

[LOGIN](https://twimlai.com/login/)(<https://twimlai.com/login/>) [ABOUT](https://twimlai.com/about/)(<https://twimlai.com/about/>)

=

[\(https://twimlai.com/\)](https://twimlai.com/)[\(https://twimlai.com/article/responsible-ai-in-practice-with-sarah-bird-transcript/\)](https://twimlai.com/article/responsible-ai-in-practice-with-sarah-bird-transcript/)

Responsible AI in Practice with Sarah Bird (Transcript)

December 4, 2019

SHARE:



Sam Charrington: Hey, what's up everyone? This is Sam. A quick reminder that we've got a bunch of newly formed or forming study groups, including groups focused on Kaggle competitions and the fast.ai NLP and Deep Learning for Coders part one courses. It's not too late to join us, which you can do by visiting twimlai.com/community. Also, this week I'm at re:Invent and next week I'll be at NeurIPS. If you're at either event, please reach out. I'd love to connect.

All right. This week on the podcast, I'm excited to share a series of shows recorded in Orlando during the Microsoft Ignite conference. Before we jump in, I'd like to thank Microsoft for their support of the show and their sponsorship of this series. Thanks to decades of breakthrough research and technology, Microsoft is making AI real for businesses with Azure AI, a set of services that span vision, speech, language processing, custom machine learning, and more. Millions of developers and data scientists around the world are using Azure AI to build innovative applications and machine learning models for their organizations, including 85% of the Fortune 100. Microsoft customers like Spotify, Lexmark, and Airbus, choose Azure AI because of its proven enterprise grade capabilities and innovations, wide range of developer tools and services and trusted approach.

Stay tuned to learn how Microsoft is enabling developers, data scientists and MLOps and DevOps professionals across all skill levels to increase productivity, operationalize models at scale and innovate faster and more responsibly with Azure machine learning.

Learn more at aka.ms/azureml (<https://aka.ms/azureml>).

All right, onto the show.

Sam Charrington: [00:01:52] All right everyone, I am here in Sunny Orlando, actually it's not all that sunny today, it's kind of gray and gray and rainy but it is still Sunny Orlando, right? How could it not be? At Microsoft Ignite, and I've got the wonderful pleasure of being seated with Sarah Bird. Sarah is a principal program manager for Azure Machine Learning platform. Sarah, welcome to the TWIML AI Podcast.

Sarah Bird: [00:02:15] Thank you, I'm excited to be here.

Sam Charrington: [00:02:17] Absolutely. I am really excited about this conversation we're about to have on responsible AI. But before we do that, I'd love to hear a little bit more about your background. You've got a very enviable position kind of at the nexus of research and product and tech strategy how did you create that?

Sarah Bird: [00:02:37] Well I started my career in research. I did my PhD in machine learning systems at Berkeley and I loved creating the basic technology, but then I wanted to take it to the next step and I wanted to have people who really used it. And I found that when you take research into production, there's a lot more innovation that happens. So since graduating I have styled my career around living at that intersection of research and product, and taking some of the great cutting edge ideas and figuring out how we can get them in the hands of people as soon as possible. And so my role now is specifically focused on trying to do this for Azure Machine Learning and responsible AI is one of the great new areas where there's a ton of innovation and research, and people need it right now. And so we're working to try to make that possible.

Sam Charrington: [00:03:33] Oh, that's fantastic. And so between your grad work at Berkeley and Microsoft, what was the path?

Sarah Bird: [00:03:42] So I was in John Lankford's group in Microsoft research and was working on a system for contextual bandits and trying to make it easier for people to use those in practice, because a lot of the times when people were trying to deploy that type of algorithm, the system infrastructure would get in the way. You wouldn't be able to get the features to the point of decision or the logging would not work and it would break the algorithm. And so we designed a system that made it correct by construction, so it was easy for people to go and plug it in, and this has actually turned into the Personalizer cognitive service now. But through that experience, I learned a lot about actually working with customers and doing this in production, and so I decided that I wanted to have more of that in my career.

And so I spent a year as a technical advisor which is a great role in Microsoft where you work for an executive and advise them and help work on special projects. And it enables you to see both the business and the strategy side of things as well as all the operational things, how you run orgs and then of course the technical things. And I realized that I think that mix is very interesting. And so after that I joined Facebook and my role was at the intersection of FAIR, Facebook AI Research and AML which was the applied machine learning group with this role of specifically trying to take research into production and accelerate the rate of innovation.

