# SDS PODCAST EPISODE 909: CAUSAL AI, WITH DR. ROBERT USAZUWA NESS

Jon Krohn: 00:00:00 This is episode number 909 with Dr. Robert Osazuwa Ness, Senior Researcher at Microsoft Research AI.

00:00:14 Welcome to the SuperDataScience Podcast, the most listened to podcast in the data science industry. Each week we bring you fun and inspiring people and ideas, exploring the cutting edge of machine learning, AI, and related technologies that are transforming our world for the better. I'm your host, Jon Krohn. Thanks for joining me today. And now let's make the complex simple.

00:00:48 Welcome back to the SuperDataScience Podcast. In today's episode, we're covering the fascinating field of causal AI and we have exactly the right guest, Dr. Robert Osazuwa Ness, to be leading us on that journey. Robert is a senior researcher at Microsoft Research AI. His research focuses on statistical and causal inference techniques for controllable human aligned multimodal models. He's also founder of Altdeep.ai where he teaches professionals advanced topics in machine learning. He holds a PhD in statistics from Purdue University in Indiana. In addition to the above, Robert is the author of the book Causal AI, which was published by Manning in March. I will personally ship five physical copies of Causal AI to people who comment or reshare the LinkedIn post that I publish about Robert's episode from my personal LinkedIn account today. Simply mention in your comment or reshare that you'd like the book, I'll hold a draw to select the five book winners next week. So you have until Sunday, August 3rd to get involved with this book contest.

00:01:51 Today's episode will resonate most with hands-on practitioners like data scientists, statisticians, and AI engineers. In it, Robert details the three rung ladder of causation that determines what types of causal questions you can actually answer with the data that you have, the surprising connections between Bayesian networks,

graphical models, and modern causal AI, why AI systems have been dominated by correlation-based learning and what's stopping them from adopting causal reasoning like humans and animals naturally do, how tools like PyTorch, Pyro and Du AI are revolutionizing causal inference by separating statistical complexity from causal assumptions and how LLMs like GPT4.0 Can act as causal knowledge bases outperforming traditional causal methods in some scenarios. All right, you ready for this deep episode? Let's go.

00:02:43 Robert, welcome to the SuperDataScience Podcast. It's a delight to have you on the show. So we have a fantastic episode scheduled today, I can't wait to get into the questions. We also had a lot of audience questions. We're going to try to get through as many of those as we can at the end the episode, so those who asked their questions on social media beforehand, yeah, we'll get to you. But in the meantime, I've got tons of questions. So our researcher, Serg Masís, prepared amazing research on this episode, I was so excited when I worked through it.

00:03:13 You're a researcher at Microsoft Research focused on causal AI and probabilistic machine learning. You're also founder of an educational platform called AltDeep, I'll talk about that more a bit later, and you're the author of the book Causal AI, which was released just a couple months ago by Manning, fantastic book. And in a glowing recommendation of your book, the Turing Award winner and creator of causal calculus, none other than Judea Pearl, who many of our listeners will know, said that your book Causal AI is a timely resource for building AI systems that generate and understand causal narratives.

00:03:48 Could you unpack for us, Robert, what it means for AI systems to generate and understand causal narratives? How does the word narrative fit into all of this? What does

it mean for it to be causal? Yeah, this seems like a kind of an interesting starting point.

Robert Ness: 00:04:04 Yeah, I mean, I think when Judea Pearl was saying that he was looking at the chapter in the book's last chapter that is looking at the kind of intersection between causality and foundation models, namely in this case large language models. And I talk a little bit about multimodal models as well.

00:04:22 And it's interesting because I think the first time I got the inkling to write the book was after reading his book, The Book of Why. So most of that book is really kind of bread and butter causal inference, but written as a popular science book. But he has one snippet in there, very short, but it talks about what it would be like if there were a robot that could understand causality, there was an artificial intelligence that had the ability to do causal reasoning. And it was a cool example, it had to do with a robot waking you up in the morning because it decided to vacuum and then you were upset about it and it had to understand why you were upset in order to improve its behavior. And I was thinking to myself, "Well, yeah, and? What else?" So I said, okay, well we need a book that's really taking an AI approach to looking at causality and is using tools like PyTorch to actually write algorithms. And so maybe that was just the point that the book was initially conceived.

