

Sean Gallagher:

Okay now, please help me welcome to the stage Lee Hutchinson, our Senior Technology Editor. Lee, come on down.

Lee Hutchinson:

Thank you, Sean.

Sean Gallagher:

You're welcome.

Lee Hutchinson:

Much appreciated. Hey, everybody. Okay, we're going to pivot a little bit away from the world being on fire and security being a disaster, to talking about machine learning a little bit. Machine learning continues to push the frontier of what's possible. Now, we are going to have sort of a little fireside one-to-one chat with Amazon Web Services Technology Evangelist, Dr. Nashlie Sephus. Dr. Sephus, Tech Evangelist at AWS. She focuses on identifying biases in AI. She formerly led the Amazon Visual Search team as an applied scientist and in 2018, Dr. Sephus founded the Bean Path nonprofit. And in 2020, she established the Jackson Tech District to bring tech training and workforce and economic development to the area. Please join me in welcoming Dr. Sephus.

Lee Hutchinson:

All right, question number one, machine learning is great, but as with all technology, not everyone has equal access to machine learning. Can you give us an overview of why there are barriers to ML accessibility and what those barriers are?

Nashlie Sephus:

Yes, absolutely. I do want to say, thank you all so much for having me today. I'm excited about being here and seeing everything that's going on. In terms of access to just technology in general, I mean, we've seen during the pandemic especially, how many people don't have access to the internet or even have capability of getting access to internet. We've seen a lot, this really impact youth in schools if you think about our K through 12, even just the public school system, especially in a lot of rural areas. And I'm originally from Mississippi, and so I've seen it firsthand, a lot of kids drop out of school just because they don't have access, maybe not have had access to laptops. And the school system not being able to supply them with that. And so there's a lot of things going on there, even before the pandemic.

Nashlie Sephus:

You think about there being a lack of access to these type of tech experts in certain communities in the United States and also access to the technology itself. And so fast forward that to machine learning, you're talking about a concept that a lot of people think is pretty intimidating. They're intimidated by it. A lot people are scared. They feel threatened by the technology. We've seen a lot of issues in terms of privacy. We talked about that today earlier. Other issues about possibly invading someone's rights with technology that may not be as accurate. And so a lot of times we don't have people in the community to help people understand what these issues are, how to overcome them, how to be aware of them, how to look out for them. So there's a big access issue. If we focus on that alone, we could go all day.

Lee Hutchinson:

Right. I mean, in some ways, this sounds sort of like the same issue as the digital divide that we heard about in the past few years. It's an old and unfortunately familiar problem at this point. Are some of the same issues that prolong the traditional digital divide issue also in play here when it comes to trying to get access to machine learning?

Nashlie Sephus:

I believe so. I think, because I'm also a consumer of this technology as well as a technologist, a developer of the technology, I see and understand both sides. I understand that, as technologist, a lot of times we just like to build cool stuff, right? We're not thinking about sometimes the longer term impact, but that's why it's so important to have that diversity of thought at the table and those different perspectives. I heard someone mention hiring people in the humanities. That is actually something that we're doing at AWS. We're hiring sociologists and psychologists to join our tech teams, to help us understand what is the long term impact of some of these things. And in terms of the digital divide, understanding how do we meet these people where they are.

Nashlie Sephus:

Machine learning is a pretty complex intimidating topic, but you can break it down into very simple ways for people to understand. Everybody has a cell phone for the most part, where everybody at least understands that these technologies exist. When you're talking to your phone, when you're unlocking your phone with your face, when you're trying to get recommendations to watch something on a streaming platform, this is AI, this is machine learning. A lot of people, they use it every day, they don't know what it is. And so kind of meeting them where they are, showing them how these technologies affect them in their everyday lives and having programming out there in a way that's very approachable. I think that's something that we should focus on.

Lee Hutchinson:

Do you find that the ways and the amounts that people use machine learning and AI types of technologies breaks down along socioeconomic lines? Is this something where like a predominantly minority community, a community of persons of color, are less likely to be utilizing AIML tools?

Nashlie Sephus:

I think so, I've seen that firsthand. That's one of the reasons why a lot of people say, "Well, what is the tech evangelist at Amazon?" It took me maybe a year to figure it out as well, but it's essentially outreach and education. And so bringing these topics to places where we normally wouldn't go at AWS. I was an applied science manager for four and a half years at Amazon, and moved over to this role last year, because I felt like we needed someone to fully represent in a more diverse way what this technology is and how do we meet the people. And so I wanted to make sure that, in terms of having talks like this for example, coming to DC. We have, at AWS, our second headquarters is now in the Crystal City area. And so now more than ever, it's so important for us to reach the DMV communities near and far, all backgrounds, ages, ethnicities, and understanding, because we actually need feedback from people in order to make products that work for people. And somewhere along the way, sometimes people lose that.