So I started the Onyx Project as a part of that, enabling us to solve a tooling gap where it was difficult to get models from one framework to another. And then also worked on PyTorch and enabling us to make that more production ready. And since then I've been working in AI ethics.

Sam Charrington: [00:05:34] Yeah. If we weren't going to be focused on AI ethics and responsible AI today, we would be going deep into Personalizer, what was Microsoft Decision Service and this whole contextual bandits thing. Really interesting topic, not the least of which because we talk a lot about reinforcement learning and if it's useful,

and while it's not this deep reinforcement learning game playing thing, it's reinforcement learning and people are getting a lot of use out of it in a lot of different contexts.

Sarah Bird: [00:06:05] Yeah. When it works, right? It doesn't work in all cases, but when it works, it works really well.

It's the kind of thing where you get the numbers back and you're like, can this be true? And so I think it's a really exciting technology going forward and there's a lot of cases where people are using it successfully now, but I think though there'll be a lot more in the future.

Sam Charrington: [00:06:25] Awesome. I'll have to take a rain check on that aspect of the conversation and kind of segue over to the responsible AI piece. And I've been thinking a lot about a tweet that I saw by Rachel Thomas who is a former guest of the podcast, long time friend of the show and currently the UCSF Center for Applied Data Ethics head. And she was kind of lamenting that there are a lot of people out there talking about AI ethics like it's a solved problem. Do you think it's a solved problem?

Sarah Bird: [00:06:58] No, absolutely not. I think there are, are fundamentally hard and difficult problems when we have a new technology, and so I think we're always going to be having the AI ethics conversation, this is not something that we're going to solve and go away. But what I do think we have now is a lot more tools and techniques and best practices to help people start the journey of doing things responsibly. And so I think the reality is there are many things people could be doing right now that they're not. And so I, I feel like there's an urgency date to get some of these tools into people's hands so that we can do that. So I think we can quickly go a lot farther than we have right now.

Sam Charrington: [00:07:41] In my conversations with folks that are working on this and thinking about the role that responsible AI plays and the way they "do AI," do machine learning. A lot of people get stopped at the very beginning like: Who should own this? Where does it live? Is it a research kind of function or is it a product function, or is it more of a compliancy thing for a chief data officer or a chief security officer? [Is it] one of those executive functions and oversight, or compliance is the better word? What do you see folks doing and do you have any thoughts on successful patterns of where it should live?

Sarah Bird: [00:08:33] Yeah, I think the models that we've been using and are thinking a lot about... the transition to security, for example. And I think the reality is it's not one person's job or one function. Everybody now has to think about security, even your basic software developers have to know and think about it when they're designing. However, there are people who are experts in it and handle the really challenging problems. There is of course legal and compliance pieces in there as well. And so I think we're seeing the same thing where we really need every role to come

together and do this. And so one of the patterns we are seeing is part of the challenge with responsible AI and technology is that we've designed technology to abstract away things and enable you to just focus on your little problem, and this has led to a ton of innovation.

However, the whole idea of responsible AI is actually, you need to pick your head up, you need to have this larger context, you need to think about the application in the real world, you need to think about the implications. And so we have to break a little bit of our patterns of 'my problem is just this little box,' and so we're finding that user research and design, for example, is already trained and equipped to think about the people element in that. And so it's really great to bring them into more conversations as we're developing the technology. So that's one pattern that we're finding adds a lot of value.

Sam Charrington: [00:10:07] In my conversation with with Jordan Edwards, your colleague, many of his answers were all of the above. And it sounds like this one is an "all of the above" response as well.

Sarah Bird: [00:10:19] Yeah. I think doing machine learning in practice takes a lot of different roles, as Jordan was talking about, in operationalizing things, and then responsible AI just adds an extra layer of more roles on top of that.

Sam Charrington: [00:10:32] Yeah. I guess one of the challenges that kind of naturally evolves when everyone has to be thinking about something is that it's a lot harder, right? The developer is trained as a developer and now they have to start thinking about this security thing, and it's changing so quickly and the best practices are evolving all the time, and it's hard to stay on top of that. If we're to replicate that same kind of model in responsible AI, what sounds like the right thing to do? How do we support the people that are on the ground trying to do this?

