BCavello_FinalCut

Tue, 11/3 10:20AM 44:01

SUMMARY KEYWORDS

work, ai, people, technology, concerns, bias, conversation, area, world, space, machine learning, ways, thinking, data, future, partnership, questions, understand, organizations, exploding kittens

SPEAKERS

B Cavello, Rebecca Scott

R

Rebecca Scott 00:00

How informed do you feel about the future of artificial intelligence? Have we addressed his challenges and biases? Or do we need more conversations to improve his trajectory? In this episode I speak to B Cavello, Research Program lead at the Partnership on AI, on the opportunities to improve AI, the importance of educating the public on its use, and why it's so important to bring more people together to discuss its future.

B

B Cavello 00:24

It's our responsibility to welcome people into this space of magic and power and wonder and remind people that this you know, with great power, there is a sense of responsibility needed in the space of AI, and when we demystify it, not through dismissing people's concerns about its power and magic, but instead by embracing them, we can empower people to make better informed decisions about what the possibility space looks like, and also hopefully encourage them to feel a little bit more energized and welcome toward actually contributing their viewpoints or their creative impulses to the space as well.

R

Rebecca Scott 01:00

B Cavello is a Program Lead at the Partnership on AI, a nonprofit multi stakeholder initiative focused on advancing the benefits and addressing the challenges of machine intelligence. B leads PAI research on areas such as fairness, transparency, AI's impact on

labor, and works closely with PAI partner organizations. B's personal projects focus on social good, and range from textile art, to computer education, insecurity, and even a game to help NASA researchers study climate change. So ready to dive into the current and future state of AI? Let's discuss. I'm Rebecca Scott. and this is Humans, Now and Then.

R

Rebecca Scott 01:50

B Cavello, thank you for joining me.

B

B Cavello 01:52

Thank you for having me, it's a pleasure to be here.

R

Rebecca Scott 01:55

We had a conversation a few weeks back, and the work that you're doing around AI, machine intelligence, machine learning is fascinating. But your journey to get there, a journey to get to your place right now with your level of expertise and some of the experience you've had, including this whole amazing adventure around Exploding Kittens, is fascinating. And so I'd love to start the conversation with a little bit about your journey and how it got you to this place of having these tremendous opportunities around shaping the future of artificial intelligence and machine learning.

B

B Cavello 02:25

Thinks Yeah, it's a funny thing. You know, I have so many people reach out to me, especially as one of these perhaps less conventional people in the space. People are always wondering how did you get to where you are. And there's some degree to which I feel like my journey is so strange that I don't have great advice for them. But I also try to pull out some high level themes. So for some context, I've always been really interested in these topics. Actually, I was going through some old school papers and found something I'd written in middle school about the potential and concerns about artificial intelligence and society. But I never was someone who was so passionate about it that I, you know, like, studied computer science in school and got a masters or a PhD in the topic. Instead, my journey, you know, involved studying economics in undergrad, I became really fascinated with storytelling and game design, and ended up moving out to Los Angeles where a friend of mine and I ran a nonprofit organization that had an arts and education charter, so teaching people about technology through games and storytelling, and through the games world, I ended up making friends with this guy Elan Lee, who was