Jon Krohn: 00:05:31 Very cool. And so the concepts of causality, it seems to me like they should be more intuitive and analogous to human reasoning than something like correlation. And so in that example that you just gave there where you're talking about a vacuuming robot, when a human or even when my dog makes a mistake, it's pretty easy for us to have an intuition around what's causing that. My dog knows that I'm annoyed at him because he's barking, he knows that I'm not just arbitrarily annoyed and he knows

that if he stops barking, I will stop saying, "Stop barking." And so animals seem to kind of intuitively have these built-in systems for understanding causality, yet the AI systems that we've built up until now, they're dominated by correlation-based learning. So why do you think AI systems developed like that and what's stopping it from adopting more causal building blocks?

Robert Ness: 00:06:36 Well, if we look at traditional causal inference, as you might see in a econometrics textbook or a causal inference textbook that's targeting people who work in epidemiology or taking more of a statistics approach, and so there are these models that are very much rooted in classical statistics and they're thinking about, okay, well, there's some kind of causal relationship between these variables, but there might be some confounding, which means that they might share some common cause, statistical association is coming both through this causal relationship that they share as well as through this common cause and we want to figure out how to distill the statistical association coming from a direct causal relationship from the overall background noise of statistical dependence that includes that causal dependency as well as a non-causal dependency through that confounder.

00:07:40 So everything I just said there is very statistical language and packed in with some causal assumptions about the structure of causality between these variables. And insofar as we do have some use of, say, for example, deep learning and causality, it was still very much based on this kind of basic problem of having this kind of statistical association that's due to causality and due to non-causal relationships. And then what deep learning was doing when it was introduced into this space was to say like, "Well, let's make sure we can scale it up with more data and work with, say, non-linear relationships,"

maybe you work in higher dimensions, etc. But still basic underlying framework.

00:08:32 Now, kind of going back to your example about how humans or animals reason about causality, there is a space in research that thinks about this thing a lot, the ways that animals and people reason causally or otherwise quite a bit, which is in cognitive science, right? So cognitive science researchers try to understand, hey, how is it that people are reasoning about cause and effects? How are they making causal decisions? How do they understand why some outcome happened? In other words, what are the causes or what is the main causes that led to this outcome, et cetera? And it's a fascinating area of study because it's less about understanding actual ground truth, right?

00:09:28 So for example, if you think about causality and practical terms, let's say that there's a pandemic and people are sick and I want to propose a vaccine that's going to help people. Well, you really want me to be right, right? If I say that there's a causal relationship between administering this vaccine and people not dying of this illness, the burden of proof is quite high. And so we're looking at that same type of rigor that we take in statistics that we may be thinking about statistical hypothesis tests, thinking about falsifiability, et cetera. But in this cognitive science field, the question is not what's true in the world, but the question is how do we write algorithms that do what we think is happening in those people's heads or in those animals' heads?

00:10:23 And it's an interesting space because it turns out that humans, despite what kind of conversations you might have at your Thanksgiving table, humans reason really well about causality, relatively speaking, especially about what some call intuitive physics, right? You mentioned your dog, if you walk into a room and something has been

knocked down off of the table and you can see some evidence about the spread of the debris on the floor, you might make some pretty good inferences about whether it was your dog or your kid that knocked it down and how it was knocked down, et cetera.

00:11:02 Similarly, humans tend to be really good at what some call folk psychology, which is to say you and I could be sitting together in a cafe and watching people across the cafe having an argument in hushed voices, and without actually hearing what they're saying, probably get a good guess about what it is they're arguing about, or at least a theme about what they're arguing about. And so there are certain domains where humans tend to reason about cause and effect fairly well. And so one of the interesting things that we can talk about from an AI perspective is to say, okay, how can we write algorithms that emulate those reasoning processes? But this is in contrast to, say, classical statistics, which is much more concerned about type one error, a false positive.

Jon Krohn: 00:12:03 Right, right, yeah. So I guess what you're saying is that there is a branch of study in cognitive sciences mostly where people are trying to figure out how the way that humans animals have these intuitions around causality, we try to figure out some ways of packaging that into the way that models work so that they can do some causal inference, whether that's a statistical model or machine learning model. And so it sounds like that's a relative niche, a relatively niche academic pursuit from the way that you're kind of explaining it. So if the way... Yeah.

Robert Ness: 00:12:43 I'm not sure, I think so. I mean, I think particularly in AI, people are stealing ideas from other sciences all the time, right? In some sense, reinforcement learning is like the Pavlovian learning that, you mentioned your dog, that a dog might learn in terms of responding to stimuli without actually really understanding a lot about how cause and

effect is happening under the hood. Or we borrow a lot from physics and other domains when designing loss functions and machine learning architectures. So I think this is just another type of that approach, which is to say, "Hey, these guys are doing something here. What happens if we combine it with the stuff that we're doing?"