Sarah Bird: [00:11:07] Yeah. And I think it's definitely a challenge because the end result can't be that every individual person has to know the state of the art in every area in responsible AI. And so one of the ways that we're trying to do this is, as much as possible, build it into our processes and our tooling. So that you can say, okay, well you should have a fairness metric for your model and you can talk to experts about what that fairness metric should be, but you should know the requirement that you should have a fairness metric, for example. And so we first are starting with that process layer and then in Azure Machine Learning, we've built tools that enable you to easily enact that process. And so the foundational piece is the MLOps story that Jordan was talking about where we actually enable you to have a process that's reproducible, that's repeatable.

So you can say, before this model goes into production, I know that it's passed these validation tests and I know that a human looked at it and said, it looks good. And if it's out in production and there's an error or there's some sort of issue that arises, you can go back, you can recreate that model, you can debug the error. And so that's the real

foundational piece for all of it.

And then on top of that, we're trying to give data scientists more tools to analyze the models themselves. And there's no magic button here. It's not just, Oh, we can run a test and we can tell you everything you want to know. But there's lots of great algorithms out there and research that help you better understand your model. Like SHAP or LIME are common interpretability ones. And so we've created a toolkit called Interpret ML, this is an open source toolkit you can use it anywhere. But it enables you to easily use a variety of these algorithms to explain your model behavior and explore it and see if there are any issues.

And so we've also built that into our machine learning process so that if I build a model, I can easily generate explanations for that model. And when I've deployed it in production, I can also deploy and explain her with it so individual predictions can be explained while it's running so I can understand if I think it's doing the right thing and if I want to trust it, for example.

Sam Charrington: [00:13:35] It strikes me that there's a bit of a catch 22 here, in the sense that the only way we could possibly do this is by putting tools in the hands of the folks that are working data scientists and machine learning engineers that are working on these problems. But the tools in their very nature kind of abstract them away from the problem and allow them, if not, encourage them to think less deeply about what's going on underneath. Right? How do we address that? Do you agree with that first of all?

Sarah Bird: [00:14:09] No, I completely agree with that and it's a challenge that we have in all of these cases where we want to give the tool to help them and to have more insight but it's easy for people to just use it as a shortcut.

And so in a lot of cases, we're being very thoughtful about the design of the tool and making sure that it is helping you surface insights. But it's not saying this is the answer because I think when you start doing that where you have something that flags and says this is a problem, then people really start relying on that. And maybe someday we will have the techniques where we have that level of confidence and we can do it. But right now we really don't, and so I think a lot of it is making sure that we designed the tools that encourages this mindset of exploration and deeper understanding of your models and what's going on. And not just, Oh, this is just another compliance tests I have to pass I just run this test and it says green. And I go.

Sam Charrington: [00:15:12] You alluded to this earlier in the conversation, but it seems appropriate here as well, and it's maybe a bit of a tangent, but so much of pulling all these pieces together is kind of a user experience and design. Any thoughts on that? Is that something that you've kind of dug into and studied a lot? Or are the other folks worry about that here?

Sarah Bird: [00:15:36] It's not in my background, but to me it's an essential part of the function of actually making these technologies usable. And particularly when you take something that as complex as an algorithm and you're trying to make that abstracted and usable for people, the design is a huge part of the story. And so what we're finding in responsible AI is that we need to think about this even more. And a lot of the guidelines are saying be more thoughtful and include sort of more careful design. For example, people are very tempted to say, well, this is the data I have so this is the model I can build and so I'm going to put it in my application that way.

And then if it has too much inaccuracy, then you spend a lot of resources to try and make the model more accurate where you could have just had a more elegant UI design, for example, where you actually get better feedback based on the UI design or the design can tolerate more errors and you don't need that higher model accuracy. So we're really encouraging people to co-design the application in the model and not just take it for granted that this is what the model does and that's the thing we're gonna focus on.

Sam Charrington: [00:16:53] With the Interpret ML tool, what's the user experience like?

Sarah Bird: [00:17:01] It depends on what you're trying to do, there's two types of interpretability that people think about. One is what we call Glass-Box models. And the idea there is I want my model to be inherently interpretable. So I'm gonna pick something like a linear model or decision trees where I can actually inspect the model and enable you to build a model of that, that you can actually understand. And so we support a bunch of different Glass-Box explainer or models. So then you can actually use it to train your own model. And the other part is Black-Box explainers where I have a model that is a black box and I can't actually inspect it, but I can use these different algorithms to explain the behavior of the model. And so in that case what we've done is made it easy for you to just call and explain and ask for global explanations and ask for local explanations and ask for feature importance. And then all of those are brought together in an interactive dashboard where you can actually explore the explanations and try to understand the model behavior. So a lot of the experience it's an SDK and so it's all easy calls to ask for explanations, but then we expect a lot of people to spend their time in that dashboard exploring and understanding.