actually one of the original Xbox team members who worked on alternate reality games as sort of storytelling that is this almost like large scale LARPing this work at the threshold between kind of your real life and the story world. And with him, I ended up joining Exploding Kittens as the first employee as we launched an incredibly successful Kickstarter campaign we raised almost, well, gosh, I'm forgetting the timeline now definitely a million dollars by the end of the first day, almost 9 million by the end of the campaign. It was an absolutely wild ride as the name Exploding Kittens might suggest, we were basically creating a paper card game and not in the AI space at all. But really something that was an incredible kind of community management, product design, everything that you can think about, you know, when people say they're wearing many hats, we were literally, there are a bunch of photos of us hanging out literally in the garage of his house where Exploding Kittens started. And, you know, we're all wearing a variety of different party hats. So it was truly an all hands on deck startup situation where we were doing incredibly large things with an incredibly small team. But I knew you know, in my heart of hearts, I really wanted to be back in this kind of cutting edge technology space, those early conversations I had with people doing this nonprofit Education work really made me feel like this is the space I want to work. And it was actually through a happy accident that someone I knew from the crowdfunding world, someone who had met actually, at a Kickstarter party at South by Southwest, suggested me to be a part of this very unusual and interesting team at IBM in the Watson group. So Watson is IBM's, Artificial Intelligence Division. And in the Watson group, especially in the early days of commercialization, you know, IBM had built this technology, but they didn't necessarily have the kind of educated client base to really understand what this technology might mean for businesses. And so there was this little team within the Watson implementations group. And our job was essentially as communicators, as educators as storytellers, to help companies and their leadership strategize about how AI might impact their industry, and specifically, their company and working there brought me out to San Francisco, where I ended up participating in all of these many communities around kind of tech and social impact in San Francisco and through one community, the foresight Institute, which thinks about kind of long term implications of technology. I met the then Director of Research at the partnership on AI, and they invited me to apply to the program lead role that I'm in now. So today, I run research programs, at the Partnership on AI, which I'd like to think polls from all of these very disparate experiences, my economics training in undergrad, my scrappy startup mentality of Exploding Kittens, you know the high level business conversations with executives at IBM about the future of artificial intelligence, and this kind of deep rooted passion around both education and empowerment of people, both through technology and around technology, and that really informs the work that I do today, working with my incredible colleague, Katya Klinova, working in the area of programs at Partnership on AI, that focus on AI's impact on labor in the economy.

R

Rebecca Scott 07:07

Wow, I mean, that it is a fascinating story. And I think this, it really speaks to one of the things that is one of my areas of passion, which is around how having a diverse set of experiences in life really helps you bring great perspective to your work. But I think sometimes that organizations think a lot about very specific technical ability. And I almost want to challenge folks who really, you know, look for people who can help them that have very broad experience in the world, have lots of different broad perspectives, different experiences, different kinds of things that they've gone through, and all the learnings that come with that. So I think that's really cool. But one of the other things, of course, it's really cool as this involvement you have with Partnership on AI, and I'm going to venture to guess a lot of listeners may not be aware of all the great work that Partnership in AI is doing, and some of the work that you're doing that is absolutely fascinating. So I'd love to share a little bit more about that work that you do.

B

B Cavello 07:58

Thanks. Yeah, the partnership on AI is a really interesting organization we were founded about four years ago now. So coming up by into our fifth year soon. And it's really an organization that is a coalition nonprofit organization. So we act as a forum sort of hosting conversation about the important and sometimes difficult decisions about the future of AI. And we hold this multi stakeholder approach, where the research that we do the work that we do is really around bringing together a variety of voices from you know, industry, the big tech companies, academia, like the MIT Media Lab, or Oxford Internet Institute, as well as other nonprofits and civil society organizations. You know, Human Rights Watch, the ACLU, the EFFs of the world, and you know, all of these different community stakeholders who have incredible voices to lend and perspectives, to really inform the direction that artificial intelligence development goes in. And the way that we approach this work through this kind of multi stakeholder initiative approach allows us to really host conversations and bring together people who might not always be in the same room together might not always, you know, work cooperatively or see eye to eye. and the work that we do is really around facilitating these conversations so that we can ensure that the recommendations that we make and the actions that our partner organizations take, try to uphold this diversity of values. And you know, the areas of work at PAI are as diverse as well. We have work underway on a variety of different topics. I mentioned the work that my colleague Katya and I are doing as it relates to AI's intersections with labor in the economy, we have work streams focused on kind of AI and media integrity. So how does how is AI used in generating synthetic media such as deep fakes as well as how our algorithmic recommendation mechanisms such as your kind of your Facebook newsfeed or things like that serving up media to you, and how do you navigate that landscape is quite complicated. We're also doing work as it relates to the kind of issue this big bucket

of issues around fairness and transparency and accountability in AI, especially as it relates to both things in the criminal justice system or, you know, legal aspects as well as ways in which companies and organizations developing AI can think about documentation as both an artifact and a process for cultivating transparency. Another area of work that we're also involved in is one around publication norms. So how can we be responsible and thoughtful around the risks of publishing high stakes research, learning from fields like, you know, nuclear research, or the ways in which biological research can result in runaway effects? Similarly, we want to think really critically about, you know, how can we balance our values of openness and sharing research knowledge, with thoughtfulness about the ways in which that knowledge might be misused or abused when it's out in the world across all of these different project areas, there's this core value around trying to bring together a variety of different perspectives and shapes, something that's a really strong recommendation that organizations can use to inform how they approached the development and the deployment of AI systems.

