

AI and You

Transcript

Guest: Beth Singler, part 2

Episode 39

First Aired: Monday, March 15, 2021

Welcome to episode 39! Today we'll finish the episode we started last week with Beth Singler. She is Junior Research Fellow in Artificial Intelligence at Homerton College, University of Cambridge. Prior to this she was the post-doctoral Research Associate on the "Human Identity in an age of Nearly-Human Machines" project at the Faraday Institute for Science and Religion. She has been an associate fellow at the Leverhulme Centre for the Future of Intelligence

Last time, we talked about her documentaries that, like her, explore the social, ethical, philosophical and religious implications of advances in artificial intelligence and robotics. The first, *Pain in the Machine*, won the 2017 *Best Research Film of the Year* award from the Arts and Humanities Research Council, and we talked about the ethical and moral issues that were raised. You can find [Pain in the Machine](#) and its sequels, [Friend in the Machine](#), [Good in the Machine](#), and [Ghost in the Machine](#) on YouTube, and linked from Beth's page at bvlsingler.com and in the show notes and transcript.

Let's get back to the interview with Beth Singler.

Let's talk about how you got into this field. What was the trigger that made you go, oh, that's fascinating, I want to work on that. I've I'm going to devote my intellectual, uh, energy to that.

Well, like I say, I'm a little unusual in my background, being more of an anthropologist of contemporary forms of new religious movements and religion. But I've always been a huge geek, and I think some of the earliest influences on myself - actually, someone asked this question the other day, and I remember that that probably the first time I thought this was an interesting area to get into was watching *War Games*. This is dating quite how old I am. But *War Games* from the 1980s, where it comes down to this question of nuclear thermonuclear war, and actually there is no winning position in the AI system, and that comes up with this as playing it as a game. There is no winning position. And I thought that interaction was just fascinating as a child, but then also seeing a more personified AI in Commander Data in *Star Trek the Next generation* as well. That was one of the kind of touchstones for my early conception of how we illustrates an artificial being in the media and that how those stories were used to reflect various aspects of humanity back. So throughout the course of the series he developed hobbies and habits, and he tries to have a pet, and he tries to form romantic relationships, he tries to form a friendship relationship. And all the times at which it doesn't quite go right for him as an artificial being says an awful lot about what we think those relationships should be like, and also an awful lot about what kind of limitations we imagine artificial beings to have. So I think my early science fiction interests, obviously, and then just finding this a really fascinating area of development, everything seems to be changing almost every week that there's another bit of progress that people are responding to. So I just did another podcast, actually on AlphaFold. And

this solution for the protein folding problem from Google DeepMind. And this again is being, you know, a big trumpeted moment and online people reacting for my interest in my area of anthropology, people reacting with existential despair, thinking that this means any moment now AI is going to develop the technology to replicate bodies for itself. So you can see - you're rolling, your eyes, I can see it - but I think that's really fascinating that as soon as there's any kind of advance - and part of the podcast was discussing quite how much of an advance this is - but as soon as there's any kind of level of advance with AI, so many people, instant reaction go to some of the sci fi tropes and the fear and the dystopia and, of course, the utopia as well. So there could be some interpretations to say Well, now, AI can build a wonderful world by using different proteins and, you know, nanotechnology. So there's these kind of it, two extreme reactions I find fascinating that how we tell stories to ourselves about what AI is and could be.

I rolled my eyes because it hadn't occurred to me that someone would go straight from AlphaFold to *WestWorld*, uh, which is another dystopian idea there about artificial intelligence. Although I like some of the questions that it asks and some of the things that it portrays. In fact, we've seen a number of forays in fiction on visual media recently that have done, I think, quite a good job. *Humans*, *Black Mirror*, *Westworld* will ask different questions, but often the same question. And that I think, as you suggest, is priming people to have these fears and thoughts about advanced artificial intelligence that we have no evidence that were close to. But they don't have anything else to fill that gap. Are you trying to fill that gap and help them realize what's realistic?