Sam Charrington: [00:18:32] I did a really interesting interview with Cynthia Rudin who you may know she's a Duke professor and the interview was focused on her research that essentially says that we should not be using black box models in, I forgot the terminology that she used, but something like mission critical scenarios or something along those lines where we're talking about someone's life or liberty that kind of thing.

Does providing interpretability tools that work with black box models, like encourage their use in scenarios that they shouldn't really be used in? And are there ways that you advise folks when and when not they should be using those types of models?

Sarah Bird: [00:19:19] So we have people who do publish best practices for interpretability and it's a very active area of work for the company. And we work with the partnership on AI to try to make industry-wide recommendations for that. I don't think it's completely decided on this idea that models should be interpretable in these settings versus, well, we want other mechanisms to make sure that they're doing the right thing. Interpretability is one way that we could be sure that they're doing the right thing, but we also could have more robust testing regimes. Right? There's a lot of technologies where we don't understand every detail of the technology, but we've been able to build safety critical systems on top of it, for example.

And so yeah as a company we do try to provide guidance, but I don't think the industry has really decided the final word on this.

And so the mindset of the toolkit is enabling you to use these techniques if it's right for you. But that doesn't specifically say that you should go use a neural net in a particular setting.

Sam Charrington: [00:20:27] So in addition to the Interpret ML toolkit you also announced this week here from Ignite, a Fair Learn toolkit. What's that all about?

Sarah Bird: [00:20:39] So it's the same spirit as Interpret ML where we want to bring together a collection of fairness techniques that have been published in research and make it easy for people to use them all in one toolkit with the same spirit that you want to be able to analyze your model and understand how it's working so that you could make decisions around fairness. And so there's famously, many different fairness metrics published. I think there was a paper cataloging 21 different fairness metrics. And so we've built many of these common ones into the toolkit and then it makes it easy for you to compare how well your model works for different groups of people in your data set. So for example, I could say does this model have the same accuracy for men and women? Does this model have the same outcomes for men and women?

And so we have an interactive dashboard that allows you to explore these differences between groups and your model performance through a variety of these metrics that have been published in research. Then we've also built in several mitigation techniques so that if you want to do mitigation via post-processing and your model, then you can do that. For example, setting thresholds per group. And in a lot of cases it might be that you actually want to go and fix the underlying data or you wanting to make some different decisions. So the mitigation techniques aren't always what you would want to do, but they're available if you want to do that.

And so the name of the toolkit actually comes from one of these mitigation techniques from Microsoft research where the algorithm was originally called Fair Learn. And the idea is that you say, I wanna reduce the difference between two groups on a particular

dimension. So you pick the metric and you pick the groups and the algorithm actually retrains your model by re-wading data and iteratively retraining to try to reduce that disparity. So we've built that into the toolkit. So now you can actually look at a variety of your versions of your model and see if one of them has properties that works better for what you're looking for, to deploy.

Sam Charrington: [00:22:59] Again, I'm curious about the user experience in, in doing this. How much knob turning and tuning does the user need to do when applying that technique you were describing? Or is it more, I'm envisioning something like contextual bandage reinforcement learning where it's kind of tooling the knobs for you.

Sarah Bird: [00:23:18] Yeah, it is doing the knobs and the retraining, but what you have to pick is which metric you're trying to minimize. Do I want to reduce the disparity between the outcomes or do I want to reduce the disparity and accuracy or some other there's many different metrics you could pick, but you have to know the metric that's right for your problem. And then you also need to select the groups that you want to do. So it can work in a single dimension like as we were saying making men and women more more equal, but then it would be a totally separate thing to do it for age, for example. So you have to pick both the sensitive attribute that you are trying to reduce disparity and you have to pick the metric for disparity.

Sam Charrington: [00:24:10] Were you saying that you're able to do multiple metrics in parallel or you're doing them serially?

Sarah Bird: [00:24:17] Right now the techniques work for one, for just one metric. So it will produce a series of models, and if you look at the graph, you can actually plot disparity by accuracy and you'll have models that are on that Pareto optimal curve to look at. But then if you said, okay, well now I want to look at that same chart for age, the models might be all over the place in the space of disparity and accuracy. So it's not a perfect technique, but there are some settings where it's quite useful.