R

Rebecca Scott 11:34

Yeah, and I think that is critically important work. And, you know, we've had different conversations. And of course, I've had conversations with folks out in the world, whether they be guests on the podcast, or just people I meet, and whenever the topic comes up in relation to artificial intelligence, machine learning, or even just like, you were talking about things that get served up in our Facebook feed, for some folks that don't understand the technology behind it, you know, there's obviously an education gap, I think, for the general public, on how that works. And so some folks will think it's just downright creepy. You know, other folks might think it's convenient. And in some respects, that might be, but I think there's more and more questions around the amount of messages that we received, and maybe some level of growing concern around the decisions that are starting to be made on our behalf in relation to artificial intelligence, all the data that's out there about each of us or our demographic, so to speak, and then thinking if we venture out in the future, education of, you know even just organizations coming together and thinking about how they can ethically move forward in this space. But how does Partnership on AI approach or think about the education required for the general public to really understand the direction that we're heading and why it's really important for them to stay kind of informed on the topic?

B

B Cavello 12:49

Yeah, I'll speak a little bit to Partnership on AI's perspective, but also to my own because this is something that I'm deeply passionate about. So the way in which we approach this work is with a recognition that a lot of these topics are frankly quite specialized, quite

niche, quite intimidating, sometimes to people who aren't familiar with the industry terminology. And I think that the work that we've done around for instance, policymaker education speaks to this, especially some of my colleagues have done work on both issues around facial recognition and understanding, what is this technology? What are its limitations, where are its flaws? And how is it misused? As well as thinking about kind of effective technologies or emotion recognition systems and thinking about all of the different ways that these are applied today, as well as how they're being developed for the future in ways that could be concerning or, you know, promising, depending on both the application and your perspective on the subject. So I think that public and policymakers especially education is incredibly crucial. Given that there are really big decisions being made right now, I happen to live in California, where I did not live in California my whole life. So when I moved here, I was quite surprised by the ballot propositions that there are so many of here, which basically, you know, using a sort of direct democracy put decisions about lawmaking into the hands of citizens. And that's incredibly powerful. And it could be promising, but it also can be quite concerning when policymakers or when the public don't understand the technologies or the potential implications of the technologies about which they're asked to make decisions. And certainly, there are many different forces out there, you know, that have particular viewpoints on these topics. And so giving people an opportunity to make a choice is difficult when they don't have all of the information. So, you know, something that really motivates me in this work is this fundamental kind of metaphor or analogy that I have, which is that, especially today, look at modern computers or smartphones. They're so beautifully designed and sewn up and neat and tidy. And all the screws are so tiny. I remember when I was buying my first Mac Air with the new, you know, aluminum casing and everything like this. And when I was in the computer store, I was asked, oh, which you know, which version do you want? How much RAM do you want or something like this? And you know, I was like, Oh, I'm not worried about it, I'll get the less the smaller one now. And if I need to throw more RAM in there, I'll just do that later. And the guy at the computer store was like, I don't think you understand, you are never going to open up this computer and install more parts. No, that's not how they work. They're all perfectly designed just so. And I think that that's the direction that we've seen technology headed in recently, with an eye toward ease of use, and you know, elegant design and things like this, which I think comes from a good place. But the challenge here is that these technologies are so sealed up and so beautifully polished that it feels like magic. I can't tell you how many times I've seen videos of parents astounded by their toddlers' ability to navigate through an iPad, which you know, now today, I think people are so familiar with toddlers navigating iPads, but these technologies are wondrous and amusing and magical. And in many ways, I actually think that that makes people feel shut out from them. They're so magical, that people say, Well, you know, I'm not magical, like this is tough stuff. This is magic. I don't know these things. And one of the things that I see

