EP. 49: HEALTH EQUITY IN THE AGE OF DATA AND AI

AUDIO TRANSCRIPT

Nitesh Chawla [00:00:00] This is a collective challenge and a collective problem to address. So, we need partnerships, we need voices, we need viewpoints, perspectives to get it right.

Arnab Chakraborty [00:00:17] Hello, everybody. Hope everyone is doing great. We are here with another episode of AI Leaders podcast and I'm thrilled to have Nitesh here with us today. Nitesh is the Frank M Freeman Professor of Computer Science and Engineering, and he's also the founding director of Lucy Family Institute of Data and Society, University of Notre Dame.

Nitesh Chawla [00:00:40] Thank you, Arnab.

Arnab Chakraborty [00:00:41] Great. So, you know, today, Nitesh, that's what we were planning to do is with your expertise and experience around health equity. We wanted to make that as our main topic and really look at the topic from a daytime the lens in terms of what opportunities it can unlock in this piece of health equity. And we have been hearing about the health equity and the fairness and the disparities that are there today, you know, not just in the United States, but across the world. And I would love to start there, you know, in terms of what progress we have made over the last many years on the topic of health equity. And what do you believe are the biggest challenges that are in terms of creating a fair opportunity for everyone across the world to get access to health?

Nitesh Chawla [00:01:24] Arnab, that say, first of all, thank you again for hosting me here and especially hosting me at your office in Chicago, at central office in Chicago. It I must tell the folks listening, it's a fascinating place. It's a delightful place to be here. It's beaming with a lot of energy and excitement. So, hey, so health equity, I think it's an important concept. And we've been talking about it for several decades now. I am not a health professional. I'm a computer scientist who does machine learning AI. But the notion of health equity is important. But because essentially the key definition here is how do we help enable folks to achieve their best potential for health, which not doesn't because we are recognizing health potentially different, my attainment of what it could be is different. However, are there opportunities available for me to attain the best I can? So, if we just start with that notion of health equity, so that means do I have access to resources that will allow me to achieve that sense of attainment and just attainment for my for my health? Do I have access to the health care systems of providers? Do I have access to recreational opportunities? Do I have access to transportation? Do I have access to a place where I'm living which does not have lead poisoning in the household? Do I have access to nutritional food opportunities in my neighborhood? So, if those don't exist for an individual, if that access by itself is limited. Then how do I achieve my best potential for health? I am starting with, you know. A significant overhead ad to begin. So that's what we've been thinking about, looking at it. And it's a complex issue, as you rightly said, And it's not

just in the United States, in the communities and even within the cities or communities of live. You go a few blocks here and there and the health dynamic changes and there are multiple underlying root causes. I you know, even access to education is one, for example. And these things become more profound globally as well as we have seen in our research, where certain things we may take for granted in the US and you look at it in a in a low medium income country and you go, Oh. It's a very different perspective. So. So yeah, so that's sort of is how do we get an individual to attain.

Arnab Chakraborty [00:04:10] It towards you mentioned attaining and access, right? And when I think about all of us in the United States, we all know that depending on the zip codes, you know, things change dramatically. You know, and the whole concept of access to the right facilities, right, health care, even having the right nutrition varies by zip codes. And then, you know, when I look at back in developing countries like India, if you are in one of the rural areas or in in in in obesity or CCD, you don't have access to the best of the health care. So how do you how do you still make that access possible? And to me, it starts creating almost like a, you know, mathematical model saying I have got these resources available in different parts. How do I how do I understand that individual, you know, who is living in that particular zip code on that particular city in that particular area? And how can they make that resource available? And it is a mathematical problem. And today, you know, I feel is an opportunity with all the data that is coming around us and the explosion of data, structured data and structured data, can we take advantage of that and create those, you know, optimizations that is required to attain the health care that an individual needs at the individual level and maximize the benefit? So, you know, I'd love to get your perspective with the you know, you say that, you know, you are a computer science practitioner, right? So, AI, ML background, what opportunities are you seeing today with the AI and ML being activated for health equity.