Yeah, I hope so. I think by unpacking some of these dystopian fears and thinking about where they come from, where the cultural trends come from, what the repeating patterns, why we react so viscerally to apparent minds in places that they shouldn't be if we have an understanding of should and shouldn't, sort of our kind of conception of the taboo of minds in places that they shouldn't be. If we can unpack those things, then it's like being in the dark and waking up and seeing a shape and deciding it's a person in your room. But then being able to flick on the light and realize it's a coat on the back of a wardrobe or something that the more we illuminate what's actually happening with the technology - also to illuminate what could be really quite potentially hazardous now, before we even get into the sci-fi future of the robocalypse, what are the mini robocalypses that individual people will experience now, so the decision about whether or not you get a mortgage or change the course of your life. The decision of whether or not a diagnosis is made for you could change the course of your life. I think we don't have to go all the way to the extreme dystopian future to get these sort of outcomes. And I think that's why you know, *Black Mirror* is quite successful. That yes, it does have some of those more kind of extreme ideas in it, with simulation theory scenarios and things, but some of the some of the outcomes of very near future as well. And they highlight some of the stories that we need to pay attention to.

Yes. And by taking it to extreme, it can cause us to reflect on things like impact of automation on jobs that are current concerns and that otherwise creeping up on us, and maybe we haven't paid enough attention to until they turn exponential and it's too late. You told me earlier. You have a nine-year-old child.

I do, I have a little boy.

What do you think about his world 10 years from now? Are you helping him prepare for that, do you think about what your responsibility is for helping him get ready for a world that could be so different as a result of AI? And I ask because I've got children myself, eleven and seven.

Yeah. I mean, it's difficult because I don't think I have any really very good predictive powers to be able to say I'm definitely certain that this is the direction that everything's going in. I think there's general trends that we should try and prepare for, like, as you say, the increase in automation and the increasing automation in knowledge-based industries as well. Physical automation is not the end of the story. We will see more and more and more cases of thinking work being handed over to systems as well. It's really hard because part of me just wants him to not have to worry about this, you know, and not just fill his head with concerns of, make sure you get a career in a particular path and try and limit his options. I mean, no parent actually wants to really stifle their child's dreams. Unfortunately, my son's main dream at the moment as a nine-year-old, which is shocking, is to be a YouTuber, which - I think of all the things that might be automatable. What the videos that he really likes. I don't see a huge skill set there, but this is me being a cynical, older person. But yeah, I don't know. It's really difficult to know. I know my parents kind of suggested in various ways I should concentrate on more practical things. I always wanted to be a film writer, and they knew, as most people do, that that is an incredibly difficult industry to succeed in. So they suggested I did some other things, and I've ended up in an academic career that's enabled me to do some of those creative things without being, you know, the Oscar-winning screenwriter I always dreamt I was going to be. But I think there's ways to help people towards their dreams without limiting them. It's just a concern how much of the future will be curtailed by these more reductionist systems that, as we said, put people in boxes and decide almost on their behalf what's possible for them. That is also a big concern.

What does the place you're at in Cambridge, the environment, the college, the university, the institute that you're working with, the people there, what does that provide in terms of resources and inspiration to further your goals?

So I'm what's called a Junior Research fellow at Homerton College, one of the many colleges of Cambridge. And it's a privileged position that we get 3 to 4 years. I have four other - others only have three. We have 3 to 4 years basically to pursue our research. We apply with a specific research plan, and then within those 3 to 4 years, we have free reign pretty much to do as we wish, as long as we stay obviously within that area of that research plan. So yeah, I think that is a very privileged position to be in, and the freedom that gives me. And Homerton is very supportive of public engagement work supporting their own efforts to highlight what they're up to and what they're able to give to students as well. So I think that's a good relationship to being. It's a temporary one, obviously and precarious in its own ways, being an academic. But I think that's I think possibly the most anyone could hope for in this kind of field. Lectureships exist. But then you're limited a little bit more on how much time you have for straight research versus teaching. So I think at the moment, I'm probably in the best position. I just actually have to get going and do the things that I've decided that I'm gonna do during this four years.