Jon Krohn: 00:13:32 But I guess where I'm going with my question is would you say that the predominant approaches to causal AI today, maybe they don't really have to do with these cognitive science approaches, maybe some of them do, but what are the predominant ways, and this is maybe too big of a question and maybe we'll kind of get to this answer through other questions that I ask today, but what are the predominant ways that we add causal elements into an AI system?

Robert Ness: 00:14:05 So one of the things that I observed in practice that made me really want to write the book was that people weren't drawing the connection that seemed pretty obvious to me between kind of graphical causal inference, like causal inference with a DAG directly, basically a graph, and probabilistic graphical models in statistical machine learning or more generally probabilistic models, including deep probabilistic models insofar as they had very common origins. We already talked about Pearl, and Pearl was a huge contributor to the world of Bayesian networks and Bayesian networks in terms of if you take a Bayesian network and you interpret the edges in the Bayesian network as being causal, you get a causal model. In fact, a lot of the earlier kind of causal graphical models were based on the theory, it was very much part of probabilistic graphical modeling theory. And the du calculus, for example, that you mentioned is based on ideas of what happens when you fiddle with a graph in ways that simulates an intervention, for example, actually doing an action on a data-generating process.

00:15:34 And so at some point that area of probabilistic graphical models, we started developing this software network which implemented, right? So some of your listeners might remember tools like JAGS or BUGS that where we take something that looked like a graphical model and write it out as a program, and it was using inference like just sampling-based inference. And the interesting thing about some of these probabilistic programming approaches is that they would use for loops, they would use control flow that you wouldn't really be able to do in a conventional Bayesian network, but this technology continue to develop into tools like Stan, for example, and then also tools like say WebPeople or Gen and Julia, other probabilistic programming languages, languages like Pyro, which use PyTorch or in the case of NumPyro use NumPy and JAX and the inference engines allowing you to import learning algorithms from deep learning, for example, using stochastic variational inference.

00:16:55 But these all had a common route. And some of those probabilistic modeling approaches, they were very much adopted, in fact developed by people in the cog sci space. In fact, even going back to BUGS and JAGS, and there was some famous BUGS or JAGS book back in the day that was written by cognitive scientists, even though most statistics departments were using it. And so now this technology has developed quite well. You can go into PyTorch and using a language like Pyro or an extension of PyTorch like Pyro, you can write fairly sophisticated deep latent variable models that are using modern deep learning architectures, doing things like variational, using deep learning to do inference with things like stochastic variational inference, but are just as capable as of modeling any Bayesian network you saw in the '90s and thus capable of building a causal model that's built on a graph.

| | | |
|---|---|---|
| | 00:18:00 | And the fact that they can deal with latent variables makes them pretty powerful because I think one of the reasons that causal inference in the field kind of moved away from generative models is because they didn't do a very good job at dealing with latent variables, which is what a confounder is, it's something that's confounding your causal inference because it's unobserved and you need to deal with it. But now that's no longer an issue, right? These models can handle that kind problem fairly well. |
| Jon Krohn: | 00:18:26 | Okay, nice. Let me try to summarize back some of the things that you said for our audience. So a lot of it made sense to me. I come from a time when I was doing my PhD many years ago now, JAGS and BUGS were the kinds of tools that people were using for Bayesian inference. Now we have come into a time where tools like Stan are more common. You also mentioned Pyro, NumPyro. For those of us who work in Python, which is a lot of our listeners. |
| Robert Ness: | 00:18:51 | PyMC too, PyMC also has some causal abstractions. |
| Jon Krohn: | 00:18:55 | For sure, PyMC. So if you want to learn more about those kinds of tools, we had an amazing two-hour long episode on Stan with Rob Trangucci back in episode 507 some years ago now, four years ago now. And we've more recently had episodes on PyMC, so I can really quickly look that up. So for example, episode 585 with Thomas Wiecki, so Thomas Wiecki is CEO of PyMC Labs. Yeah, so we've had a great episode on PyMC there as well. |
| Robert Ness: | 00:19:30 | And to plug his stuff, they added the main causal abstraction being something that can model an intervention called a du operator, they added a du operator to PyMC. And so if you're a PyMC fan and you search that, you'll find some pretty good PyMC tutorials |

for causal reasoning, one I believe for uplift modeling or multimedia mixture models.