technologists often do is they will hear that, and they'll say, Oh, no, no, no, it's not magic. Look, look, you know, let me show you all the ways in which this isn't magic. It all makes perfect sense, once you know how it works. And I actually think that that's not the way to approach these things. I think that there's value actually and saying, you know, who cares if it's magical, you are a wizard, you will learn to master these technologies, you can learn to become familiar with all of their mysterious inner workings. I think anyone who learns about, you know, EM physics is, has to have some appreciation of the sort of magic of it all. And to that extent, I think that I see my personal role as a storyteller. And as someone who thinks about these human experiences of technology, it's our responsibility to welcome people into this space of magic and power and wonder and remind people that this, you know, with great power, there is a sense of responsibility needed in the space of AI. And when we demystify it, not through dismissing people's concerns about its power and magic, but instead, by embracing them, we can empower people to make better informed decisions about what the possibility space looks like, and also hopefully encourage them to feel a little bit more energized, and welcome toward actually contributing their viewpoints or their creative impulses to space as well.

R

Rebecca Scott 17:52

Oh, yeah, I think that's so critically important. And I think you're right, some people do kind of become intimidated by technology, or some of the articles they read about technology, I think some of the language that, you know, folks might understand in relation to some of the terminology about technology, you know, it intimidates folks, they feel like they can't learn it, or it's not for them. It's just, you know, where they'll give it up and assume everyone else is taking care of it. Even if they have concerns, they're not feeling like they're a part of the conversation, because quite frankly, they don't really understand it to the right level. But you're right, this is where that gap of education is critically important. And even inviting people into that conversation, say, hey, you don't have to be a technical expert to talk about how you want your experience to be, or to raise the concerns that you have, because you'll either learn something that will help alleviate those concerns, or your concerns might be perfectly valid. And being part of the conversation, and being able to express those concerns is really important.

B

B Cavello 18:45

Yeah, I think, you know, especially as it relates to the responsible AI space, while educating people about technology is important. You don't have to be a technology expert, to be an expert. On the experience of living in a world with AI, you don't have to be a technology expert, to be an expert in thinking about those implications in your life and how your work is being affected. And so it's crucially important that those who are technology experts

hear from and listen to, and work with the people who are often kind of on the direct receiving end of of our innovations, and make sure that they're not just the recipients, but they're the CO creators.

R

Rebecca Scott 19:26

Yeah, I love that. That's so true. And so one of the things that I know folks have concern about, I mean, there's obviously a lot of data to back this up to, especially in earlier algorithms in relation to AI machine learning, the level of bias that tends to be built into these algorithms and their awareness of that bias, you know, starting to kind of raise up a little bit more and more where the general public is more aware that that bias exists. What are some of the things that you know, you think about relation to ensuring, you know, we think about things like diversity and inclusion in artificial intelligence, and how do we make sure to address or prevent some level of bias In algorithms as we kind of move forward into the future.

B

B Cavello 20:02

It's a great question, the topic of bias in AI and machine learning, as you've mentioned, is growing in kind of attention and popularity, as we see the negative impact of kind of lack of mitigation being applied out in the field. There are so many ways in which having a diverse community of stakeholders informing the development of the technology is crucial. One of them is around bias in data. So one of the things that I think has been getting a lot of attention recently, is how the data sets how the mechanisms that are used to train to build these machine learning models that ultimately guide the behavior of AI technology, those data sets are often quite skewed in terms of representing the world. And frequently, those representations are not only biased in the kind of objective sense that they may represent some groups more than others, but they're also biased in this kind of more societal sense in that the ways in which they over represent or under represent certain groups. Kind of perpetuates the ways in which systemic biases have caused terrible harms in our society, to one area that thinking about diversity and inclusion is incredibly important is in representation around data sets. And there are a lot of conversations happening about this, I think in some of the more visually obvious spaces. So a lot of the computer vision work around facial recognition. In particular, there's a big conversation about race and gender bias especially precipitated by the work gender shades, which called out some significant underperformance in commercially available machine learning offerings. That being said, there are a lot of types of bias in terms of representation, that I think go under acknowledge as well, especially around languages. So one of the ways in which bias is perpetuated through kind of a lack of diverse data is a ways in which machine learning kind of keeps trending toward dominant

