SDS PODCAST EPISODE 1011: THE MATH STILL MATTERS: DEEP SKILLS IN THE AGE OF AI, WITH DR. CATHERINE WILLIAMS

Jon Krohn:	00:00:00	Today's guest was solving black hole equations with pen and paper before she ever wrote a line of code. And in today's episode, she makes a compelling case that going deep on the math underlying machine learning matters more than ever, even now that AI can do the math for you. Welcome to episode number 1011 of the Super Data Science Podcast. I'm your host, Jon Krohn. My guest today is Dr. Catherine Williams, chief data officer at the nonprofit Candid. Catherine earned a PhD in math researching general relativity and black holes, did postdocs at Stanford and Columbia, and then became one of the very first data scientists anywhere. Joining AppNexus back in 2012 around the same time data scientist became a job title at all. Since then, across more than a decade of senior data leadership at AppNexus, Xander, Qualtrics, and now Candid, she's watched our field get born and then reinvent itself again and again.
	00:00:53	In this episode, she traces that evolution from Bayesian models to Burt to today's LLMs and shares sharp, hard won guidance on which skills will still matter as machines take over more of the technical work. Enjoy. This episode of Superdata Science is made possible by Anthropic, Cisco, Excel Data and Gurobi. Catherine, welcome to the SuperData Science Podcast. It's great to have you on the show. How are you doing today?
Catherine W.:	00:01:17	Thank you. Do well. Glad to be here.
Jon Krohn:	00:01:19	And where are you calling in from roughly in the world?
Catherine W.:	00:01:22	I am precisely in Seattle, Washington this morning.
Jon Krohn:	00:01:26	Okay. Nice. West Coast. Where are you? Well, when you and I chatted last week to plan this episode, I was in New York. I was in the Lightning AI office, but now I'm actually, I'm near Toronto, Ontario visiting my family where I have

	00:01:42	Exactly the same studio set up as I have in New York. So yeah, so things should look exactly the same and sound exactly the same to listeners regardless. Yes, thanks for asking. Almost nobody ever asks that. Anyway, let's get into the technical stuff, Katherine. So you're currently the chief data officer at a nonprofit called Candid, and we'll get into that. But you also have a storied career, several major data leadership roles in things like high velocity ad marketplaces, but you began your career with a PhD in math researching general relativity and black holes. And then you did postdoc research at Stanford and Columbia. Regular listeners will know that I usually jump to what people are doing right now and have people focus on that and then we work our way backwards. But what you were doing for your research was so fascinating. I just wanted to hear more about it and I'm sure our audience will too.
Catherine W.:	00:02:40	Sure. Well, it is my first love, my first career, my first set of ambitions. I loved math from childhood and wound up majoring it in college. Took a brief deer tour to work at Microsoft for a couple of years while I got my act together to apply to grad school. But then yeah, went to math grad school, explored a bunch of different subfields of math, but eventually found my way to geometry. I think the visual intuition always meant a lot to me. I was able to get further in that space. And then the advisor that I wanted to work with at University of Washington had recently made the switch from differential geometry into general relativity, which really is just a sub-flavor of differential geometry.
Jon Krohn:	00:03:21	I did not know that.
Catherine W.:	00:03:22	Einstein's true famous equation, not E equals MC squared, but the real Einstein equation, it basically says math and curvature stuff equals physics matter energy stuff. There's geometry on one side and physics on the other side. And so mathematicians study the curvature and the geometry side of that. And I did as well,

specifically black hole space times. And it was very gratifying both mathematically and to sort of know that there was this connection, at least theoretically, to the real world.

Jon Krohn: 00:03:52 Wow. That is cool. And I had no idea about that connection between geometry. And I guess I basically would've assumed that if you're studying relativity and black holes, you'd be an astronomer.

Catherine W.: 00:04:08 I mean, people do. There are people who take detailed measurements and there are people who do really sophisticated numerical simulations and then compare the measurements to the simulations in order to derive inferences about what's going on. But then you also need to know sort of theoretically what's even possible to simulate. And black hole space times are black holes are caused by a concentrated amount of matter and energy that causes very intense curvature around them. And so modeling that becomes very tricky and it's very important to understand the geometry. Yeah, the geometric aspects of that.

Jon Krohn: 00:04:44 Cool. You said simulations. Was some of your work computational in getting large scale simulations of behavior happening? Because when you and I were talking last week, yeah, you were saying you were a paper and pencil mathematician.

Catherine W.: 00:04:58 Pen and paper. I learned somewhere along the way that I don't like to erase. I'd rather scratch out or cross out than erase. But yeah, no, I was pen and paper only. I went to conferences and talked to people who were doing the numerical simulations, which at that time were pushing up against sort of some of the hardware limits of the time. It basically boils down to systems of partial differential equations. And so that is what I did my work on is figuring out properties of solutions to the very specific

		partial differential equations and what one can say about space times as a result.
Jon Krohn:	00:05:27	So could you kind of end up working through derivatives that are like multiple pages long?
Catherine W.:	00:05:33	I had lots and lots of intense calculations, but I will say, so maybe this goes deeper than you're interested in.
Jon Krohn:	00:05:39	No.
Catherine W.:	00:05:40	But a lot of my work, I studied spherically symmetric space times, which means that you can suppress two degrees of freedom because you know there's the symmetry involved and really just look at two dimensions. And so a lot of the work I did was actually just two dimensional PDEs.
Jon Krohn:	00:05:58	Oh really?
Catherine W.:	00:05:59	And I was looking at systems of them and properties of the solutions. So sort of some geometric analysis type things specifically looking at an alternate characterization of black hole boundaries. Everybody's familiar with the famous event horizon and that does have a geometric meaning, but it's hard to detect, shall we say, locally geometrically. And so I was looking for more local characteristics of solutions, space time solutions where you could detect where there'd be a black hole present. So yeah.
Jon Krohn:	00:06:29	Let's go into this a little bit more. So I'm going to try to define black hole and event horizon and then you're going to tell me where I'm wrong. Okay. But basically it's a point in space that has become so heavy I believe from a star collapsing that it starts to suck in everything around it. And the event horizon is where if you get to that point far away from the black hole, you are kind of like inevitably sucked into the black hole from that point.