Do you have any idea what comes after that four years?

No. Yeah, I don't I don't know. Specifically. I'd like to continue as much as possible to focus on research, but yeah, there's limitations on that. That probably mean another short-term position somewhere. So we'll see. I don't know. I think the main thing for me, I'm working on a book at the moment, and that's gonna be the next thing. It's not anywhere near finished, so I'm not going to talk about it too much, but yeah, to have that is the kind of culmination of some of the research I've been doing over the last couple of years is the game.

I find that when I'm speaking to general audiences - adults, not Children, that I endeavor to present the optimistic outcomes and the pessimistic outcomes the good and the bad side of artificial intelligence equally because I perceive them equally weighted. And I go to great pains to weight them equally, but people glom onto the negative ones about 10 times as much. Is that something that you have found? And do you have any ideas about whether we should or how we can, make that better?

Yeah, I think I think that's been my experience as well. Certainly I also think I want to try and present positives and negatives a critique both of the utopianism and the dystopianism and find out what is actually happening now and where whether genuine concerns now should lie where we can also identify the genuine benefits of artificial intelligence at the moment. Yeah, I'm not sure It's really difficult because I think as with science fiction, the dystopic tensions make in some ways for better stories, like If you had a happy, utopian story where everyone's just milling about being happy, there's no dynamism, there's no tension, there's no plot. And I think people are looking for a plot in the story of AI in the real world as well. I talk sometimes about how we perceive ourselves in the story of AI that with the influence, the cultural influence of stories like the Terminator, we often position ourselves in that kind of almost survivalist, like righteous warrior position of that if we can understand what's happening, we can try and change it. If it's going to be this topic, we can be a part of that story individually and collectively, to try and turn things around. I think that's very attractive to audiences, both when you're presenting science fiction and when you're talking in terms of science facts as well. So getting people to kind of step out of the way of being the protagonist in the story, I don't entirely know how you can do that without just really emphasizing education on the realities that this is not a story where you need to step in with an AK-47 and gun down the Terminator. This is a story where you need to think critically about the applications of AI, think critically about who's applying ai. We mentioned Elon Musk earlier. Again, I wouldn't I wouldn't straight down the line stick him down as you know, the antagonist of the story or the protagonist of the story. Uh, he can be a very controversial figure, but we shouldn't necessarily hold him up as *the* reason why something is going in a particular direction is a combination of many different charismatic authorities. And AI itself as a charismatic authority because it's it has so many variable interpretations for audiences, they can place so many of the aspirations and their hopes and their fears onto this entity that we called AI, that sometimes in public's imaginations bears no relation to what the technology is and what it can do at all. It's an imaginary being like a liminal creature, and that gives it narrative strength that's sometimes hard for us to kind of push back and say, "I should know the reality is this."

Yes, we're talking a lot about, I think, the collision of human nature and technology and - not that this is the first time that this has happened, but it makes me wonder whether we are drawn to using a I like anything else to cause havoc just so we don't get bored.

Yeah, I wonder. Yeah, actually, boredom's an interesting concept to think around. I talked about earlier about loneliness and this idea of conceiving, creating AI consciousness because we are in some ways lonely, which completely ignores the diversity of voices and entities on our own planet, but we sort of get into this speculative space of wanting to create for transhumanists the concept of the mind-children, that which will come after us if we're doomed to die out as a civilization, at some point, perhaps our mind-children will go to the stars and be our progeny. I think that's an interesting reflection on our conception of AI and then boredom that you know, for some around the kind of survivalist prepper culture this idea of a pending society in order to then have a situation where you can rise to the top and be the Sarah Connor be the John Connor fighting up against the robot uprising is also an interesting interpretation of why we focus on AI. Could it be this great disrupter? Disruption is a such a key word in the technological marketing around AI. But is it a desire for some sort of new order out of chaos? There's another question that's quite interesting.

