Raskin: You know, in Silicon Valley, there's this fixation on this singularity, the place where technology gets better than the things that human beings are best at, and that's the place we should focus all of our attention. What that misses, and I'm really hearing you say, is there are many ways that technology starts to affect us and undermine the places that we aren't looking, and that's the real danger of technology, is this kind of invisible rewriting of the rules of our society, the menus from which our lives are being chosen.

Aza Raskin: I know just a couple of the examples. There's a now fairly famous example of Obama, like a picture of his face getting blurred out and then the AI being asked, please reconstruct this face, and it doesn't return a picture of Obama. It returns a white face. If you take pictures of women and you ask an AI to auto complete, you just show the top half of the face and a little bit of the shoulders and say, please complete the image, an AI will show, auto complete the woman into a bikini. And, if you have a man's face at the top, it'll auto complete the man into a business suit.

Aza Raskin: And, there are all of these invisible ways in which these systems are making decisions about us that I think the film does such a good job of highlighting.

Shalini Kantayya: Oh, absolutely. And, I think that we think of technology as our gods and I think they're more like our children, flawed reflections of ourselves and even the things that we don't want to pass on sometimes. And, I think the thing that I have grown in compassion for is that bias is not something that's in a few bad people. It's actually

an inherent human condition that we all have and is often unconscious to us. And, the scary part of when that gets encoded in technology and what was so alarming to me about Joy's discovery of racial bias in facial recognition technology... And, she's just trying to get an art project to work... Was that this was not a technology that was sitting on a shelf somewhere. This was technology that was actively being sold to the FBI, actively being sold to ICE or immigration officials, actively being sold to law enforcement, with no one that we elected, no one that represents we the people, giving any kind of oversight to that.

Shalini Kantayya: There are no laws that would make this information transparent to me here in the U.S., so I actually had to go to the U.K. It was Silkie Carlo, who's also featured in the film, and they found that with police use of facial recognition, 85% of the people being stopped were misidentified. And, I'm using the most conservative statistics and I think I sort of almost never have recovered from seeing a 14-year-old child who is stopped by a five plain-clothes police officers, never asked for his ID, fingerprinted, and doesn't understand why this is happening to him. It's only because there was a human rights observer there that explained to him, "You've been wrongly identified by facial recognition."

Shalini Kantayya: And, I think it's those moments in the making this film where I really see like, "Oh, that's the moment where technology oversteps on civil rights. There it is. That's the line being crossed." And, I think that was most frightening to me.

Tristan Harris: What was fascinating in hearing your example was the idea that Joy was discovering flaws in these systems that were already working in police departments or in the FBI, that this was after the fact discoveries of consequential biases. I mean, to not even recognize... You said, is it 85% of people were misidentified in the other example you gave?

Shalini Kantayya: Yes, absolutely. In the U.K., a study by Big Brother Watch U.K. And, those are conservative statistics. I mean, it's upwards of 90, some of their statistics around the misidentifications.

Tristan Harris: Yeah. Well, and so, the thing that this makes me think of is a similarity between your work and ours, is that we could have these systems that are right underneath our noses, that are already running our lives, whether it's a Facebook algorithm that's already determining the news feeds that we're seeing or TikTok already sort of ranking bad content for sexual predators or things like this, that we don't even realize until after the fact. And, the idea that we can only tinker with it after the fact and why wouldn't we have discovered some of these problems upfront. What does that say about the production processes that govern what technology gets out there?

Tristan Harris: Imagine if someone were to say, "Hey, I'm going to give you a robotic heart and then a robotic lungs and then a robotic liver. And, I'm only going to test afterwards if it's somehow wrong in some highly consequential way." With FDA or with drugs, we have a system to vet drugs and their safety upfront before we ship technology. Now, of course, people are going to balk at that because they're going to say, "Well, how else are we going to have an innovative society that's shipping technology really quickly?". But, software can be more damaging or more consequential than drugs

and we're seeing places around the world in the Facebook and social media cases where genocides are getting amplified and we're not testing to make sure that it's not doing that. In fact, we're optimizing for growth and distribution faster than we're optimizing for safety. And, that just seems like a recipe for disaster that's represented in both the areas that we're looking at.