Lee Hutchinson:

We're going to come back to this, because I'm going to ask you some questions specifically about the nonprofits that you've worked on and founded. But real quick, there's another side to improving access to machine learning, kind of switching away from people and looking at organizations too. There are plenty of situations where businesses can realize huge improvements by applying AIML to business problems. Do you find much institutional resistance at companies to embracing this brave new world machine learning and AI?

Nashlie Sephus:

I think people want to. They definitely want to do it, because I don't think there's much resistance there. They just don't know how a lot of times. I know in general, if you think about just cloud computing, for example, which shameless plug, AWS has some greats cloud computing, check it out. But especially if you think about in the government world, in the public sector, a lot the times things are not in the cloud. Some may think of the current processes as sort of antiquated, but that is a culture change that the leadership has to adopt, the grassroots has to adopt, in order for those things to happen. I think the same is true for... If that even should happen, because there's many cases where it probably shouldn't, to actually ensure security at a certain level.

Nashlie Sephus:

But in terms of machine learning, I think that same mindset has to happen. You have to have buy-in from leadership. You have to have buy-in from those at the grassroots level. And again, that just doesn't come if people aren't aware of how the technology works. And I know that there's a lot of buzzwords out there, AI, cyber security and the metaverse, but what does all that really mean at the end of the day? And so understanding that we have a responsibility as a large company to help people understand that, because it does us no service to sell products that people don't understand.

Lee Hutchinson:

Mm. Yes. Building on this a little bit. There's a difference in how a small business builds an AIML process when it's small versus a big company with big data needs. How do you create, maintain and scale AIML processes sanely without introducing inadvertent bias? And I mean, at scale, like when you're talking about a big installation, that's going to make big models. How do you do that?

Nashlie Sephus:

Yeah. So there's a lot that goes into just preparing data itself in terms of before you even get to actually training the AI or machine learning model. That includes thinking through a lot of questions and understanding what's the worst possible thing that can happen, and work backwards from that to make sure it doesn't happen. One is, are you adequately supplying a evenly distributed diverse set of data that represents the people that your technology is going to be, not sold to, but used on. And that's a big difference there. So the customer is not always the same as the person that the technology is being used on. And understanding that we have to educate the customers even on proper uses of the data that they're using to train these models with.

Nashlie Sephus:

I think, secondly, you shouldn't just take publicly available data sets for face value. There could be a lot of issues with the data. There could be mislabels in the way the data is set up and arranged. You'd be surprised what we've seen, and just going back and looking through some of the most commonly sold data sets out there, especially those dealing with humans, which are very mission, we call mission

critical. And so it may be an opportunity to create your own data set if that's needed, if that's necessary. And then even when doing that, you want to make sure you're ethically sourcing that data. And in my opinion, I think that involves if you need to get data from humans. Let's think about how we can adequately compensate them for that, because their data is a huge part of bettering and improving our product. And so that's something that we don't see a lot of, but again, it all starts with the data. That's the most important thing, garbage in, garbage out.

Lee Hutchinson:

Right. And I want to expand on that just for a minute and discuss bias in some of these tools and sets, because expanding access feels like potentially, that's only half of the issue, especially if we're expanding access into underserved areas, potentially minority communities, communities that are predominantly persons of color. If they're using tools that have biases built into them, then the tools aren't necessarily going to be doing potentially what they think they're going to be doing. Is that a good capture?

Nashlie Sephus:

Yes. Well, there's a lot to be said about potential biases. If you think about a product that has many different features, like for example, in the FinTech world, there are algorithms that recommend whether or not someone should be approved for a loan. That's a lot more critical than recommending the right show on Netflix or not. And so for those particular use cases, we could spend our time digging into every single use case and if you have the ability and the resources to do that, then by all means do it, but oftentimes we don't. And so let's focus on those most important use cases to understand what's going on and then understanding for every single feature that is available, it can be from the smallest thing to the largest thing, it needs to be tested. We cannot assume until we test.

Nashlie Sephus:

And there are proper ways to go about that, to be able to compare apples, to apples and oranges to oranges and understanding if those tests are repeatable or not. They should be reproducible. There are so many areas of biased testing and bias detection that are emerging now in the research. And you're seeing papers being published on these things, and it's not one right or wrong way to do it. A lot of times people say, "Oh, well, what's so hard about testing?" Well, how are you going to test? Where are you going to get the test data from? Because the test data has to be different than what you actually use to create the product. And oftentimes people don't put that much thought into it. So it really should be something that should be thought of at the product conception stage rather than after you've deployed the product. But yeah.

Lee Hutchinson:

Are biases like this inherent in the tool? Are we always going to have them because this is people making them, or can we build around them at the very foundation of creating the tools in the data sets?