One of those possible future outcomes that transhumanists focus on is that AI becomes super-intelligent, and then we have to evolve some kind of symbiosis or evolution to coexist with it or merge with it, and it and some people talk about humans being the vehicle for an ultimate superintelligence to create itself with, makes me wonder whether dogs create humans so they have companionship in that way, we'll have the same kind of role. There's a there's a joke in there somewhere, but I think not many people are actually thinking that far ahead. This is the sort of territory of Bostrom and the like. But other than transhumanists, is that a bridge too far for too many people or not?

Well, I think it's very interesting how much of that transhumanist ideology and narrative does filter down to the general public. But again, through science fiction films that pick up on some of these ideas and that then disseminates. But again with the caveats that public understanding of AI and even awareness that it even exists is so variable. I mean, it's the old unknown unknowns that some people are not even thinking around the topic of AI at all. And when you're like myself and you yourself and we're constantly in this conversation, it's very easy to forget how little AI impacts on so many people's understandings of the world. So, for instance, I was in a taxi a couple of years ago. Very lovely, chatty taxi driver who asked me what I do. And I say I work in AI, like, broadly, and he said, "Oh, artificial insemination. That's really interesting. You help people have babies," and - artificial intelligence, but that no, that's that the assumptions that we as people who talk about AI a lot go into conversations with. We have to be aware that for so many people, this is not even on their radar. They might have occasionally a very over-the-top tabloid story about Alphago, and it was illustrated with Terminator pictures. And they said, Oh, this this looks like a terrible thing. But beyond that, there is there are vast swathes of the general populace for whom this has no bearing on their lives at all. But it obviously very much does because it will influence what decisions are made about them. And I think that, you know, we can get these moments of kind of hot points or melting point moments, like in here in the UK we

had a whole controversy over the A-level algorithms. I don't know if you followed this story, that basically because of the pandemic and students not being able to take their A-level exams that allow them into university here at the usual time, that predictions were made using an algorithmic system and the predictions were very heavily weighted, dependent on whether your school traditionally had done well for A-level results or not, so you could be the smartest student there's ever been at your school, but because your school doesn't do well historically, the algorithm, obviously with that weighting would not give you a good result. And then we had protests in the streets, students lost their places. And I don't know if I can swear on your show, so I'll say they said a lot of the students said things like, "Beep the Algorithm," right, you know, and even our Prime Minister, Boris Johnson, blamed the mutant algorithm for making these decisions when obviously it was absolutely the way the algorithm was set up, written, data that was inputted, the decisions given to the weight of that data that reflected a history of algorithmic thinking about students that what you come from, what school you come from, predicts where you go. So those kinds of flashpoints were suddenly, into the public consciousness and application of AI that has gone very wrong for people will, of course impact our understanding. But if you know, until that happens for a lot of people, this is not a day-to-day concern.

Although it seems a bit of a stretch to have called that AI; I mean, as far as I can tell, that sort of weighting is just the kind of statistics that we could have done 60 years ago.

Yeah, again. So what I'm saying about that term "artificial intelligence" that that was the reaction was. And as I said, with Boris Johnson referring to in a different speech, pink-eyed terminators a lot of these conflated and the algorithm being criticized for this. For some people, that was understood as AI, for other people, it was not understood at all. So again, as I say, this variety of levels of digital literacy and understanding of what artificial intelligence is and isn't what an algorithm is and isn't.

And it was convenient to divert the blame onto something that people don't understand as opposed to, say, blaming mathematics, where people would realize that that couldn't be at fault, and it would have to be the people that made the decisions.

Of course, because if you blame maths, then you blame mathematicians, the humans behind the scene.

I think that you've exposed something possibly quite interesting that the reactions of people to artificial intelligence news, and the way that it's presented in the press, with going straight to the Terminator every time, is a result of precisely the fact that it is not intersecting their lives yet in much of a tangible form. And so it's this thing on the periphery that may or may not be heading towards them. But if it had actually arrived and they knew about it, it would be like electricity, and it would not be this great unknown. It could still be dangerous. But you would know not to stick a fork in the socket. And yet AI has that promise for according to people like Andrew Ng of being the next electricity, so when it does arrive, it could be that ubiquitous and And so do you think that this, um, difficulty of relating to it properly and the way it's presented to them is a result of it not being that prevalent?