Jon Krohn: 00:19:55 Nice. And so things like this do functionality, what that is allowing us to do with a Bayesian model is to be able to try to simulate that a variable, so if you have a whole bunch of data where all the data have already been collected, and so you can't really run a real experiment in the real world, this du operator allows you to kind of simulate one of your variables as potentially being the instrument, being the cause like, going back to that vaccine example, if you're running an experiment, you give people a placebo, you give some people a placebo, you give some people the real treatment, and that is kind of the ideal, that's what we'd ideally like to be able to do is to run an experiment to determine whether that treatment does actually, cause in your case there a reduction in disease rates, that it's an effective vaccine.

00:20:58 But a lot of the time we've already collected the data, we just have a bunch of data. And so it sounds like what you're saying is that something like the du operator that we could implement in a tool like PyMC would allow us to in some circumstances simulate afterward post-hoc whether that variable is in fact an instrument. Is that right?

Robert Ness: 00:21:19 Yeah, so a randomized experiment, obviously it's the gold standard and the reason it's the gold standard is because, well, barring certain things that can come up in the actual implementation of the experiment, you don't need to make many causal assumptions between the treatment and the outcome. In a causal model, any model that allows you to model intervention, we can call that a causal model. And what the causal model is doing is in exchange for explicitly providing some assumptions about the causal structure of the data generating process, we're getting the ability to simulate the effects of an

intervention into that process like we would get by randomizing in a clinical trial or in a randomized experiment, a randomized controlled trial. And so it's not a free lunch here, it's not like these causal models are doing something magic that means you don't have to run experiments. It just means that, in exchange for adding some assumptions about the causal structure of the system, you can simulate what would happen if those assumptions were true.

Jon Krohn: 00:22:36 Nice. Okay, I got you. And so it's interesting because so far in this episode we've been talking about Bayesian libraries that I'm familiar with, it seems like a lot of the examples are in a Bayesian realm. So I'm hoping that there's a funny answer to this or that you find this a funny question, but based on what you said so far, why wasn't your book called Causal Bayesian Statistics instead of Causal AI? So what's the distinction there?

Robert Ness: 00:22:59 There's a little bit of Bayesianism in the book. And the way I think about it, I didn't go too far into Bayesian domain just because, well, number one, it's important to disentangle things when you can, right? And one of the things that what I like about, or what I was trying to accomplish with my book was to say, "Here are the kinds of assumptions that you're doing with statistics, and here's the kinds of problems that you're dealing with when you're trying to scale up to larger data or you're working with higher dimensions, and here's the causal problems that you're trying to solve." And rather than kind of sloshing these all together and trying to figure out, a lot of these books, it feels like you're having to get a whole new master's degree just to solve some of the problems in the book, we can say, "All right, well these things we're going to separate out. You can either use a library to do this for you, or you can rely on your existing knowledge or go in deep dive into that if you need to, and then the causal stuff is over here."

00:23:56 So I think of Bayesianism as being in that kind of statistics box, which is to say that... But another way of thinking about it is that one of the things you're doing with a Bayesian model is that you're injecting assumptions about the data generating process into your model. In this case, typically what's happening is that you're injecting your assumptions in the form of priors on unknown elements of the model, and then it's kind of reflecting your certainty about the values or the structure of those elements as a model or as opposed to what's actually true in the data-generating process. And then from the causal perspective, we're injecting assumptions, but usually in the form of causal assumptions, say for example with a causal graph or alternatively in the form of, say, mechanistic assumptions between how the variables are connected. So both the Bayesian and the causal approach are thinking about we need to inject some assumptions into our model to get better inferences.

Jon Krohn: 00:25:16 Okay, nice. So when we're talking about causal AI, what kinds of libraries would we use to do causal AI? What kinds of problems can we solve with causal AI that we might not be able to with other approaches? Maybe there's examples from your book that you can provide us with that are kind of illustrative, I feel like I'm pronouncing that word wrong, but that illustrate the value of causal AI. What kinds of circumstances could our listeners find themselves in where they should be thinking about using a causal AI approach? What do they get if they do that? And how would they do it, what kinds of libraries or approaches would they use?

Robert Ness: 00:25:54 So there's the traditional answer to the question, which is to say that anytime you need to do an inference or that's not just about, say, predicting some variable given another variable, but actually trying to understand what would happen if you intervene in that system. So if for example, again, I needed to figure out if this vaccine, if

some medicine or some supplement that people were taking were actually having an effect on some outcome that we care about, say for example, does some new exercise supplement have an effect on muscle gain, you want to isolate out all the factors that people who are already taking the supplement, are taking, maybe they're going to the same gym and they listen to the same workout podcasts and they have the same diet or something like that. And so all of these things you want to control for you would an experiment.