Nitesh Chawla [00:05:42] Now? I, I think you painted it very well on that. First of all, yes, there are, you know, disparities that one can observe at the census tract, zip code level and one can tie them to, you know, race, ethnicity, demographics. One can tie certain things to gender. One, we know if, say, for example, in the United States, the there is a significant challenge, a black maternal, significant challenge in black maternal health or infant mortality rates, especially for the black population as compared to the white population. So, we know there are these factors that come into play and at the same time, you know, there are. The action of the measures that can be put into place to say, okay, we can address this or we can correct for this. What have we truly understood? What are some of the underlying aspects, whether it's in a rural city in India or whether it's in an urban or rural environment within the United States, or if it's in a place in, say, Mexico, in Latin America, etc... So, if we can and you phrased it well, that is it a data mathematical optimization problem as well. Is it an algorithmic it an algorithmic issue? And in some ways, it is because. I'm not talking about by that by itself, what they may do is set is attainable through data AI. And I want what it what is attainable through data machine learning is an understanding of, let's say if we are in a community, we are in a city in a zip code, we in a neighborhood, we are in a rural area. And you have if and you mentioned the notion of optimization, let's look at it that way. You have a finite budget resources to spend. And if I have data, what other levers I can do? Well, you do have the most meaningful impact is that that creating more green spaces. Is it creating where I'm allowing more physical activity, recreation, more public parks? Is that creating access to more nutritional food, to those retailers, to those retailers? Is it more, you know, access to clinics or health care providers where they are, you know, adjoining or resident? And with a retailer that I go to so it's more

accessible to me? Or is it, you know, making sure that the houses are retrofitted so that the lead poisoning campaign is not happening. So, I can't do it all. So can data. And, you know, an algorithmic AI based formulation tell us, what could be done first that allows us to move the needle a bit and because at the end of the day is we're trying to have an individual attain who access and health care is complex. When we think about, I've talked about personalized health very different than what we may see, say, I go to Amazon.com on Netflix or choose your favorite online retailer, right? One may go or Apple or Spotify or you name it, right? I may go to these places and unless I'm buying a product, Amazon, all these retailers online have dropped a cookie on my laptop on the computer. They are following what I am doing. So as a customer of one, I am trackable. Let's face it, I there are yes, we can do some privacy adjustments and things like that. But at the end of the day we are trackable that a customer of one can be tracked and then, you know, my offers and things that, you know, suddenly something may show up on my social media feed is because something I was browsing or for shopping, etc. that just shows that I can be I can achieve the notion of customer of one from a retail perspective. Now imagine us achieving detailed perspective. Now imagine us achieving that as a patient or one from a health care objective and as an individual. My life is split across different providers, different clinics at home, out, not even having access, not when visiting those. The diets and the habits I have or my specialist I may see. I'm split in multiple different ways and there's no cookie being dropped anywhere to follow me along. So how do we even get to that idea of a patient of one just the way we are used to the customer one. Is a complex problem.

Arnab Chakraborty [00:10:39] Nitesh you have already activated a number of ideas here and the patient of one, you know, that's like the Holy

Grail, right? And creating the information asymmetry across the whole healthcare value chain, I think is a tremendous opportunity that unity like that will unlock value. You talked about, you know, health equity, not in the context of just the patient, but the entire ecosystem, you know, where human beings, all of us live and what interventions can happen in the community. And that creates interplay across different agencies and organizations to collaborate. And how can data inform, insights inform for you to help improve the overall community, because that has a direct impact on the health of the individuals living in those communities. So, you touched upon a number of areas and I was thinking if you could share, given your role in Notre Dame and what you are already doing in the state of Indiana, but in the larger community on this topic, maybe a couple of examples of projects that your organization is already initiating and working on.

Nitesh Chawla [00:11:38] Absolutely. So, recognizing that. So, for the last several years I've been immersed in the health and wellbeing space and research with whether it's on personalized health care journeys or through variables. And you know, you've published papers that have talked about the function of a social network and the function of, you know, a sense of well-being and in everything from job performance to our happiness to a sense of attainment of happiness inside, that we all are tied up quite closely. But I've been struggling with the fundamental question of disparities and inequities, because if you look around, there are certain things that, you know, it's easier. You know, I could get my variable and track my stats and do my, you know, do my EKG on my watch now or do my, you know, heart rates and look at what my resting heart rate is. That requires A should be able to afford that community B access to internet connectivity and things like that and C time to be able to do these things and more so some notion of literacy as well, to understand these things. What does it