Yeah, there's I think there's a famous saying that when AI works, it isn't called AI any more. So in some ways, from the very basic definition of how algorithms work and robots work, we have those around us already, and they become so, as you say, ubiquitous and invisible. Some people give us the example of the washing machine as a robot. I mean, again, the language is changing there from AI to robot. But the things we're familiar with that use programming aren't as spectacular and as amazing as our imaginations of the robot from the 1940s 1950s, *The Jetsons* cartoon where we had the robot maid that never really manifested. But the washing machine is the robot maid its just in a form that actually makes much more sense than the ambulance robot that walks around your house picking things up. All those people still make attempts at developing those sorts of technologies, but yeah, absolutely. I think the more ubiquitous AI is it serves some purpose is not to highlight where it exists as well. You brought up *Westworld*. The really obvious AI in *Westworld* is of course, the synthetic entities who live in *Westworld* and then escape if you watch season three - again spoilers. Sorry. But when they've escaped, the more prevalent and the actually more dangerous AI is Rehoboam, which is the kind of world computer like Asimov spoke about with his great continent mind computers that could make decisions for humanity all around the globe. And Rehoboam does the same thing based on predicting what your future is going to be and also shifting people around like pieces on a chessboard. So I think science fiction gets there earlier, perhaps? We start seeing more and more instances how AI is becoming ubiquitous in decision-making systems that aren't visible and that just becomes part of our daily life, unless we start questioning why things are happening.

Well, that's a good point to start wrapping up. What sort of questions would you like people to be asking more of, that would be something that they could think about from day-to-day and observe their world a little differently as a result?

Yeah, I'm really interested in questions around trust and agency. So some of these recommendations systems and people's understanding of them means handing over a certain level of trust because they assume that these recommendation systems are some ways smarter or more rational. And that's something you see in the wider discourse around AI people assuming that humans are messy, problematic creatures and AI might in some way be cleverer and smarter and more rational and more neutral and more objective and make better decisions. One of the examples I often get quoted to me is this idea that actually we should allow AI to make decisions in court cases because judges are more likely to find people guilty before lunch because they're hungry. And when you unpack that and you look into it, it's not actually the case, there's a complete debunking of it. But this idea that humans, perhaps with our millions of years of messy evolution aren't possibly the best ones to be making decisions. And that kind of handing over of agency and trust is something I think the general populace should question. How much do we want humans to be involved with decisions, and how much do we want to actually allow AI systems to be making those decisions on our behalf, and how much those decisions are still being informed by normative values that come from humans? So those are the three questions that I'm quite interested in there. And more broadly, what do we want the future to look like? Do we want to hand over so many of our responsibilities that we are kind of almost, a side story to the main plot, or do we want to be the main characters and whether that whether that means for some

people, that they're going to be warriors or whether it's about taking back responsibility and agency? I think that's an interesting tension that we can keep talking about.

It reminded me that a few episodes ago we had Ted Parson, the law professor at UCLA. And we agreed that AI could do a better job than judges or juries and at least in many cases in the American South, for instance, in racially charged cases that have happened. How to get to that point, though, is obviously, a very different and difficult question.

With the concern there being that we already have examples of where parole hearings using algorithmic decision-making systems have been more detrimental to people from ethnic minorities than to white people. So we already know some of the outcomes, again because we have these normative values being inputted into systems based on data that has nothing to do with the individual. Goes back to what I was saying about the A-level students. So you are an individual. You should not necessarily be judged on the activities of your school or on people who look like you historically, and that's one of the problems of algorithmic thinking; it reduces people to an output of an input that they had no hand in.

Right. Of course, if you train your AI on examples of human decisions that are biased, it will replicate the bias. But you can equally well expunge the bias from it if you know what to look for and then AI has the benefit that you can copy it *ad infinitum* and export it to places that could use that lack of bias.