Catherine W.:	00:06:58	Yeah. But geometrically that's very hard to characterize. And so the person who did it was Roger Penrose back in the '60s.
Jon Krohn:	00:07:06	Roger Penrose strikes again.
Catherine W.:	00:07:08	Yeah. Yeah. He's all over the place. Interesting character. I met him years ago in Cambridge.
Jon Krohn:	00:07:13	No way.
Catherine W.:	00:07:14	Fascinating guy. In order to say like, yes, the common sense or the common way of saying it is like, yeah, once you cross the threshold of an event horizon, you can never escape. Okay. But how do you make that mathematically precise? In order to make that mathematically precise, you have to have a precise way of saying what the difference is between being able to get back out and in is. So Penrose has this definition where you sort of extend your space time for all future infinity and then attach a boundary. And a black hole is where you can't get to that boundary if you're inside the black hole. So it's weird. You have to look at the whole future of the space time. You have to know everything that ever happens in order to know where the black hole was. There's no way that in crossing it, nothing happens.
	00:08:05	You can cross an invent horizon and have everything be totally normal. You wouldn't know that you can't get back out of it. Nothing happens for a long time. So that's where the geometric characterization breaks down. So the quasi local version that I was looking at was one where like, no, something happens physically when you cross this boundary. Light starts going out, it starts going in, which I know sounds super weird, but you would know immediately and your bones would start getting crushed immediately with this other sort of definition of black hole. And my specific thesis work and a couple of my other papers after that were looking at to what extent

		does this quasi-local notion of the surface of a black hole, the boundary of a black hole coincide with that event horizon? And I was able to show under some reasonable assumptions, they actually coincide at infinity.
	00:08:52	So they're asymptotic to each other eventually. So it doesn't... Yeah, they're well behaved, shall we say.
Jon Krohn:	00:08:58	And so you were talking earlier about symmetry. And so is it the case that black holes, event horizons are often symmetrical and so this makes modeling easier?
Catherine W.:	00:09:09	I mean, I think the real world shows high degrees of symmetry, but it's imperfectly symmetric. So absolutely this does not characterize the full four dimensional or higher dimensional space time, but it does have some geometric properties that we think probably do hold outside of that. It's a very common mathematical technique to study a highly symmetric constrained problem and then try to say, "Hey, if we perturb this away from symmetry, does the property still hold?" And then you say, "Well, how much does it hold?" And like, oh, lo and behold, it holds even outside of the symmetric constraints. And so this was studying the simple version. But no, we don't believe that reality is in fact spherically symmetric.
Jon Krohn:	00:09:46	Right, right. Gotcha, gotcha, gotcha. So very intense academic question for you coming up next. Have you seen the film Interstellar starring Matthew McConaughey?
Catherine W.:	00:09:57	Yes. Yes, I have.
Jon Krohn:	00:10:01	What do you think of it?
Catherine W.:	00:10:03	Well, the physics was really good. The whole time dilation - Oh, okay. This is what I wanted to hear. The time dilation aspects of it were dead on. I mean, they had Kip Thorne advising, I believe, from Caltech. So he's one of the founding modern fathers of general relativity. And

		yeah, I think he was able to steer them in the right direction about the actual practical implications of getting close to the black holes and the time dilation that happens such that you then can't travel backwards. And my understanding is that in advising on that movie and helping with some of the visual simulations of black holes, he actually did some good research. Some papers came out of his advising for that movie about -
Jon Krohn:	00:10:47	Whoa.
Catherine W.:	00:10:48	There's the iconic image of the round thing, but then with a loop over the top of it, which apparently was something that came out of some of the modeling that they were doing just for the CGI for the movie, but they said physically, actually, this is sort of what it would look like. So I don't know. It's very cool.
Jon Krohn:	00:11:04	That is really cool. As
Catherine W.:	00:11:05	For the whole rest of the plot, whatever.
Jon Krohn:	00:11:09	I loved it. I've only seen it one time and I saw it pretty recently. I'm looking up the film year. So it's over 10 years old. It came out in 2014. Actually, wow, the cast is incredible now that I'm looking at it here. I remember Matthew McConaughey because he's the main character, but it also has Matt Damon, Jon Lithgow, Jessica Chastain, An Hathaway, Timothy Chalamet, Mackenzie Foy, Casey Affleck, Topher Grace. It's actually an insane cast. It's insane.
Catherine W.:	00:11:38	Yeah.
Jon Krohn:	00:11:39	And I thought it was really good. I really enjoyed watching it. And I'm not going to give away any spoilers for our listeners, but it's top recommendation for me film-wise. I think you'll really enjoy it. Touching some science ideas to think about and also just climate change and where

		we're going in the world. 10 years on. I think the core messages probably come across more clearly than ever.
Catherine W.:	00:12:14	That's funny. I don't remember the rest of the plot very clearly. I remember some of the physics pieces. That's
Jon Krohn:	00:12:18	All I really wanted to ask you about.
Catherine W.:	00:12:20	Don't worry. I don't want to spoiler
Jon Krohn:	00:12:21	Anyway. Won't spoiler. Yeah, we won't spoil anything. We don't remember enough about it to spoil the movie for you. Cool. Well, so yeah, back to the math. Doing that research, doing a PhD, doing a postdoc, did anything from that time, do you think that there were intellectual habits that proved useful in all of the senior leadership that you've had in industry since? Do you think that there were aspects of that that were useful? I guess obviously there's technical things and we'll kind of get into how backgrounds for data science have changed over the years. But yeah, do you think that besides the technical stuff, do you think that there are aspects, intellectual habits from all the rigor that you had in your education and the postdocs that has proved useful in industry since?
Catherine W.:	00:13:18	Yeah. I have I guess two examples and then maybe even a counter example. So one of course is yes, rigor. I think when one does a PhD in math, you stop being scared of hard technical subjects or going deep. And so I am mathematically fluent. You give me a paper in machine learning and I am confident that with enough time and diligence, I can get through it and understand it down to the bottom. So there's just some facility with that kind of, I don't know, it's not just numerical literacy, but that kind of abstraction and that kind of thinking that I think comes along with it. Another one that I think came along that I didn't realize till later, until I started practicing it more intentionally in my career has to do with building

mental models. I implicitly did this along the way in my PhD thesis of over time, as I was studying reading papers and whatever, I was sort of building up this picture in my mind of the possibilities and where the degrees of freedom were and what the interesting questions were and so forth.