languages and language support for both spoken models, as well as natural language processing of text. These are places where lack of data sets can kind of cause this, you know, secondary digital divide effect where new technologies emerging leave behind different groups of people because their languages or their ways of speaking, aren't supported. So I think that that's another area in which having a greater diversity of data sets can be really powerful. That said, I think that the question of diversity and data sometimes also can lead to problems. There's such an emphasis on diversifying datasets that sometimes people don't pause to really examine how those datasets are being created. And whether the people who are being added in for increasing the diversity are actually empowered in the way that that's happening. One of the projects that we're working on, at the partnership on AI is what around responsible sourcing for AI supply lines. So thinking about the data that goes into training our machine learning models, you know, giving them examples of what we do and don't want those systems to do. And that training data comes from people it comes from our lives, it comes from, you know, whether it's generated data, such as people taking photos, or creating posts, or whether it's data that is enriched through data labeling, or annotation processes, where someone goes in and, and looks at the picture that you've posted of your cat and labels it as relevant as a cat, or even highlights the particular pixels of the image that are the cat in the photo. And this is another area where there is a lot of work to be done. That isn't necessarily just around adding more data to the system, because that's, you know, something that I think a lot of people are hungry to do, but thinking about what where does this data come from and what work is involved in actually making it available. So the AI supply lens responsible sourcing project is really around making sure that the people who are doing that work who are generating the translations who are labeling the images, who are tagging emails as happy or sad or angry to teach these machine learning models, what these different things are, that those people have quality working conditions, and that their voices are represented in the design process. So Mary Gray and others wrote a book called Ghost Work, which focuses on this kind of behind the scenes work of generating all of this data, this labelled data for our machine learning companies, and that's an area where increased representation and both diversity but also just power is really important. And then there's a third area that I think is important to think about as we think about diversity and inclusion in the AI space, which is really around thinking about whether things should be built at all. So it's one thing, you know, to say, Oh, this facial recognition model, for instance, is biased in that it doesn't represent darker skinned individuals well, or it, it doesn't recognize women or things like this. that's problematic. But also we should step back and say, should we be developing this technology in the first place, there was some research, I want to say a couple years ago now that folks had created that was around essentially the the quote unquote gaydar of machine learning. So based on looking at people's images, we're going to predict their sexuality. And this is something that, you know, maybe they went into it with good intentions, but many of us in

the queer community would say, that's pretty scary, right? Like, even if it works, even if it really well represents queer communities in the data set. That's a technology that maybe shouldn't exist at all because of its capacity to be used for harm. And so when we think about diversity inclusion in the AI space, I think, in addition to thinking about data sets, in addition to thinking about, you know, whose voices are included in how those datasets were developed and where they are in the world and what they're being asked to do. There's this third area of, well, how can we have diversity inclusion in the decisions about which technologies we actually prioritize building. And I mentioned the publication norms work happening at responsible at the Partnership on AI, led by my colleague, Rosie Campbell. But this work also touches on another project that we're working on in the AI labor and economy space, which is the AI and shared prosperity initiative. And this initiative essentially says, you know, it's well recognized that technological innovation has an impact on labor, on people's economic prospects on our ability to get a job, right, the availability of jobs, we've seen, you know, past industrial revolutions, we know that technologies can dramatic transform that landscape, many people would acknowledge that with technological innovation, especially in the AI space, there are certain jobs or tasks or people who will be more impacted by the types of innovations that we're seeing. And some people have even suggested things like universal basic income, recognizing that there might be a net negative amount of jobs. So there might be some new jobs that are created, but they may be fewer than the number of people who are looking for work whose work has been displaced by automation. And there's this kind of additional layer of this conversation where we say, well, that's terrible right now, there are all these kind of future of work things well, how can we rescale people? How can we make sure people can find work, because what we recognize is that there's a skill bias in this automation, meaning that as people are in work is displaced, there are new jobs that are created, but they're not the same as the jobs that people were doing before. So there's this energy to help move people from working in the factory to the coal miners to coding boot camps kind of narrative. But the problem with that is that one, there may not be enough of these new positions available. But even if there are, the reality is transitioning from you know, the coal mine to the coding boot camp, is not an easy thing to do. And right now, it puts the entire burden of that transformation at that shift on the workers themselves, who are already the most kind of marginalized players in this space, instead of asking the question, Why are their jobs being automated in the first place? And what say do they have in that process, and increasingly, the AI and shared prosperity initiative is around harnessing this recognition that, Wait, maybe there are some technologies that actually help increase shared prosperity, help increase opportunity. And maybe there are some technologies that don't and rather than thinking of technology as this determinist course of progress, that we have no ability to shape, we say, actually, as innovators, as those empowered magicians who understand the magic, that is technology, we have an opportunity to actually make a choice about what type of innovation we're going to pursue. And that's a