Shalini Kantayya: Absolutely. I don't think we've really examined the fact that democracies are picking up the tools of authoritarian states with no democratic rules in place to protect people from its impacts and I think we're missing the point of humanity because in the making of this film, I've thought a lot of it, what it means to be human, and is the goal of human civilization to go as fast as possible and to be as efficient as possible? And, I've thought a lot about what human intelligence is and I've decided it has something to do with empathy and something with our ability to have compassion, and I think that we're living in an age where it's almost like a world with the automobile with no seat belts and no car seat for your baby or pharmaceuticals don't have a label of how much you should use. It's just a lawless Wild West and we don't have any health and safety standards.

Shalini Kantayya: I'm often asked, "Don't you believe there are good uses of technology?". And, I'm like, "I love technology." And, I think this sort of idea of an FDA is the idea that we should have certain health and safety standards for technology. And, the scary part is that this stuff has real impacts for civil rights, for people's lives, and I've seen it in the making of this film, whether you're talking about a school teacher like Daniel Santos, who, if you stood 10 feet away from this teacher, you would know what a passionate, committed, dedicated teacher he is. And yet, an arbitrary algorithm says that he's a bad school teacher and we just give it blind faith and he has to defend himself. He has 10 years of evaluations that say he's a good teacher, but against one algorithm, he has to defend himself.

Shalini Kantayya: And, I think right now, we have a system where we deploy these things at scale and then they hurt people and then we pull them back, and I really think that there should be some sort of process of ethicists and policy makers and other people in the room before these technologies are deployed at scale.

Aza Raskin: Mm-hmm. Just like there are environmental impact reports before we deploy technology that affects the ecosystem from which we derive our life support, I think we absolutely need societal impact reports, right? As we put technology into the field that changes the environment from which we draw support or... Yeah. Social environment. And, it's not like the harms can actually be walked back because once you start down this path, you create a new set of conditions. Like, you harm the teachers. They get fired. Then, the next time they go to get a job, well, there's already a bad mark on their resume for having been fired by this AI. And so, it's a cascading set of harms that, unless we get out in front of right now, we're just going to continue to live more and more and more in this sort of the detritus of these poor decisions.

Shalini Kantayya: Absolutely, especially something like facial recognition. I mean, there's one case in the U.S., the only one we know about because there are no laws that make it transparent, and this Detroit man was arrested in front of his neighbors and his family, held for 30 hours in a cell, and never asked for his license. And, in spite of

that wrongful arrest, the Detroit Police Department continues to use that technology. And, that's the kind of stakes that we're dealing with and it's kind of astounding to me that three black women scientists who were all graduate students at the time somehow found bias in commercially available technologies that Amazon, IBM and Microsoft missed. That's astounding to me and I think that sort of speaks to what I would call an inclusion crisis in Silicon Valley.

Shalini Kantayya: When you have less than 14% women be AI developers, I think half the genius of the room is missing and I think oftentimes, we think about inclusion as sort of the social service announcement, the thing that's good for the pictures. But when you're trying to control for bias as an innate human condition that we all have and something that we have to be perpetually vigilant about when we're building technologies, having inclusive teams and not just racial and gender diversity, but maybe not everyone goes to Stanford. Maybe some people come from San Jose State. Maybe some people's first language isn't English. I feel like having those kinds of inclusive teams are really important.

Shalini Kantayya: The other thing I want to say is I know that you speak to a lot of technologists and I'm concerned that I see this pattern of independent science that highlights bias being dismissed and attacked at these companies. And, I'm speaking... You know, Joy's work was first dismissed. The firing of Dr. Timnit Gebru at Google certainly. And, I was very heartened to see 2,500 of her coworkers do a virtual walkout and resignations that followed. But what I've seen as a recipe for how change works is that we need brave science unencumbered by corporate interests. We need ethical scientists that can speak the truth and a culture that encourages dissent within these companies so these voices can be heard and these technologies can be made more ethical and fair.

Shalini Kantayya: Something happened after Coded Bias was released at Sundance that I thought was remarkable, that I never dreamed would happen, which is that IBM said that they would get out of the facial recognition game. They're not researching it. They're not deploying it. They disrupted their whole business model. Microsoft said that they won't sell it to law enforcement. And, Amazon said they would put a one-year pause on sale of facial recognition to law enforcement. We're good for, like, two more months there. But that was sea change that I never thought was possible and I think that happened because of these brave scientists in Coded Bias. Because of science communication, the public understanding, because I think AI literacy... I can't underestimate how important that is to the public.