00:14:24 And I have discovered and now intentionally practice in my professional career doing the same thing with each sort of business setting that I'm in, of building out a mental model of what is the data flow? Where is it coming from? What are all the different parts of the system? Do I understand this correctly? Can I repeat it back to the person who told me and they nod or do they say, no, you have that a little bit wrong and do I need to dive a little bit deeper to understand it? And so I think intentionally building that kind of... And I can do it now, not by reading papers, but just by talking to people like, explain this to me. Do I understand this? What about this part? Okay. And then that winds up being really, really helpful for reasoning about when you're interacting with people or making decisions or figuring out what to build or whatever that is.

00:15:06 So those are maybe two big things that came out of that academic career. And then I think the one that I'll cite as a counter example is actually a pattern I kind of had to unlearn.

00:15:17 I used to think of myself as a bottom up thinker, meaning that mathematically I had to get all the way down into the nitty gritty and feel like I understood every last piece to be confident that I understood the big picture. You know what I mean? I wasn't happy with calculus until I learned real analysis and understood exactly what all the limits and the theorems were saying about what was true and what wasn't true, which is great for a math PhD. It's great if you want to be rigorous. It's not great in the business world. You have to be able to sometimes reason, this is where that mental model comes in. You have to be able to reason about things without going all the way down to the bottom and understanding every single last detail. In fact, that will trip you up and you will waste a whole bunch of time and not be effective.

00:16:01 Becoming a top down thinker or learning how to build that muscle and be confident that I can still make good decisions without understanding all the gory mathematical details is definitely something that I've had to learn and practice.

Jon Krohn: 00:16:14 Really cool. Those were really interesting examples. And in them, at the end, they were talking about going deep into the math, going deep into calculus in particular. And it used to be common in data science, machine learning, AI for people to be studying advanced calculus, linear algebra, statistics. And those were kind of seen as essential prerequisites for being able to do professional work in our space. But it seems like in recent years, there's been a shift more towards engineering. Software engineering generally, but also specific subdisciplines within that data engineering, ML engineering and AI engineering of course. Do you think that having a really deep understanding of the underlying mathematics is still really useful today? I'm guessing the answer is yes, because you talked about, for example, being able to understand ML papers in detail about being able to work through them. And obviously if somebody doesn't have a rigorous math background, they can't dig through all that.

00:17:17 But we're also in this interesting time now where large language models are getting particularly good at. If you look at the meter charts of capability, those are based around capability on programming tasks, math tasks. It's a very interesting time that we're in right now where I wonder how those technical skills, any of the ones I just mentioned. I said engineering seems to have supplanted kind of mathematical background in importance in our field, but both of those things, programming and math are the two things most vulnerable to disruption by large language models. So anyway, very long question. My apologies for that.

Catherine W.: 00:18:03 Yeah. I think about this a fair amount in different flavors and I don't have a great answer for where it's all going. I mean, I think it's clear that there's a hierarchy of different levels of complexity and these things build on each other. So there's arithmetic that gives rise to algebra, then you have linear algebra and then you have systems of things that lead to machine. I mean, these are bad examples, but there's just layers and layers of conceptual hierarchy involved conceptually. And then also on the physical side, right you have electron, you have circuits and you have chips and then you have hardware and then you have motherboard. I mean, you can walk the stack of complexity. And I think studying any one particular piece of that can be very valuable because each piece in that chain plays a role and has gotchas and has value to being deeply understood and researched and so forth.

00:19:02 And as machine capabilities start being able to do most of the work in those areas, it's less necessary to go deep. And so you then move to the next higher level of the hierarchy and focus your attention there or on assembling the pieces. But that doesn't mean that understanding lower levels isn't still really valuable. And in fact, it's really important. In my mind, it's parallel to the argument of should you teach kids arithmetic when they all have calculators? Well, yeah, because you need to have that sort of fluency with conceptually, even if you're not going to do the arithmetic yourself, you need to have the conceptual fluency in it. Sometimes I think about it as like we're training our own neural networks so that we have the right subsystems to then create the abstractions, to create the abstractions on top of that, that then lead to the right understanding of the world.

	00:19:49	So anyway, the transition from sort of math to engineering to me reflects one movement sort of up the hierarchy. And now that we have LLMs who can do the engineering part, what's the next level on top of that? Well, presumably some kind of a systems view, but you're still going to need to be able to dive down into the different layers and understand how they fit together, I think. So I think if anything, the best data professionals going forward are going to be ones who can really move up and down that chain and build their mental models and continue to update them as they learn over time rather than specializing in one particular slice.
Jon Krohn:	00:20:27	You articulated all of that really well and it makes it so easy for me to agree with you on all the points you made. And I look for the confirmation bias in these kinds of questions. Let me find people with strong technical backgrounds and ask them whether it's going to continue to be useful to have a technical background of the future. I agree with everything you're saying to be able to dive deeply and to even just have an intuition for what's possible. It's one thing to go into an LLM and ask, "This is the problem that I'm facing." And you can provide lots of context and have it generate ideas as to possible solutions to a problem. But I wonder if by being able to go deeply technically into questions as opposed to... So when you make a mental model, the way that that mental model is stored is very different from the vector embedding next token prediction approach that an LLM has.
	00:21:32	And you giving that example of kids with calculators and yeah, I guess you could theoretically in grade one or whenever kids start doing arithmetic, you could just give them a calculator and show them how to use calculator and never have them kind of learn what numbers even mean and just have them kind of recognize the symbols and enter those into a calculator. It would be possible, but it limits how creative you can be in terms of solutions. And given how differently humans think

		differently from machines, hopefully there will still continue to be a lot of value, a lot of creativity in the way that we create our mental models and the way we can have intuitions around concepts for years to come.
Catherine W.:	00:22:22	I think that's exactly right. And I wonder about that as a frontier for AI too. I know, is it Lacoon who's interested in world models and just like pulling more and more different pieces into AI, helping AI build its own conceptual frameworks. And maybe it starts to approximate what we have, but I do think there's many more layers to go before we're done.
Jon Krohn:	00:22:47	For sure. For sure. And that's these places, math and programming have been the easiest places for machines to have superhuman capabilities because those are the easiest areas to create training data. Because if the solution works out, it's a pretty good indicator that the algorithm solved it correctly. And so you can create lots of training data there that is in the real world problems don't have that same kind of, you can't simulate data as reliably because you can't be confident that an answer makes sense just because you generated it. So yes, yes, yes. The world models thing, as you say, Yalnika, Faifei Li and others trying to, they're raising billions of dollars in seed rounds to be able to create these huge data sets and it'll definitely make a difference. But I still think the way that information flows through machines has differences to the way that information flows through our own brains.
	00:23:56	And yeah, hopefully we'll continue to provide some value to machines for a while.
Catherine W.:	00:24:01	Well, it is interesting. So raising all this money for world models and so forth seems to point to the future being just like the machines getting smarter and smarter and smarter. Where actually I think the really interesting frontier is machines into humans.That interface between

their thinking and our thinking. The way I like to use LLMs is use them to help me improve my own mental models. And then I prompt back. I don't know. It's not a very well articulated though, but I do think that there's a lot there that I haven't heard well explored there other than people sharing their prompts and comparing notes.