00:26:48 And then you say like, "Well, I can't run an experiment here, but I can make some pretty good assumptions about the way that the system's set up. Given those assumptions, what can I do?" Right? Now you're asking causal questions. Once we get into the realm of AI and the kinds of questions that we typically think about in machine learning, it's interesting, it gets a little tricky because, in my experience, tools like deep learning are really good for us brute forcing their way through a lot of problems with a lot of data. And so really what you're trying to do is to say, what are the sets of questions that can't be answered in this particular domain with just more data or with just some kind of clever architecture? When would you want to be thinking about using causal inference or causal AI?

00:27:47 I think the easiest place to start is to think about when would you use causal inference in the first place? And then thinking about how you could use AI to either scale it up to larger data sets or higher dimensional data sets or to automate it, right? Because one thing that you can do with artificial intelligence is automate decision processes. And so, again, when would I use causal inference is when I'm thinking about an experiment that I like to run, and maybe the experiment itself is expensive to run or it's infeasible, or I can run all the experiments at once, but there's some kind of opportunity cost, right?

And so anytime I can simulate the outcome of an experiment in exchange for providing some causal assumptions for how the data-generating process is set up, then I'm deep into causal inference territory.

00:28:45 And then I think there's the question of where do we start to apply causality and causal reasoning in AI problems? And a way I think about that is in machine learning, we often use this term inductive bias, right? So we're thinking about what inductive bias means is that given some data and some induction, and I was going to say, for example, a prediction that I want to make, I need to have some kind of assumptions that are guiding the direction of the inference, right? And so that could come in the form of, say, as you mentioned, Bayesian priors, could also come into form of, say the architecture of the model, say for example using convolutions with max pooling to look at invariances, the identity of shapes as they have different positions within image, for example.

00:29:52 And so what causal inference theory does is it tells us what kinds of causal questions can be answered with a given set of data and a given set of assumptions. There might be some causal questions that you can't answer given your assumptions, even if you get more data, right? Or you might say, "Well, given my data and my assumptions, I can't answer this question, so I need more assumptions or I need more variables in my data," rather than having a greater quantity of data, you have a greater spread of the data across observables.

00:30:34 And so that type of intuition you can bring into the typical workflows that we're using in say, for example, deep learning where we're saying, "All right, well, look, my algorithm seems to be able to answer some causal questions. According to the causal inference theory, these conditions must be satisfied in order for that to be true." So if I know that if I'm doing my ablation study or I'm

trying to understand if my results are going to transfer to a new domain, that intuition from the causal theory is definitely going to help me as opposed to just kind of crossing the river by feeling for stones to quote Deng Xiaoping.

Jon Krohn: 00:31:26 Nice. Great quote there. So the idea here is, so we had already covered kind of earlier that any kind of circumstance where you want to be really careful about how a variable is impacting another, not just that one is correlated with the other, but you want to be able to determine that X actually causes Y. And so it sounds like for the most part, is it often the case that there's more heavy lifting required or that you need to be more careful about the way that you collect your data or maybe the particular data that you collect in order to be able to do causal AI? Or can you just use data that you've already collected, make some assumptions?

00:32:14 I guess what I'm trying to get at here, if somebody wants to do causal AI, are they going to have to do something special with their data or the way that they collect their data, or is it kind of just a matter of making assumptions? And I realize I'm asking a big question, I don't know the answer, but maybe there's some kind of examples in your head that you can provide that either they could help answer my question and give us a clear picture of when we can use causal AI or when we can't.

Robert Ness: 00:32:42 So one thing that I think happens a lot in practice, particularly in industry, say you're working as a data scientist, you often don't get a lot of control over what data is collected, right? Somebody sends you an Excel spreadsheet or points you at some data warehouse and they say, "Make some sense of this," right? And we all know that's not idea;, that ideally we would like to be there at the point of data collection and perhaps provide some input into what variables get collected. If we're going

to be thinking about causality, yeah, we need to make sure that we are making causal assumptions that reflect what we honestly believe about the data-generating process as opposed to what's convenient to what variables we happen to collect in our dataset, right?

00:33:35 So one thing I often tell students when I teach is to say the causal inference doesn't care about what data you're stuck with. If your data is insufficient to make certain types of causal conclusions, you can't coerce the data into doing so. But we know this from statistics, right? When I was doing my PhD, one professor, he said something that stuck with me, which is to say, there is no method of statistics that's just going to remove bias from your data, right? You can get more data, you can use a fancier model, but it's not going to extract the bias. The only way you can deal with that is by, say, modeling it explicitly with some assumptions that aren't coming from the data but are coming from you.