Shalini Kantayya: But I also think it was because engaged people acted on that science and we had the largest movement for civil rights and equality that we've seen in 50 years and people making those connections between racially biased, invasive surveillance technology in the hands of police and the communities that are hurt and brutalized the most and have the most to lose. And, I think that the more that we can encourage brave science, science communication and activism based on science, I think that we have a moon shot moment to call for greater ethics in these technologies that will define the future.

Aza Raskin: I'm curious. What kinds of pushback do you get against the film when you screen it at big tech companies or elsewhere?

Shalini Kantayya: I think the most common thing is that... That it'll somehow kill innovation, and I think that's actually the opposite. I think that when you change the business model and you create health and safety standards, it unleashes a new type of innovation. And, I wonder sometimes what it would mean to design technology not for efficiency, but around the inherent value of every human being, if that's even possible. And, that could mean that we need a slower approach to technology and I know that's not what technologists want to hear. But I think often, too, that when I talk about bias in artificial intelligence, especially to technologists, there's often this impetus to say, "It was just the data set. It was just garbage in, garbage out. We'll just fix the data set and then everything will be fine."

Shalini Kantayya: And, I think that I really want to resist that because it's not about building the perfect algorithm. It's about building a more humane society and changing our entire way of what the technology is doing and trying to make it in service of our humanity instead of us being, for better words, like, enslaved by our technology to its click bait and to its bells and whistles, and I think that there's a whole different way that these systems can work that we haven't even begun to explore. I don't think here in the U.S. we even know what AI for public good could look like.

Tristan Harris: You're speaking our language on so many levels. I completely agree. I mean, there's certainly... I want to make sure we credit. There's a lot of people, I think, who are working on public interest technology for a while, but I do think there's an imagination gap and one of the things you talked about earlier is, in part, this is due to an inclusion crisis, that the other minds, the other possibilities, are not present. And, could you speak to who some of those people are? And, you featured many of them in your film.

Shalini Kantayya: When I was talking with Safiya Umoja Noble, author of *Algorithms of Oppression*, she basically talked to me about a whole different way the internet could work as an artistic tool and maybe there could be some transparency around it and how you could maybe see the funding backers, that there might be a way that you could select, okay, I want news sources here that are from vetted resources. Maybe over here, you are on the commercial sort of section of the internet. And, she talked about the ways in which that process might be more transparent to us.

Shalini Kantayya: I think Zeynep Tufekci is one of the smartest people I've spoken to and she's just brilliant about talking about how we just love the internet. We just hate the invasive surveillance of it. And, is there a way where we can have some data rights and bringing that balance of power?

Shalini Kantayya: Coded Bias centers the voices of women and people of color because I think this is a community of untapped genius within tech. To change tech... I think there are seven or eight Ph.D.'s in the film, all incredibly astute data scientists and mathematicians. But they were also women, people of color, LGBTQ, religious minorities. They had some identity that was not centered and I think because they had that view from the margins, they could shine a light on bias in technologies that Silicon Valley missed and I think that that's really important, that we need each other to shine a light on into each other's biases, particularly when you're developing technologies that you're deploying on the world.

Shalini Kantayya: I know there is a lot of conscientious, well-meaning, brilliant people who work within these technology companies and it's just my hope that they will obey their sense of true north, their own moral compass, that when they're in rooms where they feel that something isn't right, that they'll actually speak up in spite of what may happen, to have that sort of bravery. And, it's also my hope that when you hear someone like Dr. Timnit Gebru at Google speaking out, that people will make space to listen.

Shalini Kantayya: We need people like Joy Buolamwini and Cathy O'Neil and Timnit Gebru in the rooms where these decisions are made. I mean, three black women, data graduate students, essentially changed the policies of IBM and Microsoft and Amazon, and I feel like that's a rally cry for more inclusion in the kinds of voices that are making these decisions that impact the world.