Jon Krohn: 00:24:40 Oh no, I know what you mean. And I think that this kind of thing, I don't think we'll be able to get billion dollar seed round funding for the kind of innovation that you're describing. I don't know. There isn't the same kind of obvious enterprise use case. There theoretically is. You could have a learning and development department be investing in these kinds of educational tools that allow us to better understand calculus and algebra and engineering and tie all those things together to be better ML engineers, for example. But it doesn't seem to me like the TAM, the total addressable market there is nearly as big as there is for developing role models and kind of this idea of everything being fully automated. But what you're describing to me as somebody who likes learning stuff does sound way more interesting. And yeah, I mean, there's lots that we can do today, but having tools that make it...

00:25:43 We're obviously at a time where it has never been easier to learn. And some people are taking advantage of that, but it seems like it's probably a minority of people who have access to these tools and probably some multiple. So of the people who are using large language models to understand things more deeply and say, be able to do better on a college paper or on some work project, there's probably a multiple of those people who are using them just to spit out an answer and circumnavigate learning.

Catherine W.: 00:26:21 Exactly. You don't have to build your own mental model because you can just copy paste something instead and get by. Yeah.

Jon Krohn:	00:26:29	Yeah.
Catherine W.:	00:26:30	Which is an approach. I mean, but yeah.
Jon Krohn:	00:26:34	Yeah. It's an approach. So there's an economist article that came out this week at the time of you and me recording that is about how at UC Berkeley, which is a tough school to get into, it's a prestigious university, one in eight people, first year undergrads have math. And it was in math, I think it was in a program where math matters. Math was a prerequisite for getting into this program. One in eight people that get admitted to Berkeley for that technical degree program have math abilities below high school level. Wait,
Catherine W.:	00:27:14	What is happening?
Jon Krohn:	00:27:15	I
Catherine W.:	00:27:15	Don't get it.
Jon Krohn:	00:27:16	Exactly. And so it seems like it's hard to know what the causal factors are, but it seems like it's probably a blend of LLMs. And also I think the pandemic had a big impact on learning.
Catherine W.:	00:27:33	Backtracks, but wow.
Jon Krohn:	00:27:34	Yeah. Anyway. All right. Well, fascinating conversation there about your background, Catherine. Let's now dig into the journey since. So from pure mathematics research, you went to ad tech. So companies like AppNexus and Xandr, which were doing revolutionary things in AdTech at the time. And then you went to Enterprise SaaS at Qualtrics, another really well-known name. And now you're in the nonprofit sector at Candid. Each of those transitions represents a significant shift in industry, culture and purpose. And something else that happened over that time is data science changed a lot. So you started as a professional kind of data scientist in

2012 at AppNexus. And so you started as a quantitative analyst at AppNexus it looks like from your LinkedIn profile and then manager of quantitative analytics, head of data science after that, then chief data scientist. And by the time you left AppNexus after almost seven years, you were chief data and marketplace officer.

00:28:48 And so you have about as much experience as a professional data scientist as anybody could because I believe it was around 2012 when you had your first kind of data science role that that term even came into being. And I'd love to hear kind of simultaneously, if this isn't too crazy a question, how your career journey personally evolved while our industry data science was kind of born and evolved to where it is today.

Catherine W.: 00:29:20 Yeah. They are very, very tightly intertwined. And I think I wouldn't have been able to have the career trajectory that I had if either one had been offset by even a couple of years. So remember I said I was a pen and paper mathematician. So I taught myself how to code. I knew a little bit of data. I learned some SQL online. I wasn't a computer scientist. I was hired at AppNexus because they wanted people who could do probability math because at that time, all the data that was backing up the real time bidding systems was stored in aggregate form. And so they were using complicated sort of Bayesian Predictions in order to build out the algorithms to be used for the bidding. And so that mathematical fluency, the fact that I had a PhD in math, even though I hadn't done anything around probability, I could pick it up because like I said, I'm mathematically fluent.

00:30:14 So they hired me despite a complete lack of data skills, but because I was mathematically fluent. But within a couple of years of being there, the technology was changing really rapidly. So Hadoop came out and all of a sudden our systems, we were able to start getting access to log level data. So now we're talking big data, not aggregated data, raw data. And that requires a different set of tools and it opens up a totally different set of possibilities. And so I was at that time managing then a team that had been quantitative analytics, but increasingly it became clear that there were more techniques. And so the team and I learned together. We sort of bootstrapped ourselves up. We read papers and went to conferences and learned what one can do with data and started building out different kinds of systems that weren't Beijing. They were more sort of what we consider as traditional ML, logistic regression models, et cetera, decision trees, not decision trees, random forests, that kind of thing based on the different flavor of data.

00:31:17 And that then continued. My own career then kind of diverged a little bit from data science and I took on additional responsibilities within AppNexus that were still sort of mathematically flavored. So I was looking at marketplace dynamics and got deep into auction theory, which is actually super interesting. So not as much a traditional data science path, but that's what then took me to be chief data marketplace officer. I was marketplace czar for a while at AppNexus.

Jon Krohn: 00:31:45 So it was around the same time. I mean, I also had my first data science job in 2014. So a couple years after you and relatively early on and you talked about there being a traditional path and I don't know what that would have meant at the time. We were all kind of figuring it out together. I remember in that first data science job, Dr. Amit Batacharia, he was like the existing data scientist at Omnicom and he was like showing me Jupyter Notebooks and I was getting into Python with him because I had previously been programming in R and Matlab and yeah, it's funny.