Sam Charrington: [00:24:48] So going back to this idea of abstraction and tools versus deeply understanding the problem domain and how to think about it in the context of your problem domain. I guess the challenge domain or your problem domain, I don't know what the right terms are. But you mentioned that paper with all of the different disparity metrics and the like. Is that the best way for folks to get up to speed on this or are there other resources that you've come across that are useful?

Sarah Bird: [00:25:23] Yeah, I think for fairness in particular it's better to start with your application domain and understand, for example, if you're working in an employment setting, how do we think about fairness and what are the cases and so in that case we actually recommend that you talk to domain experts, even your legal department to understand what fairness means in that setting. And then you can go to the academic literature and start saying, okay, well, which metrics line up with that

higher level concept of fairness for my setting.

But if you start with the metrics I think it can be very overwhelming and there's just many different metrics and a lot of them are quite different and in other ways they're very similar with each other. And so I find it much easier to start with the domain expertise and know what you're trying to achieve in fairness and then start finding the metrics that line up with that.

Sam Charrington: [00:26:22] You're also starting to do some work in the differential privacy domain. Tell me a little bit about that.

Sarah Bird: [00:26:27] Yeah, we announced a couple of weeks ago that we are building an open source privacy platform with Harvard and differential privacy is a really fascinating technology. It was first published in Microsoft Research in 2006 and it was a very interesting idea, but it has taken a while for it, as an idea, to mature and develop and actually be able to be used in practice. However, now we're seeing several different companies who are using it in production. But in every case the deployment was a very bespoke deployment with experts involved. And so we're trying to make a platform that makes it much easier for people to use these techniques without having to understand them as much.

And so the idea is the open source platform can go on top of a data store, enable you to do queries in a differentially private way, which means that actually it adds noise to the results so that you can't reconstruct the underlying data and also then potentially use the same techniques to build simple machine learning models. And so we think this is particularly important for some of our really societaly valuable datasets. For example, there are data sets where people would like to do medical research, but because we're worried about the privacy of individuals, there's limits to what they can actually do. And if we use differential private interface on that, we have a lot more privacy guarantees and so we can unlock a new type of innovation and research in understanding our data.

So I think we're really excited and think this could be the future of privacy in certain applications, but the tooling just isn't there, and so we're working on trying to make it easier for people to do that. We're building it in the open source because it's important that people can actually ... It's very easy to get the implementation of these algorithms wrong and so we want the community and the privacy experts to be able to inspect and test the implementations and have the confidence that it's there. And also we think this is such an important problem for the community. We would like anybody who wants to, to be joining in and working on this. This is not something that we can solve on our own.

Sam Charrington: [00:28:58] Yeah, differential privacy in general and differentially private machine learning are fascinating topics and ones that we've covered fairly extensively in the podcast. We did a series on differential privacy a couple of years ago maybe and it's continuing to be an interesting topic. At the Census Bureau I think is using differential privacy for the first time next year and it's both providing the

anticipated benefits but also raising some interesting concerns about an increased opacity on the part of researchers to the data that they wanna get access to. Are you familiar with that challenge?

Sarah Bird: [00:29:41] Yeah, absolutely. So the reality is people always want the most accurate data, right? It doesn't sound great to say, well, we're adding noise and the data is less accurate. But, in a lot of cases it is accurate enough for the tasks that you want to accomplish. And I think we have to recognize that, privacy is one of the sort of, fundamental values that we want to uphold, and so in some cases it's worth the cost. For the census in particular, to motivate the decision to start using this for the 2020 census they did a study where they took the reports from the 1940 census and they were able to recreate something like 40% of Americans' data with the result of just the outputs from the census.

Sam Charrington: [00:30:33] Meaning personally identify 40% of Americans?

Sarah Bird: [00:30:37] Yeah, he talks about this in his ICML keynote from last year. So if you want to learn more you can watch the keynote. But yeah, basically they took all the reports and they used some of these privacy attacks and they could basically recreate a bunch of the underlying data. And this is a real risk, and so we have to recognize that yes, the census results are incredibly important and they help us make many different decisions, but also protecting people's data is important. And so some of it is education and changing our thinking and some of it is making sure that we use the techniques in the right way in that domain where you're not losing what you were trying to achieve in the first place, but you are adding these privacy benefits.

Sam Charrington: [00:31:21] There are a couple of different ways that people have been applying differential privacy one is a, a more centralized way where you're applying it to a data store. It sounds a little bit like that's where your focus is. Others like Apple's a noted use case where they're applying differential privacy in a distributed manner at the handset to keep user data on the iPhone, but still provide information centrally for analysis. Am I correct that your focus is on the centralized use case? Or does the toolkit also support the distributed use case?