mean? And. And I'm talking about not just literacy, but also health literacy. So, all of these factors come together. So. So we having, you know, cut conditions deeply in anyone the mental health paradigms that we see here. We decided to launch a health equity data lab in the Lucy Family Institute that you mentioned earlier last year. And the idea was that and there's a reason we called it health equity data. We didn't call it healthy. Health equity is such a broad and we wanted to, as you mentioned in your opening remarks, that. What data? The role data has to play. The role algorithms analytics have to play in not just saying, okay, this will solve the health inequity, but being awareness, understanding and something that we talk about in the analytics journey all the while, as it starts with getting the data insights and the insights inform us, then the insights inform action. That sort of is the playbook and you can get insights, but there's no action. You know, analytics if you try and if you don't get good insights, there's no informed action. So, what can we do to apply that to really understand? So, we did we launched that lab in last year fall, and then we had the good fortune of Accenture partnering with doesn't really bringing their own leadership and partnership in this space. And amazing, it's Accenture, you know, who it's leadership. And that just shows to us the leadership that Accenture has. Is it hands on said, you want to partner with you. Let's go. This is an important topic that we can address and we can only address as a collective. And then. This all happened really quickly. As I said, we launched the lab in fall. By December, we had this amazing partnership with that, except we decided to host a first forum. And the idea of this forum was let's bring in thought leaders from different, you know, corporations and builders and providers and, you know, help departments of health get them all together. Let's get them all together in a room and convene around this topic of health equity data. Let's ask each other what can be done

and why some of these corporations and amazing partners in this process, including, you know, as we think about the HR within these organizations, etc., these corporations are among sometimes among the largest employers in a city or a community. So, actions that may. Flow through to the employees may go from that individual to the family and from the family to the community and perhaps family to the community. And perhaps that becomes a movement of change. Worse is taking a very macro approach that we go to a community or a county or a city or an or a neighborhood or a zip code with ease. So, we are starting to think about that. What does that mean from that individual perspective? So. So that's an example of what we are trying to do is it's starting with the dialog, starting with the key partners. The key stakeholders in the room were interested, and it was evident through that day that the everybody cared about this too much. Everybody, you know, we had some of the amazing leaders in this space in the room with us who were in that because they believe in it. They want to solve it. So, so we are excited about what that may lead. So that's one example. I have many others I can share.

Arnab Chakraborty [00:16:44] Not this.

Nitesh Chawla [00:16:45] Unless you want. To drill.

Arnab Chakraborty [00:16:46] Into that, because first of all, you know, very proud of the partnership with Notre Dame and Accenture on the whole Health Equity Data lab. And I had the privilege of being there with you and when we did the event and a number of industry leaders and I remember it was industry leaders coming in from the peer provider space, as you said, there were retailers actually there. Right. And there were policymakers and academicians. So, it was really and I think a few things came to mind as we were discussing about the

opportunities as to what comes in the way, you know, as we are on this journey, on this endeavor with this passion to create equality for every individual, for access to health. I think it boils down to, you know, collaboration, you know, across the whole healthcare value chain. That was a big topic. And, you know, you already brought different parts of the value chain together. So that was very evident. Everybody was voicing that, that we need very strong collaboration across the healthcare value chain. And one thing that comes in between that is the privacy of data and how we are able to and how we are able to, you know, take care of the data privacy, especially as we think of data being shared across organizations for better patient outcomes and how do we enable that? There was a whole topic around responsibility, I think, as we develop the algorithms to make decisions, how are we taking care of the responsibility behind these decisions and not leaving it to the algorithms? And then there was a big topic around talent and the readiness of that talent. And you mentioned about literacy. You know, how do we actually elevate the literacy, you know, around data and digital across the organizations. So, I wanted to bring that together for you as well as we are working through the Health Equity Data lab in these challenges, some practical solution ideas, you know, around these challenges.

Nitesh Chawla [00:18:30] So I know I agree with you on that as we think about health equity data. One thing is understanding advances about this problem. Absolutely. So that's the role that data can play. And what was amazing is that you all settled in the forum as a priority, that once there is this data, we have to have a strategy and an approach for responsible use of this data, for responsible development of the algorithms, AI algorithms for responsible deployment of the news, because we could quickly this could become a Pandora's box. If not, you know, cut to it from the beginning. I just know this ethics and a sense of social responsibility has to is not in