Nashlie Sephus:

There absolutely has to be a trade off at some point, because you're essentially trying to trade off between bias metrics and accuracy metrics. And sometimes you can't get both to be what you want them to be and so you have to make a design decision on what that is, and that's a very difficult decision to make. That's why it's so important to understand what is the use case, who is the customer, who are the users, again, what is the worst thing that could possibly happen? And let's work backwards from

that. I think that all the use cases again, are important, but I think one other thing I'll mention is that oftentimes when you're training these models, you have something called a confidence score.

Nashlie Sephus:

So every machine learning model, if you're training in a traditional deep learning or machine learning way, you have this thing called a confidence score. And if you understand, like make some rule based on, okay, if it's above this confidence score, then I will use this reply. If it's below it, I will not use it. So if it's below 80% confidence, then you know, only based on that sample set that you've done this testing on, and this testing has to happen iteratively, it's not just a one stop shop. And there are other rules that you can come up with too. But again, there's so much involved in just how do you even test this system and come up with a test that is going to be used over and over again.

Lee Hutchinson:

One more on biases while I'm thinking about it. When we talk about this, are the biases that you see, do you think are they typically the result of the biases of the people who are building the tools and sets? I guess what I'm getting at, is this a process problem, or is this an endemic problem?

Nashlie Sephus:

Well, I always tie it back to the diversity of the tech teams. And if there aren't diversity... Even if you aren't able to hire a diversity team, at least reach out to focus groups, there are lots of options out there. There are universities and schools and community organizations that you can lean on to get that information and that feedback from. And I will say too, that sometimes you may want a bias system. It depends on, if I'm in a particular area or location, it's only going to be this certain demographic of people. Then I want to tune my algorithm to where it's extremely accurate for that particular intersection of demographic. So that may be an example where you do want bias, but you still have to be as objective as possible and the only way you can do that is to have a system that will allow you to see all of those metrics. It's not just one person, train it and it's saying, "Hey, this is good." You have to actually show and prove that you've done the testing to show that it is good.

Lee Hutchinson:

This is kind of a difficult question maybe to answer, but what's your criteria for success to determine that you've reduced biases to an acceptable level, I guess, or is there one?

Nashlie Sephus:

There are several metrics, I know for those who train models, there's a tool AWS has called SageMaker and basically, if you're a machine learning or a data scientist, it allows you to look at your data, your training data, and it automatically shows you some metrics and some balancing suggestions on that data. It also allows you to monitor how well that algorithm is performing over time on your test data set, even after it's been deployed, because sometimes you can have something called model drift where the environment changes, your model stays the same, but now your environment changes. So some of the assumptions that went into developing that model are not withstanding. And so it just helps you to better just visualize everything. And so we're a big proponent of just looking and visualizing as much data as possible. That's one way to do it. Clarify is actually the name of that tool. SageMaker Clarify. And so that's kind of what I would suggest.

Lee Hutchinson:

Okay. So switching back away from companies and back to looking at individuals and communities, can you take us through some of the outcomes you've seen when AIML access is opened and people can use tools they otherwise wouldn't have had access to?

Nashlie Sephus:

Yes. So we've all seen during the pandemic, I don't know if y'all have grandmothers and aunts who now use QR codes at restaurants. Now we all know what Zoom is, because before we're like, "Zoom, what is that?" And so I think that's a perfect example. I probably couldn't pay my grandmother to use Zoom before the pandemic.

Lee Hutchinson:

Yeah, my parents did-

Nashlie Sephus:

But now it's a necessity, right? So necessity brings about all these changes and ways of life that are now, it's such a different world, which I actually appreciate. Now we're able to do so much more, I think in terms of... And I think, we talked earlier about, I have a nonprofit in Jackson, Mississippi called the Bean Path and we try to bridge the tech gap. We do all sorts of tech exposure and education. We saw this shift during the pandemic as a blessing and a curse. One, because now we could do all of our programming virtually and reach people across the world. But we lost connection with the people that we wanted to reach the most, the people who will still probably never get on a Zoom call and it's not as easily accessible to them unless they get up and go somewhere that has internet.

Nashlie Sephus:

And so, whereas we used to be in person, we did kind of lose that demographic over the pandemic. And so there are definitely still barriers and access barriers, but it takes several different, I call it the tech ecosystem, in order to resolve those issues. And tech ecosystem is made up of technologies. It's made up of policy makers. It's made up of government. It's made up of the K through 12 education system, as well as the higher learning education system, the larger companies, the smaller companies, everybody has a responsibility and everybody plays a part in ensuring that no one is left behind.

Lee Hutchinson:

Mm-hmm. Looking at Jackson as an example, you said that sometimes you have areas where people actually don't have good internet access, they can't participate in Zoom calls. This may be kind of an obvious question, but going back to the digital divide thing, access to AIML tools, I mean, effectively rides on your access to internet services, right? Do you still see a lot of hugely underserved areas there?