space in which having the diversity of perspectives, the understanding of impact on stakeholders, the power at the decision table from those communities is so important to make sure that we're actually building the technologies that we want to see in the world.

B

B Cavello 20:02

Right. There's so much of what you just said, that is tremendously important for organizations especially to understand. And I think some of it is interesting at this point in time, because of course, we've been living in this world that becomes faster and faster. Organizations are continuing to try to be faster and faster and faster. But the reality is, is applying the type of deep thinking about things like should that technology exist or not? What's the value of this technology in the world, in society? What's the benefit? What's the potential risks or drawbacks to any type of technology, artificial intelligence in the world, are we going to have unintended consequences based on these technologies. These are responsibilities that organizations need to take very seriously. But it also takes time and effort and the right people in place to make these kinds of determinations to really understand the value and impact on the technology people are putting on the world. And I think in this day and age, it's almost becoming more and more important that companies in particular think about their responsibilities, both morally and ethically, about putting some new technologies in the world and the value that they'll have to society.

B

B Cavello 29:25

Yeah, I couldn't agree more. And I think that this doesn't have to be something that feels like a reprimand either. I think a lot of the responsible AI conversation to industry practitioners can feel like being scolded right, like don't do that, don't make these technologies. But I think that as much as that is the case that there are some things maybe that we shouldn't pursue. I think it's also the case that there are many really, you know, juicy, hard problems to be solved in the world. And tech, people love to talk about solving hard problems. And I think some of this is around also just like shifting focus and attention and recognizing like, wait, these are really interesting, tough, juicy problems for us to work on. I think one space in which I see a lot of momentum building there is around climate technologies. We don't have any initiatives underway at the partnership on AI right now around AI's impact on climate. But many of my colleagues have been quite passionate about this area. And I've had the opportunity to advise with an organization called the Frontier Development Lab, which partners with the European Space Agency and with NASA bringing together space researchers and AI researchers to solve these big global problems. And one of the areas of focus a lot of space research, a lot of satellites are up there are pointing back down here at us on Earth, they're Earth Observing

satellites, and Earth observation imagery is a really, you know, rich data set for people who are doing climate science and understanding weather patterns and longer term climate patterns. And I'm seeing increased energy there in the tech community to say, Wait, these are tough, interesting, juicy problems, and I have skills that can be applied to these things. So that's something that I'm really excited about. And as we move forward with AI and shared prosperity initiative, I hope we'll have that same energy to have people recognizing, you know, how much opportunity there is around, for instance, accessibility technology, and thinking about working with communities of disability to think about, well, you know, how can we build technologies to actually help bring more people in and help give more people opportunities, rather than thinking about how can we automate away the people involved in these tasks or these scenarios? So I think that there is promise and I'm really excited about that energy that I see. And I hope we'll keep that momentum going forward.