Shalini Kantayya: And, I actually have been shocked that tech companies have been so receptive to the film. I was lucky, as Tristan said, to screen at Sundance, where we actually had the rare opportunity to see the film with an audience, and someone who worked at Google came up to me afterwards and said, "We've been having this conversation with ourselves and you made a conversation we can have together." And, it's my hope that that will happen.

Tristan Harris: I love that and I was going to ask you about the impact that you've seen since the film and you just shared so much and I actually want to offer that for listeners because while some of the topics that we can talk about on this podcast in general and in the work that we're talking about today can feel really bleak and, oh my God, how do we ever change these systems? And, oh my God, isn't the massive economic interest at play going to suppress change? And, it's true that they do actually, as you said in the couple of examples that you gave. But one of the things that makes me also optimistic and I didn't know about some of the screenings you've done at tech companies, is that something that actually our listeners can also do, is to screen this film in the places of power and to create a shared conversational object.

Tristan Harris: And, I... I imagine, as for you, as it has done, I think, for us, there was actually many people inside of technology companies who had a lot of the concerns that we've raised also in our work and didn't actually feel like there was an avenue or a safe way to bring it up. In fact, it was dangerous to talk about it. And, one of the interesting things that a film can do, it seems to me, is to broker space for that missing conversation. And so, I'm just so excited and hopeful hearing you talk about that because it makes me think films really can make this difference. This is not just to create an hour and a half experience and then people go back to their day jobs and nice thing. But really, really change.

Shalini Kantayya: Absolutely. The making of Coded Bias itself and I think why I make documentaries is that it really reminds me that everyday people can make a difference and that not all superheroes wear capes. And, I've seen that perpetually in the making of this film. If you told me three years ago when I started making Coded Bias that three of the largest tech companies in the world would change their policy of selling facial recognition to law enforcement by their own volition, I would never have believed you. And, I've seen time and time again, whether you're talking about Daniel Santos, the teacher who challenged an algorithm scoring, the value-added model which is

still being challenged all over the country to score teachers, or Tranae Moran, who not only organized with her friends and her neighbors to keep her landlord from putting facial recognition in her building, but also inspired the first legislation in the state of New York that would protect other housing residents to do the same.

Shalini Kantayya: Fast Company called Coded Bias *An Inconvenient Truth* for algorithms, for big tech, and I hope that it will be that kind of film that translates science to the public so that we can pass policy. Let's hope it goes a little faster than climate change. But it's really my hope. I feel like with films like Social Dilemma and The Great Hack and Coded Bias, that we are sparking a conversation and it's my hope that it will lead to a culture of change. When you sit in the dark and you empathize with a character and you go on a journey, you come to care about something and to me, that spark of empathy is how social change happens and films are a place, to me, where they give a safe space where we can have this kind of civic dialogue, where we can have safe discussions with people who think differently. And so, it's my hope that that's what the film will spark.

Shalini Kantayya: And, I am so grateful to be working in coalition with the Center for Humane Technology and The Social Dilemma and so many others like the ACLU, The Electronic Frontier Foundation, The Algorithmic Justice League, Mijente, so many incredible organizations that are working for change. And, you can go to the codedbias.com Take Action page and there's so many wonderful organizations that are doing work in the field. There is further reading. All of the authors from the film, all of their work is listed on that site. There's an action page and a discussion guide if you want to host a screening.

Shalini Kantayya: And so, it's my hope that people will use the film at their companies, at their dinner tables, in their schools, to spark a new conversation.

Tristan Harris: I really just hope everyone leaves here and watches the Coded Bias and just thank you so much for coming on.

Aza Raskin: Thank you so much.

Shalini Kantayya: Thank you so much for having me. Such a pleasure. And, I hope everyone watches it. Coded Bias is on Netflix April 5th.

Aza Raskin: Your Undivided Attention is produced by The Center for Humane Technology. Our executive producer is Dan Kedmey and our associate producer is Natalie Jones. Noor Al-Samarrai helped with the fact checking. Original music and sound design Ryan and Hays Holladay. And, a special thanks to the whole Center for Humane Technology team for making this podcast possible.

Tristan Harris: A very special thanks goes to our generous lead supporters at the Center for Humane Technology, including the Omidyar Network, Craig Newmark Philanthropies, Evolve Foundation and the Patrick J. McGovern Foundation, among many others.