Nashlie Sephus:

I do, unfortunately, but I think it is better than where it used to be. At least there's awareness around the issue. And there are a lot of potential solutions in the works. Again, sometimes people just don't understand how to solve it, but they're at least trying to. I've seen now there's a lot of the utility companies such as the power and light companies are now able to add internet access as a part of their offering and their services. And a lot of them have grants and waivers for people that fall below a certain income amount within certain school districts. Within, for example, in the southeast, there are... If you're a part of this particular public school system, you get access to this citywide broadband for free in

order to log on and be able to use the internet to do homework, et cetera, have work from home days and things like that.

Nashlie Sephus:

So I think... And one other thing I've seen is the wifi parking lots. So they'll go to a local park or even a school and just set up this huge broadband system there, where people can come in and plug in their laptop in the parking lot, sit in their car or ride your bike, and now you have access to the internet or pavilion at least where you can sit and get some work done.

Lee Hutchinson:

Okay. We have a couple of minutes left here. So I wanted to ask a little bit about, I noted in your intro that you had established the Bean Path, the nonprofit, you mentioned a second ago, and also the Jackson Tech District. In the time we have left here, which is about five minutes, can you tell us a little bit about those projects and closing the tech gap as it were?

Nashlie Sephus:

Yeah, so I always say I never really... I have a bit of a different story getting to Amazon. So I was actually the CTO of a startup company that we sold to Amazon in 2016. It was called Partpic based in Atlanta, and I'm originally from Mississippi and so I said, when that happened, and I've been at Amazon over five years now, so when that happened, I first joined. I was like, "Okay, now it's time to go back to my hometown and see if I can have some impact or give back since I've been somewhat successful in the tech industry."

Lee Hutchinson:

I have to interrupt just a second, because you mentioned Partpic and I was Googling it, because I was trying to be a good interviewer, so I could research you before we talked. Partpic had a really cool product that also used AIML and image recognition. Can you, just like 30 seconds, describe that because it's totally neat.

Nashlie Sephus:

Yeah. So we created this machine learning algorithm where you would take a picture of a replacement part like a screw, a nut, a bolt or a washer, anything like that, hundreds of parts, and we would identify the parts. So not only identify this is a hex bolt, but also this is a hex bolt, two inches long, half an inch based diameter, three inch per inch 15, and you can buy it at this store. And so we got you the exact part, and I was actually presenting at a conference in Boston in May of 2016 and Amazon was in the audience. And so-

Lee Hutchinson:

I was going to say if anybody's thinking that sounds like a cool idea, amazon thought so too, and then they bought you.

Nashlie Sephus:

Who'd of thought that we're definitely bigger than my dreams ever could imagine, but he was in the audience and we exchanged cards and one thing led to another, we ended up selling the company. And so with that capital, I went back to my hometown at Jackson, Mississippi. I started the Bean Path

nonprofit. So we've helped over 1,500 people in the last three plus years, gained exposure and access to technology and emergent technology, things that should be household names, but oftentimes they aren't in certain communities. And so we were working out of the local libraries, and so I wanted to scale, one thing Amazon has taught me is how to scale. And so I said, "Well, hey, we need our own space." And so I went downtown, I found one building. I found out the same owner owned everything around it. And I purchased, I have a total of 14 acres of land and eight buildings now in downtown Jackson. And we're creating the first Jackson Tech District in downtown Jackson, Mississippi.

Lee Hutchinson:

Excellent. What is... Yeah. And for the members of the community that are close to it, or if you were to visit, what can you do there?

Nashlie Sephus:

Well, it is a new development project, so I'm adding real estate developers to my resume. But we are opening our first building, which is a makerspace. So people will have access to 3D printers, laser cutters, woodwork, computers, things like that. So again, it's all about how do you bridge that access to these technology and this programming, learning about AI and NFTs and all these things, cloud computing for the average everyday person. And so we're opening that this summer. We have a big summer camp we're starting. So the rest is, sky's the limit from there.

Lee Hutchinson:

Yeah, it's neat to think about all the things that you could potentially do there that folks can do when they visit. But I think the cooler thing are like all the things that we can't think of, because there's going to be some smart folks who show up who you've given access to to all of these tools who are going to be able to make things that we can't even guess at.

Nashlie Sephus:

Yeah. Yeah. I love it. I think it's really important to me. It's something I'm really passionate about. It actually goes hand in hand with my title as tech evangelist at AWS. And they've been extremely supportive of a lot of these things as well.

Lee Hutchinson:

Excellent. Well, Dr. Sephus, I appreciate you coming by today. Thank you very much for sharing your wisdom with us. Thanks a lot.

Nashlie Sephus:

Thank you all.