the purpose of us working on health equity. That is a very inclusive topic. But at the same time as technologists, what we can sometimes fall victim to is we get enamored by the technology and want to build the next big thing that you know is going to be a technological marvel. But. We need to pause and reflect before we start building these things. We need to develop some guiding principles so that notion of responsibility AI is absolutely paramount. And then something else that the notion of literacy is twofold. One is the literacy of the new workforce that's joining the field of medicine, because sometimes those encounters could be a challenge, that those encounters could not intentionally. But those encounters good in American heights. All right. I do need to go back to that physician. There's something how I felt. And so, what does that mean? What does that mean? And even not just from a health equity perspective, but also training of the workforce on how this data AI would be is because it's here, it's not going anywhere. It's only going to increase more. And it's and the whole business of health, which includes, as you rightly said, retailers be as providers, patients, policies, makers and academicians, citizens all have to all either party. So how do we sort of create that? And then the other literacy as well is Volkswagens at the consumer end of health care. It was a time when, you know, we used to talk about that financial literacy is important. Every kid, whether that finishes high school or doesn't finish high school, must know what you know. For some of the Financial Times, that repetitive balancing check accounts and things like things like that. I believe they would be now the same idea would fall in data literacy awareness of how my data, what it could do, what it may be used for, and as well as health literacy. We have not paid attention to that. So, to that. So that's another thing that's going to be absolutely important. And then something else that was apparent in the forum and is becoming a key topical focus for us as well, is as we mentioned earlier in our conversation, that health is a function of so many different dimensions, what we call the social determinants of health, whether it's, you

know, where I live or work, you know, my access to food, my access to health and many other factors. But what we realize as well is it's the social determinants of health is a concept that's been around and folks have talked about that, how that influences more than what your genes influence on your health. But why haven't we done much? It hasn't moved forward as much as a concept in our practice of health and health care both itself. So, this idea of what we talked about is precision. So, we need to be way more precise. What does it mean for you, your family, your community in as a social determinants of health, a given example. You sort of support that view. Working on this project in the city of South Bend on childhood lead poisoning and the source of lead poisoning is to paint in the house. It's not the water as consistent as much as the pipes or things. It's the paint in the house. But there's not enough awareness. If I am as a tenant to go into a house, I don't know that this house has a paint poisoning. And then there's all this policy that's complicated, as can communication between, you know, first, if I screen that is led in the house, I have to get it diagnosed and then I can apply to the city for a mediation grant. But then the landlord has to sign up to do it. These are complexities and all these complexities. I as a family, I go, how do I navigate this? I don't have the literacy to fill out that form or do some legal things, etc. and as a result. The child whose struggling could be struggling from a health inequities challenge to begin with is now has a developmental delay with that boys. So, these are and then the house, which is the source of this boy's lead poisoning, doesn't have a place in the electronic medical record. It is not EMR that says the source of this is the house or the house should be recorded. We only track the people, so we have to change the way we frame. This is the data play again changed the way we frame our electronic medical records of systems out of recordkeeping, where these social determinants of health and attached social determinants of health are attached to the persons and also maybe live individually because they affect people. As a collective.

Arnab Chakraborty [00:24:43] I think that is a great example. You know, you just can't rely on lead poisoning. And you know what data AI and technology can enable. And, you know, this triggers me at heart, which is very important in today's conversation, you know, especially with the frenzy around ChatGPT and generative AI. No conversation is truly complete without ChatGPT.

Nitesh Chawla [00:25:05] Absolutely. Now, Hey, hey, hey, now.

Arnab Chakraborty [00:25:07] So, you know.

Nitesh Chawla [00:25:08] And by the way, for the listeners, this broadcast is just two generative AI bots talking to each other.

Arnab Chakraborty [00:25:14] That it is it is really fascinating. You know, and I would love to kind of, you broadly close out our discussion with your perspective. But, you know, let me just share an anecdote. In in in in January, in Davos, you know, when Satya Nadella was talking about, I was talking about the generate AI I and the example they decided was kind of connected with health equity. You had an engineer out of India, took the, you know, open API models and trained it on, you know, hundreds and thousands of government papers and policies. You know, back in India. And this was done in few weeks and was made available as an API that a rural farmer who is looking for some of these government directions for services that he wanted to consume can now can access that in the local dialect because of the large language models that is making it possible to transcribe and translate into local languages, the policies. And by the way, this shows how technology moves from West Coast back to a rural village in India within a matter of few weeks and makes any information intelligence available for a rural farmer and to consume those services that otherwise was not possible. So, I would love to kind of give you a perspective in generate AI special and consume those services that otherwise was not possible.

So, I would love to kind to get your perspective into generative AI, especially in the health care field. And as we think about healthcare equity, health equity, what do you think are the big opportunity that organizations should be looking for?