00:34:31 And that's the same issue that we have in causality, which is to say, when we talk about confounding, does this supplement affect muscle gain or are there some confounding factors? What we're saying is it's a question of bias, there's a confounding bias, we look at this association between the treatment and the outcome and we want to interpret it causally, but there's some other signal leaking in that's biasing that conclusion. And so in that case, we either have to rely on assumptions or we have to collect more data that's covering things that we're not yet covering.

Jon Krohn: 00:35:12 Okay. Okay, cool. So are you able to give a couple case studies maybe of real world situations? It could be examples from your book or maybe examples from your research or examples from Microsoft or, I don't know, some other job you've had in the past or some client you've worked with where you're able to kind of explain to

us, "These were the data that we had, these were the assumptions we made, and this is the conclusion that we were able to draw, we were able to make this causal conclusion, and if we hadn't done that, we wouldn't have been able to achieve this commercial outcome or this research outcome," or something.

Robert Ness: 00:35:55 That's an interesting question because all the practical examples that I could think about from Microsoft are definitely confidential because they all involve products.

Jon Krohn: 00:36:04 But there must be something from your book. Everything in your book is hands-on, it's got lots of PyTorch, lots of real-world business examples, so there must be lots of examples that you've already published in there even?

Robert Ness: 00:36:18 Yeah. So in the book, one example that I lean on a lot is this, and it's a simple example because it's meant to be there to be understood, but it's an example of imagine that you're a data scientist at a gaming company, say like an online game, an online role-playing game, for example. And in this game you can go on quests, but you can also do kinds of side quests with your clan or your... What do they call that? Your crew or whatever. I'm sorry, I'm bad at video game lingo, but with your guild, I don't know. And so let's suppose that your question is, does engaging in more side quests increase the amounts of money that a player pays for on virtual items within the game? Right?

00:37:23 And so in the example I show just kind of, and the statistics all laid out pretty clearly, that if you just look at the correlation between engagement in side quests and amount of money spent on in-game items, it sounds like that there's a belief in the book, it's a positive relationship such that more side quest engagement leads to more purchases. But when we account for a third variable, in this case it's membership in the guild, so we can assume that people who are in the guild or probably going to go

engage in side quests together or discourage each other from engaging in side quests, so guild membership being a cause of side quest engagement. And of course maybe people in a guild pull the resources or say, "Hey, John, you should buy the sword and I'm going to buy the healing potions." And so it is going to be a cause of in-game purchases. And so this becomes a confounder if you're not measuring it directly.

00:38:30 And in the book I walk you through how you would get the answer wrong and report the wrong answer to your boss if you were to just say, "Look at the correlation between side quest engagement and in-game purchases," how if you ran an A-B test, a randomized test that you would get the right answer. But this would involve essentially forcing people to engage in side quest engagement because somebody logs onto the game, you're like, "Hey, here's some side quests do it," and now because you've randomly assigned them to that group and they don't want to do it. Or maybe some people are getting a better experience and other ones, and so sometimes in a real world running the randomized experiment is infeasible.

Jon Krohn: 00:39:17 Yeah, it sounds like an economics kind of situation where the human knows that they're in an experiment if they're being forced to do one thing or another. What you're trying to get at is if the human in the wild in this kind of economic free world that we live in, which the video game is simulating there on a slightly smaller scale, if you're trying to answer some question, running a real randomized controlled trial is not going to be feasible, but you might be able to use the data that you have already collected, that you've observed. So then, yeah, in your case there that you're describing, there's-

Robert Ness: 00:40:01 One thing to jump in there, there's feasibility, but there's also ethics, right? If I wanted to understand the effects of

caffeine on miscarriages, clearly I can't ethically run that experiment, not on humans. And then there's just the old-fashioned question of cost. And this happens again in not just in theoretical questions or questions in economics and medicine, but also in industry and things like optimizing a video game. But sorry, I interrupted, please continue.

Jon Krohn: 00:40:37 No, no, no, not at all. It's your episode, I'm just trying to divine some information. Okay, so now we have a clear understanding of a kind of causal problem that we might want to solve. So I'm a data scientist at a gaming company and I want to figure out whether the users of this game, when they tend to engage in more side quests, does that cause them to spend more money on in-game assets that they can be buying? And so there are potentially confounding variables out there, things like being a member of a guild that you mentioned there. And so if we weren't collecting that guild data, we'd have to have more assumptions, we'd have to basically make the assumption that being in a guild doesn't matter or not. And so this is made clear that there's a lot of assumptions, more thinking potentially about your problem you need to do if you're engaged in causal AI. So that's a great thing to understand about this.