Sarah Bird: [00:32:02] We are focusing on the global model. The local model works really well, and particularly in some of these user telemetry settings, but it limits what you can do. You need much larger volume to actually get the accuracy for a lot of the queries that you need, and there aren't as many queries that you can do. And so the global model, on the other hand, there's a lot more that you can do and still have reasonable privacy guarantees. And so as I was saying, we were motivated by these cases where we have the data sets. Like somebody is trusted to have the data sets but we can't really use them. And so that looks like a global setting. And so to start, we're focused on, on the global piece, but there are many cases where the local is promising and there are cases where we are doing that in our products. And so it's certainly a direction that things could go.

Sam Charrington: [00:32:58] And differential privacy from a data perspective doesn't necessarily get you to differentially private machine learning. Are you doing anything in particular on the differentially private ML side of things?

Sarah Bird: [00:33:11] The plan is to do that but the project is pretty new so we haven't built it yet.

Sam Charrington: [00:33:19] And before we wrap up, you're involved in a bunch of industry and research initiatives in the space that you've mentioned, MLSys, a bunch of other things. Can you talk a little bit about some of the broader things that you're doing?

Sarah Bird: [00:33:38] Yeah, so I helped found the, now I think named MLSys systems and machine learning research conference. And that was specifically because I've been working at this intersection for a while and there were some dark days where it was very hard to publish work because the machine learning community was like, this is a systems result. And the systems community was like, this doesn't seem like a systems result and so we started the conference about two years ago and apparently many other people were feeling the same pain because even from the first conference, we got excellent work.

People's top work, which is always a challenge with research conferences because people don't want to submit their best work to an unnamed conference. Right? But there was such a gap for the community. So it's been really exciting to see that community form more and now have a home where they can put their work and connect. I've also been running the machine learning systems workshops at NeurIPS for several years now. And that's been a really fun place because it really has helped us form the community, particularly before we started the conference. But it's also a place where you can explore new ideas. This last year we're starting to see a lot more innovation at the intersection of programming languages and machine learning. And so in the workshop format we can have several of those talks highlighted, and have a dialogue, and show some of the emerging trends so that's been a really fun thing to be involved in.

Sam Charrington: [00:35:13] Awesome. Yeah, was it last year that there was both the SysML workshop and the ML for systems workshop and it got really confusing?

Sarah Bird: [00:35:24] Yeah. This year too. We have both. And I think that's a sign that the field is growing that it used to be that it felt like we didn't even have enough people for one room at the Intersection of Machine Learning and Systems. And I think this last year there was maybe four or 500 people in our workshop alone. And so that's great. Now, there's definitely room to have workshops on more focused topics. Right? And so I think machine learning for systems is an area that people are really excited about

now that we have more depth in understanding the intersection. For me, it's very funny because that is really kind of the flavor of my thesis which was a while ago. And so it's a fun to see it now starting to become an area that people are excited about.

Sam Charrington: [00:36:16] The other conference that we didn't talk about, ML for Systems is all about using machine learning within computational systems, networking systems as a way to optimize them. So for example, ML to do database query optimization. Also a super interesting topic.

Sarah Bird: [00:36:36] Yeah, I know it absolutely is. And I really believe in that, and I think for several years people were just trying to replace kind of all of the systems intelligent with one machine learning algorithm and it was not working very well. And I think what we're seeing now is recognizing that a lot of the algorithms that we used to control systems were designed for that way and they work, actually, pretty well. But on the other hand, there's something that's dynamic about the world or the workload. And so you do want this prediction capability built in. And so a lot of the work now has a more intelligent way of plugging the algorithms into the system. And so now we're starting to see promising results at this intersection. So my thesis work was a resource allocation that built models in real time in the operating system and allocated resources.

And it was exactly this piece where there was a modeling and a prediction piece, but, the final resource allocation algorithm was not purely machine learning.

Sam Charrington: [00:37:43] Awesome. Wonderful conversation, looking forward to catching up with you at NeurIPS, hopefully. thanks so much for taking the time to chat with us.

Sarah Bird: [00:37:52] Yes, thanks for having me. And I look forward to seeing you at NeurIPS.

Sam Charrington: [00:37:56] Thank you.

(<https://twimlai.com/article/enterprise-reachine-learning-with-erez-barak-transcript/>)

Related Articles