Nitesh Chawla [00:26:40] I think generative AI has been, you know, absolutely you know, a phenomenal achievement. And indeed, in the limited time that we have, you know, in about four or five, it's been about since 1718 when India, you know, we had first published paper in the academic community about talking about, you know, attention is all you need to transformers model of sort of, you know, forms the basis of these things as well. And now we have seen this open AI you know revolution or you know, the GPD for now it hits a billions of parameters. Who knows what else this could this is just what we have seen. We don't know what example that you mentioned. And then there are others where folks are looking at it for drug discovery and navigating chemical reactions. And it's active. It is How do we now bring the world's knowledge to four to do meaningful research? Because as a researcher, if you can if a tool can look at all the published literature, I may be leading literature in my domain of the machine learning sciences. But imagine if things that you publish an econ the stats or social sciences and speaking it all together from you and then creating, you know, the first draft, if you will, of thinking and even creating these mechanisms where a lot of coding is done. So, prioritization of ideas would be, if not the acceleration we would see, in my opinion, generative AI is productization of ideas would be much faster. Of course, that would be much faster, of course access to the systems to do these things, etc. I'd still would be a large platform play and figure those things out, but it'd be rapid modernization. So, it's fascinating what the what we are see with generative AI. But there's also a limitation with these large language models that what we have seen so far, they can they are representative of the training data that they are seeing. Yeah, I said what they have seen is what they know. It's like, you know, if I teach you from two books and those two books alone and I don't allow you to get information from any of the sources, you're an expert in those two books, right? Nothing else. Right. You can connect the dots between words in those two books and come up with new sentences that sound phenomenally intelligent. And that's what you experience sometimes with these tools as well. So, there's this notion of the biases that may be in and as a result, are shown in the output created by tomorrow. And that's why we also talk about, you know, there's this hallucination effect where it sounds so conscious, right, that you go must be right, but maybe not. So, as we think about the field of health care, I believe there is a next gen of LLMs That would be needed. I would be possible in my lab. We are doing research quite a bit in this space. It's a couple of directions that we are exploring. One is how do we navigate the data to establish the provenance, credibility, validity of the data to begin with. It's not something that I'm doing for a fund to create new art or create new music or create new, you know, meme or create new content or etc... This is something that is something that a person has to consume and be it meaningful for that person. So could we first figure out how do we bring. The validity counts if you think that's a going to be a job. Second is given to think about it that if I create a question that I hypotheses. Now, there could be multiple different answers to it. It could be some answer that did go. Natasha, you're going to die tomorrow. Do I? And then they would say, oh, don't worry about it. Life is good. There are contradictions because it's symptoms could manifest itself in a million ways. So how do we now. Do me put these elements in steroids where there's a mechanism for gathering evidence, for rating evidence, reading some counterfactuals and providing reasoning to the individual's ideal health consumer using these things. So, there are these we have we know for a fact that, you know, the advances that we've seen in AI, in medical imaging and other things, you know, a phenomenon where there's still a physician isn't

talking about if an individual is interacting with these tools or at home. You have to be the guys. And this comes back to the idea of the responsible AI, that we talked about it is how do we build these things for the health and health care consumption? And it's important that, you know, Microsofts of the world and academics, Accenture and retailers and health care providers, systems are working together to validate what goes in. Otherwise, we can't control what goes comes out.

Arnab Chakraborty [00:32:09] Yeah, no, but this is fascinating. This is fascinating just to hear from you. And this is fascinating just to hear from you. And, you know, I just I just want to echo that. The health equity data lab that Notre Dame has set up with the participation of the industry I think is a tremendous opportunity for us to experiment. You know, these use cases and to your point, you know, experiment them with guardrails. Right. And also establish guardrails based on the learnings from these use cases that we will be experimenting with inside the data lab. So, I just want to call that out. And, you know, I request everybody who is listening to this podcast to reach out to Nitesh myself. And if you want to engage with the health equity data lab, you know, we'll be more than happy to have a discussion on that as well. So, it is. Thanks a lot for a great conversation and really appreciate you taking the time.

Nitesh Chawla [00:32:59] I thank you Arnab and thank you again for hosting me at this fascinating place in the studio. And I would echo what Arnab said just reach us each discuss is any questions, things you have, because this is a collective challenge and a collective problem to address. So, we need partnerships, we need voices, we need viewpoints, perspectives to get it right. And let's move the needle.

Arnab Chakraborty [00:33:24] Wonderful. Thank you so much.

Nitesh Chawla [00:33:26] Thank you.