Catherine W.: 00:32:23 Pandas took the world by Storm. Wes

Jon Krohn: 00:32:25 McKinney

Catherine W.:	00:32:26	Worked out of the AppNexus office. We were like early beta testers of Pandas.
Jon Krohn:	00:32:31	Oh really?
Catherine W.:	00:32:33	Yeah.
Jon Krohn:	00:32:33	Oh, that's super cool. And if people don't already know who Wes McKinney is. So Wes McKinney created Pandas, which I presume most listeners know is the standard data frame library for Python. And if you want to hear him on the show, he was in episode number 523, but him being on this podcast for one hour is not nearly as cool as having him work in your office. Early
Catherine W.:	00:32:55	Days.
Jon Krohn:	00:32:55	Yeah. Amazing. Yeah.
Catherine W.:	00:32:57	Yeah. It's funny. You're right. It was the wild west at the time. So it is kind of funny that I referred to it as the traditional path. I think it's only in hindsight. Kaggle came along around that time. There started being sort of in hindsight, the things we were doing became kind of canonical or more canonical or more mainstream. I don't know. So maybe that's why I'm thinking of it that way. But at the time it was anything but well trodden or well understood. So yeah,
Jon Krohn:	00:33:23	There was a relatively short period, but in the beginning, kind of that 2012 to 2015, maybe 16 area where there was an expectation for a lot of data science roles that industry was offering at that time that you have a PhD.
Catherine W.:	00:33:41	I heard a lot of people with PhDs. Yeah.
Jon Krohn:	00:33:43	Yeah. And it's kind of unimaginable at this time. The industry has become so big that you couldn't possibly be only like where would you get the people if you were only willing to take people with PhD background. So you did

		have that and they were looking, it seemed like one of the primary things people were looking for in early data science roles was a PhD in a quantitative discipline, which both of those boxes tick for you. And it was kind of like it's that idea of, I can't remember who said it. I've actually had them on the show whoever coined this term, but I can't remember who it is right now. You might even remember because I think they were also kind of in that New York Wes McKinney group of people. They coined the term that a data scientist is a mathematician who's bad at programming or a programmer who's bad at math.
	00:34:36	It's like kind of both together. It's like this Venn diagram. It's like, yeah, data scientist is someone who's neither good at math or programming was the idea or it's kind of the joke. Well, it certainly
Catherine W.:	00:34:46	Is neither discipline. Yeah. It's definitely a third thing. Yeah.
Jon Krohn:	00:34:51	And the idea that we're kind of borrowing a bit from both and trying to make it work in practice.
Catherine W.:	00:34:57	I think the hiring PhDs, I mean, I hired people with PhDs, not because their research was relevant, but because I wanted to make sure that I was hiring, or it was often a good credential to signal that somebody could go really deep on hard problems and wasn't scared of not knowing a well-trodden path because it was the Wild West at the time, because I needed you to go figure out what need figuring out, even if it hadn't been done before. I think now things are better understood and that kind of frontier mindedness is not maybe as necessary. I don't know.
Jon Krohn:	00:35:36	Yeah. It wouldn't have been uncommon to have a daily standup or even kind of the expectation that as a data science manager, you probably frequently gave tasks to data scientists or related occupations on your team where

		your expectation is you're going to hear from them in a week or something on -
Catherine W.:	00:35:56	Yeah, like, "Hey, we have this big open problem, go see what we could possibly do. "
Jon Krohn:	00:36:00	And something that you and I over Zoom last week when we were chatting about what we could cover in this episode, you and I both flamented how we miss the in - person whiteboarding on problems, which for me, I mean, there's actually no reason why we couldn't still have that today, but it seems like you and I both as individuals just in kind of the professional choices we've made up until the pandemic, like I spent more time at a whiteboard working through problems with data scientists on my team than writing code and I loved doing that because it's this amazing experience to get this kind of mind meld with a bunch of different people and when we would do this at that time, the startup that I was at, I was a chief data scientist there at a startup called Untappd and we were at a WeWork in Midtown Manhattan and whenever somebody on the data science team got into a problem where they hadn't been making progress and you could kind of tell that you could kind of intuit that you should be able to make some progress here.
	00:37:04	And then so I would reserve a meeting room that had no TV and no one could bring their laptops. You would just bring a notepad, a pen and you'd kind of the person who's running the problem, I'd be like, "All right, get up at the whiteboard and just start drawing pictures and bulleting what you're trying to do, get it into all of our RAM in our brains." And 100% of the time that I did that, we came out of that meeting, it might be an hour or two and we'd come out of it with a clear direction of, "Okay, these are the experiments you need to run or this is what you need to look into." And it would always unblock the problem. And I miss, I love, love, loved having smart, funny people and having all of our brains intermingled like that.

Catherine W.: 00:37:51 Yeah. That kind of intense in - person collaborative problem solving, it's one of a kind. I also have fond memories, maybe not quite as structured as you just described, but the AppNexus offices were on 23rd Street in Manhattan above the Home Depot and these huge windows. Beautiful. Beautiful office. Down and watch people shopping at the Home Depot below. But yeah, I have various memories of being different conference rooms and everybody looking at a whiteboard and thinking hard or I don't know, really fun. I still have pictures on my camera roll from some of those whiteboards because you take a picture to immortalize it and like, oh yeah, what were we thinking about?

Jon Krohn: 00:38:30 Exactly. Yeah. But anyway, so I think I've completely taken your conversation off the tracks where you were kind of going through early stages, AppNexus, growth there and then how data science has evolved in the past decade to where it is today, but how that lines up with your career.

Catherine W.: 00:38:52 Yeah. So my career journey starting from math, moving into data science then went kind of as my teenager says, wobbly wiggly and the latter chapter of my time at AppNexus and then Xander, data science provided the foundation but wasn't as central. It was more centered around marketplaces, but because I had been leading this data science organization that sat in the middle of AppNexus and saw both all of the technology and algorithms supporting the buy side, buyers of slots on advertising and the sell side, publishers selling slots, I was able to see how both sides were shooting each other in the foot. And so that positioned me then to be, I made marketplace czar for a while at AppNexus to try to think about how we then optimized the whole marketplace, both sides rather than each side thinking of it as zero sum. And so in that marketplace role, again, not super data science, but we did do a number of things like rethink how we forecasted revenue, for instance, rather than separately forecasting revenue from our buy side clients and our sell side clients, recognizing that they're deeply coupled.

00:39:59 And in fact, we built out a revenue forecasting model that included some machine learning models to forecast what the transactions were likely to be at different price points and so forth based on the inventory and so forth. So there was some data science that kind of worked its way into that world trying to make sure that we were running a robust and healthy marketplace. So that took me through the end of my time at AppNexus. We were purchased by AT&T. It became Xandr. I did some more marketplace thinking there, but led their data science organization again. And when I left, I left all the marketplace stuff behind and went to Qualtrics to... I walked into Qualtrics to own IQ, which was their sort of branding term for all things intelligent. So I owned sort of a hodgepodge of different things, including Text IQ, which involves some natural language processing and Stats IQ, which of course is what it sounds like.