R

Rebecca Scott 32:47

Oh, yeah, I think there's we think about, like you said, the big problems, and I bring up the big problems a lot on the podcast as well, because I think if people think about, you know, the difference they can make in the world, finding a real true sense of purpose. In some respects, if you have those technical abilities, there's no better purpose to apply them towards those huge problems that society faces like climate and like you said, you know, inclusion, bringing, you know, the people in the disability community into different conversations into different work opportunities that they may have not had before. There's so many great opportunities in AI and machine learning and technology in general, to really do good things in the world that make a huge difference on the human experience, and just really society in our future. So those are exciting things to think about and look forward to. What are the some of the things that might concern you about the future?

B

B Cavello 33:34

The future? I am an optimist, the future is one of those big black boxes that we can't yet see into. And I tried to retain a positive outlook. But I think that there are things that, you know, even in the present are feeling quite concerning and worried about the direction I'll go. Certainly as it relates to the work that we're doing around AI and its intersections with labor in the economy, there's a lot of concern around concentration of power, the whether it be the processing power, or the concentration of data ownership happening in this space. I think that you know, as a as a kind of meta conversation about concerns about AI, I think that there are kind of big questions around the ways in which kind of capital concentrates where capital now refers to the intellectual capital as well of both research

knowledge as well as data sets. So that's one thing that definitely concerns me, I'm really interested in and inspired by efforts to kind of give people more control over their personal data and understand where that lives and how it can be used. I think that AI could be incredibly empowering. On a personal level, if I couldn't use my own data to understand my own health. That's wonderful. what's scary is that when that data is used in ways that maybe don't have my best interests in mind and I didn't give consent to. So I think that one area of concern is concentration. Another area certainly related to that, just sort of like privacy in general, something that I think a lot about is consent and agency, and the incredible knowledge asymmetries, the information asymmetries that exist today, and how much people are asked to, you know, assign some legalese that they never could have interpreted, or, you know, give away their rights or their data, to participate in the free ad supported services that we all know and love. And I think that that's something that will be an even greater concern. If AI systems do allow us to be better predictors, if we do see this concentration of power and capital, and those systems are good at making predictions about the future. Now we live in a world in which, you know, some people are going to have much more kind of foresight about what might happen from certain decisions and others. And I think that that's an area of concern, something else that I think we should be on the lookout for, or thoughtful about is the way in which AI technologies can kind of perpetuate, not just perpetuate the kind of obvious harms of you know, racism, or misogyny, or homophobia or things like this. But additionally, think about the ways in which kind of culture is shaped by our media, and even the ways I feel like, oh, wow, you know, sometimes I feel like it's such an exciting time to be alive in thinking about all of the incredible neologisms that are being coined, and all of the beautiful ways in which people are being so creative with language, you know, especially in queer communities, there's so much innovation in the language space. But a kind of interesting phenomena of having these algorithmic systems trained on historical data is that they kind of trend toward the the norm, they trend toward the history. And so as new movements as new ideas, as new ways of speaking, are created, they may actually be kind of difficult to sustain in the face of strong algorithmic curation. So a kind of trivial example of this is the autocorrect in your phone, or your computer that helps you to spell things correctly, or these kind of grammar assistants that help you to phrase things, since those are all based on historical information, it can be quite difficult to adapt them to new phrases, and new ways of speaking, which can inhibit kind of the spread of new ideas or new ways of living. And I think that, you know, that may be a more mundane example of it. But I think that there are really interesting questions to be asked there. There's a talk that Kevin Slavin gave, gosh, maybe back in 2012, quite a while ago on the TED stage talking about algorithms and the ways in which they run awry. And one of the questions that I love in that talk is he kind of uses the analogy of the stock market crash and these runaway trading algorithms that kind of got into these feedback loops with each other. And you poses this question, what would that mean, for a Netflix or for our cultural institutions? What would it look like

for culture to experience a crash? And would we even know how to recognize it when it happens? And you know, living, living in 2020, I feel like oh, wow, those are some poignant words. Yeah. And and I think that that's something that I see as concerning that we don't necessarily yet know how to talk about even. But it's something that I hope that we will continue to interrogate this kind of question around. What does it mean to learn from history? And what does it mean to train our models on the reddits of the world? And what are we, as we kind of normalize around those things? What is it that we're normalizing around it? Is that something that we want to actually kind of progress to them enough?