There's lots of talk about that form of cleansing of data, but in any position you take on the data, you're coming in with some form of bias. So that's a difficult one. I don't think there is a big answer to that. But again, it's about the public having the knowledge to be able to question how much of that responsibility they want to hand over to AI systems.

And as you say, we're here to ask more questions and get people to ask more questions, and you've done a fantastic job of that. What's up next for you that you want people to look for? And how should they follow your work and find out more about what you do?

Well, like I said, hopefully I'm working on a book concept and who knows when that's going to be done, so I can't. I can direct people towards that, and I think part of the reason I don't know when that's going to be done is because I'm a complete Twitterholic. So I'm very easy to find online. I'm on all social media that you might expect, but certainly definitely Twitter. It's partly also where I do a lot of my digital ethnographic work. So yes, you can find me. I have a website, bvlsingler.com. I'm @BVLsingler on Twitter, and yeah, I'm just happy to keep seeing people having this conversation and showing interest. And that's why I enjoy coming on podcasts like this.

Oh, thank you so much for doing it too. And we'll put those links and handles in the show notes. Beth Singler. Thank you for coming on AI and you.

Thank you very much for having me.

That's the end of the interview. I was really interested to hear how an anthropologist – well, an anthropologist and a geek as well – sees this cultural phenomenon of our collective reaction to AI. The way that we project or focus on the dystopian aspects in particular seems like a Rorschach test, you know, the what-do-you-see-in-this-inkblot test. One person sees a butterfly, someone else sees the Boston Strangler. I don't mean to suggest by that that all the predictions for our future with AI are the product of the subconscious of whoever's talking. A lot of thoughtful people have gone to great effort to put objective analysis and scientific inquiry and serious research behind their answers. But the void, the gap between what we do know and what we want to know – and that is vast when it comes to what AI is going to do – is something that all of us to some degree populate with our fears and hopes. The trick is to not be at their mercy. Some people can play with what we do at both ends of that scale, like Beth. Please go and find out more about what she does, at bvlsingler.com, where you can find the YouTube links to her documentaries/dramatizations starting with *Pain in the Machine*. I was particularly taken by her idea about how people react to these dystopian tropes by constructing survivalist stories where they triumph in some degree, and how that explains the popularity of the Terminator franchise and how it keeps getting dragged into the conversation about AI. We'll talk more about the science fiction angle in a few episodes when we talk with author and screenwriter David Gerrold.

In the latest news from AI headlines, there's now a robot that can pick fruit – well, we've had things that can meet that description in some form for a while, but this one is a drone. It's from Tevel Aerobotics Technologies and it's got this arm with a claw on the end that can grab peaches, apples, that sort of thing. The trick of course, for something like that to be cost-effective, is for it to pick the fruit without human direction, and that's where the AI comes in. Its camera can classify fruit as being ripe enough and big enough, and then it figures out the best way to approach it and snag it. When picking fruit, timing is extremely crucial since, for example, fruit picked two weeks late loses 80 percent of its value. You can find many news stories about fruit spoiling because it wasn't picked in time for one reason or another. So here's another way that AI can help. It brings up a constant concern about automation, that it will take away jobs from humans, and although the company says that it's designed not to replace human fruit pickers but complement them, of course there is the potential for taking over. But picking apples off trees all day every day is a grinding activity that doesn't much honor our humanness. No one should have that as their best or only choice of how to contribute or survive. Getting machines to pick fruit should one day be seen as a move as compassionate as getting machines to sweep chimneys instead of sending children up them. It's the transition to that point that's difficult, of course.

Next week, I will be talking with Peter Asaro, who is a philosopher of science, technology and media at the New School in New York. He has focused heavily on the use of autonomous weapons, contributing to many of the papers, discussions, petitions, and proposals for moderating our development and use of those devices.

Until then, remember: no matter how much computers learn how to do, it's how we come together as *humans* that matters.

<http://aiandyou.net>