00:41:39 But to kind of get into the nuts and bolts, your book does a great job of using PyTorch code, using examples to make causal AI or causality in general, which is often a very theoretical, difficult to understand topic, because of all of your examples and use of code in the book, it makes understanding causal AI I'm more intuitive. And so let's say that you are the data scientist, Robert, at this gaming company, what Python tools would you use to then do causal AI? How would you model this in order to come up with a causal conclusion?

Robert Ness: 00:42:24 So one of the things that I had mentioned that I was trying to do with my book was to separate out the abstractions that have to do with statistics and computing, scale it up, algorithmic complexity from the causality. And what's cool about the libraries that we have today is that they can actually help us, if we're able to separate those abstractions, then we get to focus on one thing while leaving the nuts and bolts to be handled essentially by the library.

00:42:58 To some extent, you saw this, you mentioned you interviewed somebody who talked about Stan, and what's cool about Stan is that that inference algorithm, Hamiltonian Monte Carlo, I mean, you can go in there and understand it... Well, it's physics, but it's not rocket science. I'm like, well, it kind of is rocket science. But you can still just specify your model, specify what are the parameters, what's the model, and as long as you satisfy a certain set of requirements, I think mainly that's all of these things have to be continuous, that the inference will kind of just work for you without you having to go to implement your own inference algorithm.

00:43:45 It's the same thing here, where as long as you can kind of specify your causal assumptions, in some cases in the form of a graph for example, then you can rely on, say, graphical causal inference, the inference algorithms from, say, probabilistic graphical models to handle the inference there for you. If you're implementing it in PyTorch, plenty of PyTorch examples in the book, as long as you can incorporate your causal assumptions in the structure of the model, in your algorithm and write a basic inference algorithm that has a differentiable loss function, then PyTorch is going to handle all the nuts and bolts of the inference for you. That's kind of why we invented PyTorch to say, "Well, if I can differentiate it, if I can get a gradient, then I can just turn it into an inference problem there."

00:45:02 And so I do have examples there in PyTorch that are saying, okay, well, let's not worry about whether or not we need to use linear regression here or propensity scores or double machine learning or instruments, but these are all different types of statistical methods for doing the inference you want. And you can learn all these things, great, and there's great books for that. I think I can name a couple off the top of my head. But you can also say, let's work with some libraries that are just going to handle that stuff for us under the hood and kind of treat it as either an objective function to be optimized or as a configuration in a configuration parameter in some model specification, and then focus on our ability to think causally and write that thinking down in the form of a model and focus on the actual domain that we're modeling as opposed to all the inference stuff that we need to do to get that to work.

00:46:11 And so to answer your question, I talk a lot in the book about using deep probabilistic modeling with libraries like Pyro, as well as some more conventional tools like du AI, du AI from the broader pyWise suite, which is a big collection of causal inference libraries. And so even in du AI, there are different types of statistical techniques that you can use to estimate a causal effect, but at the end of the day, you're thinking more about what are your modeling assumptions and can you answer the question given your assumptions in your data? And then if you can, you want to get to an answer, and then all of the various statistical approaches you can take to arrive at that answer, given your assumptions and your data, you kind of just toggle between them and see which has given you more stable results, for example.

Jon Krohn: 00:47:11 Okay, so it sounds like a lot of these things we could do, a lot of causal AI modeling we could do in Python with PyTorch, but that might be unnecessarily low level, and so there are other libraries-

Robert Ness: 00:47:25 High level.

Jon Krohn: 00:47:26 High level, sorry, exactly. And so there are other libraries out there like Pyro, which we talked about earlier in the episode.

Robert Ness: 00:47:35 So Pyro was just an extension of PyTorch, so that's still in PyTorch land.

Jon Krohn: 00:47:39 Oh, I see, I see, I see. And then PyWi, is that kind of a similar?

Robert Ness: 00:47:45 PyWi, you want to work with conventional numeric data or categorical data that you have in a data frame. That's definitely the way to go. You would do PyTorch if you really need to work with neural nets inside your causal model for whatever reason, or that you really need to work with, say your variable instead of being like treatment, aspirin or placebo, outcome, healed or not healed, maybe your treatment and the outcome variable, it's a vector or a matrix, right? Because you're working with, say, media for example, you're working with some kind of rich high dimensional data, there's still just variables. But now because you're working in higher dimensions and the relationships are all nonlinear and you need a lot of data, you want to be working with tools that were designed for that. Well, if you're working with the kind of things that come into a PANDAS data frame or an R data frame, things like du are the way to go.