R

Rebecca Scott 38:58

Yeah, and I think those are so so important. And I gotta say that, you know, I love the work that you're doing. I think, even though there are some obvious concerns, and of course, you brought up my arch nemesis, which is, you know, like autocorrect, which seems to somehow pull the most embarrassing words, you know, from my mom's texts. But that being said, I really, you know, appreciate the work that you do. I think that, you know, hearing kind of the work that you do, the work that partnership on AI is doing and some of the other amazing people and work that you mentioned, really helps give me hope about the future and really kind of tackling these big problems that we face, and really putting our focus on those things that create tremendous value and importance in the human experience as we move forward. So that just I just wanted you to know that that gives me some level of hope and comfort that, to know that people like you are out there, asking these hard questions and really looking at it from a very optimistic point of view to kind of help us you know, think about the future differently than we have these kind of unanswered questions around some of the concerns that we might have as individuals out in the world. It's really good to know that you've got folks like you out there to help us help us along our path in understanding, you know, what's going on. But one of the things that, you know, we talked about too, is getting people involved in the conversation. If one of my listeners is really interested in learning more, or if they want to ask questions about their concerns, or get involved, what's the best avenue for them to do that?

B

B Cavello 40:18

Sure, if you want to get in touch with me out on the web, I'm very online, you can find me on Twitter. That's B, as in boy, @b_cavello. So my name b_cavello @ Twitter. If you're interested in learning more about the partnership on AI. Specifically, if you're interested in learning about the Shared Prosperity Initiative, we have partnershiponai.org/shared-prosperity. And there you can both read about the initiative so far, see updates, the deliberations of our incredible steering committee, they'll be sharing some impulse talks. And you can sign up to get involved as well. One of the things really in the spirit of this

multi-stakeholder decision making processes that we invite people to participate in our work. And so if you'd like to sign up to be involved in that, go to partnershiponai.org/shared-prosperity, and you can sign up for updates. And to find out when we are going to be looking to learn from you as much as you say, I might be inspiring others, which I'm so grateful to hear you say that, I am so inspired by everyone that I get to work with in this space and inspired by you and the community of Humans, Now and Then. So I'm really excited to hear from folks and and get in touch and talk about these issues more deeply.

R

Rebecca Scott 41:35

Absolutely. So folks, go ahead and get involved in shaping a better future. So B Cavello, this has been amazing and fascinating conversation. Thanks so much for joining me.

B

B Cavello 41:45

Thank you for having me.

R

Rebecca Scott 42:01

There's no doubt about it. B's work at Partnership on AI is critically important in shaping our future. As B brings together some of the world's largest and most influential companies to discuss fairness, transparency and AI's impact on labor, these conversations will shape our experiences, our work, our interpersonal connection, and how we live our lives. Given B's orientation to diversity, inclusion and doing good in the world, it gives me great hope that the underlying ethics of technological advancement will not be overlooked or underestimated. As AI makes more and more decisions on our behalf, our individual and collective voices must be a part of the conversation. Our experiences in the future will be largely shaped by the technology used in our homes, in our communities and at our work. If we don't understand what is happening, or what will likely happen, we will lose a tremendous opportunity to shape it and make it better. Again, you don't need to be a technologist to improve the path of AI and machine learning. There's no doubt that your future is being shaped by these conversations about AI and its advancements. So why wouldn't you join the conversation? Or at a minimum, why wouldn't you learn more about it? The information about AI is out there. And there are ears that are listening for voices just like yours. So get involved. Learn more. And, of course, while you're at it, go on, go help shape the future.

R

Rebecca Scott 43:27

To learn more about B Cavello and the work happening at the partnership on AI, go to partnershiponai.org. That's partnershiponai.org. You can also find B on Twitter @b_cavello. That's b_cavello.

R

Rebecca Scott 43:46

I'm Rebecca Scott and this has been Humans, Now and Then. Hosted and produced by Rebecca Scott. Episode notes can be found at humansnowandthen.com. Thank you for listening.