00:40:56 And I think there were some other IQs that were kind of part of that portfolio and took all of that and wound up building out a machine learning function and platform and team write in time for the BERT paper and for ChatGPT to arrive and actually make that a lot more relevant and immediate. So that was the progression sort of in and out of the data science and machine learning flavored things.

Jon Krohn: 00:41:22 Back to the magic of embedding, despite me saying earlier in this episode that that kind of embedding based knowledge that LLMs have, it means that machines are quote unquote thinking differently than humans are, but it is still nevertheless pretty magical, pretty wild. I mean, it isn't literally magic, but being able to experiment with BERT early encoding large language models and being able to see... It was able to do things something that up until the BERT moment I used to always try to, like you described earlier in this episode, you liked to be able to go all the way down and kind of understand the math and be able to really understand everything. Since BERT on, I don't really understand.

00:42:19 Basically up until BERT with any of the kinds of techniques we used in data science, statistical or machine learning approaches, programming approaches, data structures and algorithms, ideas, all of those things, I could kind of, okay, let me create a Jupyter notebook and write some Python code and then kind of mess around with some variables and see how that changes things. And basically since BERT onwards, I don't have the same kind of level of intuition of exactly how these wild capabilities emerge from scaling up transformers to a very large size. It is wild to me that we get the emergent properties that we have from scaling.

Catherine W.: 00:43:06 I 100% agree and I share that sentiment. Now, some of it is that it's not made available to us, right? It's proprietary and so we don't have access to the information that probably the researchers inside of those companies do that maybe they're able to, surely some of them are better able to wrap their heads around what's happening. But yeah, I agree with you. I can understand the BERT paper, but then when I see modern LLMs and I think, is that just a bigger context window? How does the architecture do that? I don't know. So it's wild. And so the top down thinking and figuring out enough of a mental model to be able to reason on is like the one thing one can do, but the embedding stuff, which you wrote a book about and I really appreciated your book as a matter of fact is really wild.

00:43:49 When I started at Qualtrics, a lot of the text IQ machinery was still syntactically based, keyword based and rules based and some of it was really good, but that just all

		gets blown out of the water once the embeddings are good enough.
Jon Krohn:	00:44:06	Exactly.
Catherine W.:	00:44:06	I was
Jon Krohn:	00:44:07	There
Catherine W.:	00:44:07	To watch that revolution.
Jon Krohn:	00:44:10	Yeah. So for a period of a few years, data scientists had and all the related occupations, so I'm including in data science. And by the way, one of your recent responses, you used not super data science in a sentence. And I think that that is the first time in over a thousand episodes of this show that someone has said super data science on the super data science podcast and not meaning anything to do with our name, but just be saying, "That technique is not super data science." Yeah, first to organic use. Yeah. Yeah, exactly. Yeah. You talking about Burt reminds me how there was this rich period of a few years where data scientists could be overhauling processes that did not have embeddings in them and find a way to have a BERT model and even very small... It can easily run on a single GPU in real time, these small encoding models, but being able to unleash magical, not literally magical, but magical feeling capabilities within some kind of workflow or some kind of platform where previously it was just unimaginable and you had these kinds of hardcore rules like you were saying, like, "Oh yeah, you had all these if - felt statements or clever code written to be able to handle some situation pretty well." But then you're like, "Wow, if we have an embedding model in here instead, we can do so much more and it's so much more reliable." And that was a fun time and that kind of went into, that was associated with maybe a slightly longer period where downloading open source model weights, fine tuning those to specific tasks and

being able to out compete GPT3 or GPT4 on some capability was a fun thing that we could do as data scientists for a while.

00:46:08 But now how often, I don't know, listener, if you have the opportunity still to be training LLMs and having those be outperforming the capabilities of what you're getting from an off the shelf open source model or a proprietary API, wow, cherish that because I don't think many of us are doing that anymore. Your thoughts, Dr. Williams?

Catherine W.: 00:46:29 No, that's exactly right. That is what we did a lot of at Qualtrics is fine tuning because the big problems were categorizing open text comments, like customer experience type comments, sentiment, topic analysis, et cetera, and fine tuning to get the exact ontology right if it was a topic-based model or something like that, that was a big deal. And now I'm sure zero shot... I mean, I'm not there anymore, obviously, but I'm sure zero shot capabilities just wipe most of that away. Now, there still is a cost element. I do think that the bill hasn't come due for modern LLMs fully. And so I think that there will be another wave of creativity I think when true costs truly come to light and companies are really bearing them. So who knows? Stay tuned.Maybe there is some fine tuning to be done down the road.

Jon Krohn: 00:47:19 So I would love for you to be able to explain to me why my thinking is wrong on this, but I don't think that that is going to become a problem in the future because... So people make that argument with... So yes, VCs are subsidizing our proprietary LLM API calls today for sure. And so people will often... The most common parallel I hear is how cheap was an Uber when Uber first came to New York and they were hemorrhaging money and now what are you paying for Uber and how profitable is that business? But the economics of LLMs is obviously very different from when you're having to, at least for now, have Uber drivers that need to eat and expensive machinery. I mean, obviously there is still expensive machinery with GPUs, but if you take... So like say today, the state of the art at the time of us recording, the state of the art AI capabilities probably Methos fable f model, who knows what it'll be when this episode comes out in a few weeks, or hopefully we have access to those models again by the time this episode comes out.

00:48:22 But if you say... So the trend has been now for several years that if you fix some level of capability, so today we say, okay, what does Fable5 cost me today? The trend has been that two years later, that same Fable five capability will cost 1% of what it costs today.

00:48:48 And this is due to distillation techniques and other tricks, engineering tricks that people come up with and the DeepSeq popularized a lot of last year. There's always tricks coming up that allow us to be way more efficient with compute and it seems like that drives this underlying... And like that's a huge multiple, two orders of magnitude cheaper two years later and there's a lot that we could do with... Fable five capability allows me to do tons to automate tons of workflows and if I can have that same capability at 1% of the cost two years from now, I don't think it matters that much if VC subsidies are no longer happening. I don't know, but I'd love to hear.

Catherine W.: 00:49:27 No, I actually agree with you entirely and I guess so fine tuning models probably has gone the way of the dinosaur for most use cases, but it's exactly that pressure to figure out the distillation techniques, to find the cost optimized ways of running some of the current frontier things.That's the creativity that I actually mean is like there continues to be this pressure to figure out the cost optimize, the cheap, the sustainable way of using this technology, I think. So I

Jon Krohn: 00:49:59 Guess that's

Catherine W.:	00:50:00	Still... Some of that's happening inside of the big AI shops and some of it's probably happening outside, like you've got the deep seeks of the world and others outside too.
Jon Krohn:	00:50:08	Yeah. It's the open source always nipping at their heels, which I think also will provide a lot of price pressure because you're like, okay, well, this midsize Quen model does everything that I need, so I'm just going to use that. So we should get to what you're doing now at Candid.
Catherine W.:	00:50:24	Sure.
Jon Krohn:	00:50:26	So tell us about Candid and what they're up to and how you're able to improve things, automate things as the chief data officer there.
Catherine W.:	00:50:38	Sure. So first, let me tell you a little bit about Candid. It is a nonprofit that provides data to other nonprofits, essentially. So you can think of it as a meta or an infrastructure kind of a nonprofit. And so in particular, it provides comprehensive data about the social sector, like information about nonprofits, foundations and grants. And so it makes it easier for nonprofits and funders, largely foundations to connect. The organization was formed about six years ago when two preexisting longtime organizations merged. Those were called Foundation Center and GuideStar, one of which collected comprehensive information about nonprofits. The other one collected comprehensive information about foundations, the two married, and now we have sort of the world's biggest, richest collection of information about each other. So my job as CDO is I look after lots of different aspects of data within the organization. This is the first role I've had, data role I've had where the data that we work with wasn't instrumented by us, right?
	00:51:45	It's an external thing that we get largely from the IRS and some other governmental organizations. Some of it we collect directly from nonprofits who fill out profiles and

tell us information about what they're up to. But since we don't instrument it, the quality and cleaning is a much bigger issue because when you instrument your own data, you go correct the mechanisms to make sure that the data that you get downstream meets your needs and that you're able to analyze and so forth. But when what's upstream is an IRS form and you don't know what people were thinking when they filled these things out or they fill it out wrong or whatever or make errors, then that just creates a whole bunch of additional challenges. So what I focus on at Candid is making sure that the foundation works, that when we pull in the data and we're cleaning it and augmenting it, enhancing and sort of extracting meaning from it that then can be interpretable to our clients, that all happens smoothly and efficiently and so forth.

Jon Krohn: 00:52:46 Sounds fascinating and I guess it must be nice to be working somewhere that's doing so explicitly doing social good. I think a lot of people we've had organizations going back a decade in New York, we had these like data for good organizations and I think there's a drive from a lot of people in our space to be working at an organization that is so explicitly doing something good as opposed to there is value in having a more efficient ad marketplace, for example. Is it an arbitrary example of a thing that someone might do? For example, yeah. But there must be something really nice about doing something socially good, so obviously socially good.

Catherine W.: 00:53:30 Yeah. Now on the one hand, one can argue that at least at AppNexus, part of the mission was to support independent journalism and like independent journalism, ad supported would be good for democracy, et cetera. So I was able to sleep at night and similarly at Qualtrics, helping organizations have better customer experience and employee experience helps humans. So it wasn't like I felt like I was dirty necessarily, but there is then still a distinct difference between working for an organization that is explicitly mission aligned, not trying to enrich any shareholders in any fashion. And really its goal is to just try to help end nonprofits, help their communities, right, like get funding to those nonprofits so that they can do the good work that we need them to do because actually nonprofits, I think many people are unaware how much nonprofits do in society. They're largely invisible.

00:54:24 They don't have big lobbying arms who talk about all their accomplishments. They're embedded all over the place doing amazing work from arts and human services and research and all kinds of things. So helping that sector run more smoothly does feel like it's mission aligned. Now that said, Candid's role in it is a little bit indirect. So in some ways we're supporting a marketplace with information and trying to help that marketplace be more efficient and equitable and that's great, but it's the upside and the downside of being in a systems level role of on the one hand you're helping a whole sector and on the other hand, it feels sometimes like, "Oh, are we helping anybody? I want to move the needle more." So it's that tension.

Jon Krohn: 00:55:11 I see. It is a little bit abstract. So yeah, even though it's explicitly this mission and you know, you can and can't pin down like, "Oh, who is the orphan who's being safe from a fire today because of my tool?"

Catherine W.: 00:55:24 Exactly. Like this nonprofit was able to use our tools to figure out how to get a grant and they got that grant and then they were able to do their work. But that's a lot of hops of causality where I'm like, "Well, but it adds up." I will say the community of people both inside the organization and in the broader sector, that sort of broader sense of mission alignment is just really lovely. I hadn't worked in a nonprofit space before I'd volunteered, but it's nice. It's different than Silicon Valley.

Jon Krohn:	00:55:49	And you have some really cool things going on there as well in addition to some feel good factor, if even abstract, you do also have really cool things happening from a technological perspective. So for example, Candid recently announced an Anthropic partnership and the Candid MCP connector, Model Context Protocol, offers real time verified data directly inside AI tools like Claude, which is a pretty cool thing to be doing, all this cutting edge stuff.
Catherine W.:	00:56:19	Yeah. So it's really exciting to be able to be pushing out some of these modern capabilities. The history of Candid, back in the old days, Foundation Center, one of our precursor organizations published books, like literally books with the information and then they came to digital age and they published CDs and then everything went to the cloud and we have SaaS subscriptions for our customers where they can log into our system and poke around and search directly in our software. But now as people are moving to getting their information from Claude or from ChatGPT, we're trying to move with the times too and have our information be where they are and where they're finding where they're looking for information. So the stars aligned there. We were experimenting with some of the modern MCP technology and got a call from Anthropic and it seemed like a perfect opportunity to try it out and be part of that leading edge.
Jon Krohn:	00:57:10	Love that. As we start to wrap up your episode here, I'd like to take advantage of the wealth of wisdom that you've developed in 12 years of data and data science leadership, again, about as much as anybody could possibly have in the world in our field. And so I'd love your thoughts on with the amount of success you've had, things like an acquisition by AT&T when you're a C-level executive and lots of C-level VP roles, how can our listeners maybe who aren't yet leaders or like to be bigger, better leaders than ever before, do you happen to have any guidance for them?

Catherine W.:	00:57:54	The one piece of guidance I would give, which also mirrors sort of my own experience is if you want to lead, you need to think about the environment the next level up from you. What are the constraints and the problems being faced? If you're a director, what are the things happening at the VP level? What are the problems and the concerns and the constraints and people not sleeping at night because of? And think about the solutions and what you can provide in that context. If you're an IC, think about the problems that your manager is facing, et cetera. I think that is what got me invited into rooms. That is that context helped me put forth ideas that were helpful in those rooms and that's what ultimately sort of propelled my career there. It's not always the most relaxing path. Sometimes it's nicer to just think about the piece of code in front of you and doing a good job on that.
	00:58:46	Maybe you don't want to think about the problems one level up, but if you do want to lead, that's the way to start thinking about it, I think.
Jon Krohn:	00:58:52	That is a great answer. Thank you. And then we talked about the transformation of data science over the course of your career as a data science leader. What do you think is going to happen next and how can all of us be setting ourselves up for success in the future, this LLM enabled future, especially where so many of the technical skills can be handled autonomously by machines?
Catherine W.:	00:59:19	Yeah. I really do not have a crystal ball on the... We're in a complex dynamical system with a bunch of externalities and I do not know I can't magically do it, but I do have a but. I think the thing that I think will hold true for at least some amount of time is the idea of investing in your own models, right? So when you're using LLMs and you're outsourcing and you're building agents, that's all great, but make sure that you are also retaining the information about what's happening. Don't outsource, grow along with. I think that's got to be critical because then that's

		going to give the fluence You to build the next thing and understand the next thing and the next level of complexity. So I guess that would be advice I would give for navigating the tidal wave.
Jon Krohn:	01:00:09	And so when you say maintain your own models, you mean mental models?
Catherine W.:	01:00:12	I mean mental models, yes. Not your own
Jon Krohn:	01:00:14	LOL. You need to have your own open source LLM learning everything that you're learning.
Catherine W.:	01:00:20	Go get a data center and then... Yeah.
Jon Krohn:	01:00:23	Exactly. That's a good idea too. That's a pretty safe bet here is have a data center. That's a good way. Yeah. Have a cutting edge AI data center and you should be set for at least a few years. Really cool. Thank you, Dr. Katherine Williams for coming on the show. Before I let my guests go, I always ask for a book recommendation. Do you have anything for us?
Catherine W.:	01:00:45	I do. I read a book a couple years ago that lives rent free in my head and I think about it all the time and it's called A Brief History of Intelligence. I love that book. I think it's really profound in helping conceptualize what intelligence is and how humans embody it and how AI might embody it and what the future might look like. So highly recommend.
Jon Krohn:	01:01:09	The author is Max Bennett and I will have this in the show notes for people. I also do highly recommend it. It's a very popular book for being such a technical subject. It has over 5,000 reviews on good reads and it's a brief history of intelligence, evolution, AI and the five breakthroughs that made our brains. I was talking to a friend of mine who's an Oxford neuroscience professor, but he works at the intersection of neuroscience and AI. And I told him that I was thinking of writing this book

		and then he said, "Well, that book already exists." And it's pretty good. I think something that's really remarkable about this book is that the author does not come from... He's not an academic. He's a startup co-founder from New York.
Catherine W.:	01:01:56	He comes out of our world, but he got interested and started digging and he did all the digging and then shared his findings with us. And it's all the questions that I had that he just... Anyway, it's fantastic.
Jon Krohn:	01:02:08	It's really good. It's almost as though a PhD is just anybody could do the work of a PhD, whether you formally get that piece of paper or not, lots of clever people out there who could be doing amazing things. Yeah. It's just a credential
Catherine W.:	01:02:28	That describes persistence.That's it.
Jon Krohn:	01:02:32	All right. And for people who would love more of your brilliant thoughts after this episode, is there anywhere we can find you online?
Catherine W.:	01:02:39	I am not really a social media person other than I do have a LinkedIn presence. And so if you really want to talk to me, come find me there.
Jon Krohn:	01:02:45	We will have your LinkedIn in the show notes for sure. And that is the most common answer by far on this show these days. But really appreciate you taking the time. I don't think you do a lot of these kinds of interviews these days, do you? Mm-mm. So we're very lucky to be able to get your thoughts. Thank you so much for agreeing to be on the show and yeah, maybe in a few years we can check in and see how things are coming along.
Catherine W.:	01:03:15	Sounds great. Thank you so much, Jon. This has been a pleasure.

Jon Krohn: 01:03:19 What a wonderful episode today in it. Dr. Catherine Williams detailed how she was hired at AppNexus purely for her mathematical fluency despite having almost no data skills and how she and her team taught themselves modern machine learning as Hadoop and ever larger datasets arrived. She talked about what Candid does as a nonprofit supplying data to other nonprofits. Why deep technical understanding still matters even when AI can do the math for you because that fluency is what lets you build the mental models to reason at the next level up. Her view that the most valuable data professionals going forward will be the ones who can move up and down the whole stack of complexity rather than specializing in a single slice and how fine-tuning has largely gone the way of the dinosaur now. The zero shot capabilities have wiped it out. As always, you can get all the show notes, including the transcript for this episode, the video recording, any materials mentioned on the show, the URLs for Catherine's social media profiles, as well as my own at superdatascience.com/1011.

01:04:16 There's a fun binary episode number for you today. Thanks, of course, to everyone on the SuperDataScience Podcast team, our podcast manager, Sonja Brajovic, media editor, Mario Pombo, our partnerships team Natalie Ziajski, our researcher, Serg Masís and our founder Kirill Eremenko, our founder, for producing another super episode for us today, one that's really out of this world. For enabling that super team to create this free podcast for you, we are deeply grateful to our sponsors. You can support the show by checking out our sponsor's links, which are in the show notes. And if you'd ever like to sponsor an episode yourself, you can get the details on how by heading to Joncrone.com/podcast. Otherwise, please help us out by sharing this episode with folks that would love to hear from Dr. Williams and such a great episode like this one we had today. Review the podcast on your favorite podcasting app or on YouTube.

01:05:10 For some reason, it seems like text reviews on Apple Podcasts are especially helpful for people discovering this show and seeing that it might be one that they'd like to listen to. Subscribe if you're not a subscriber, obviously, but most importantly, just keep on tuning in. I'm so grateful to have you listening and I hope I can continue to make episodes you'd love for years and years to come till next time. Keep on rocking it out there and I'm looking forward to enjoying another round of the Super Data Science podcast with you very soon.