Jon Krohn: 00:48:47 Okay, cool, thank you for that. And so my last kind of big technical question that I have for you before we get to some of the audience questions is that in the last few years, AI agents, generative AI and LLMs have all been widely featured in causal conferences and papers, and you've been actively contributing to research at this intersection. And so despite the shortcomings of large language models, of LLMs, in your paper Causal

Reasoning and LLMs: Opening a New Frontier for Causality, you've found that causal reasoning outperforms existing methods. So tell us a bit more about this, about the paper, about the relationship between generative models and causal reasoning, because that isn't something, generative models, agentic AI, they're the hottest kind of topics we can be talking about right now and that hasn't come into the episode yet. So I'd be interested in hearing your thoughts on that intersection and what's possible there.

Robert Ness: 00:49:42 So in the paper that you mentioned, we're essentially trying to interrogate the causal reasoning abilities of a large language model. In that book, in that paper, it focused mostly on GPT 4.o. So we had some benchmarks for answering causal questions, we created some benchmarks for what we call causal discovery, which is to say you have two variables, is there a causality between A and B? If so, does A cause B or does B cause A? And essentially what was happening there is that it works pretty well. It was doing so without data, so as opposed to trying to infer a causal relationship based on statistical association data, it's inferring a causal relationship based on relationships it has already learned semantically... Or maybe semantically is wrong. Correlations it has learned between tokens in natural language text from the training data.

00:50:54 And so in that sense, what the large language model was doing was kind of acting as a kind of causal knowledge base, right? So if you wanted to infer whether or not there was relationship between, let's say for example, in biology, some genetic product like a protein and some condition, and some papers were written on this, and maybe the LLM is able to kind of synthesize to some extent the information across all of these papers into some kind of coherent statement about the causal

relationship between these two variables. And so there was a lot of that.

00:51:44 So in the last chapter of the book, I talk about, in the first half of it, I say, how can we use these foundation models to be oracles for causality? So if, for example, I want to know what the causal structure of this system might be, I can prompt it to propose a DAG and then giving that DAG, I can prompt it to propose some du AI code that allows me to implement that DAG in Python and run the analysis. If I get a bug on the code, I can plug it back into the AI and tell it to fix it for me. And then similarly, a lot of what we're doing in theory is to take some things that we would say in natural language like, "Hey, I think this vaccine might have a positive effect on preventing this illness," and actually what we have to do is formalize those into variables and relationships between those variables.

00:52:38 So there's this step of taking your assumptions and formalizing them into math that you can write down or symbols that you can write down and then apply operations to, so you can apply the theory and everything. And the language model is pretty good at that, right? So if I say I talk about this counterfactual idea called the probability of necessity and saying, if I were working at Netflix and I want to understand if I did some promotion and people watched this show and I want to understand if they would've still watched the show had I not done the promotion, how would I formalize this in causal terms? And then the model does a pretty good job in doing that for you.

00:53:28 Of course it still hallucinates too. So I also asked, "Okay, well, now that I formalize this in the math, how do I estimate this quantity from data?" And it gave me an answer, but the answer, it wasn't wrong, it was wrong but in a very subtle way was wrong such that it would've been

right had it specified certain very strong assumptions, but it didn't mention those assumptions. And so it was like, wow, this would've gotten you in trouble had you actually tried to apply this. And so we still have the same types of cautions that we have with language models when it comes to using it in a reason causally.

Jon Krohn: 00:54:01 I see. So overall, the interesting twist here is that when we're talking about generative AI and causal AI together, what you're describing at least there is where you're using the generative model to use its understanding of the world, which typical, if we're using Stan, we're using BUGS, those models, there's no built-in kind of understanding of the meaning of words or a feature name. And so when we're using these other kinds of approaches in PyTorch with du AI, with Pyro, with Stan, with BUGS, with all of those kinds of libraries, we have some table of numbers typically that we are specifying some assumptions or we might be able to use a graph to describe relationships between those variables. But ultimately we're working with just numbers.

00:54:56 Whereas, in the kind of evaluations that you were doing, you were seeing, okay, how does a generative model like GPT 4.o, how does it use its understanding of the world to draw some causal inference? So that's kind of more like the kind of intuition that you were describing where, okay, you walk into a room, your dog is on the floor with peanut butter on his face, and the peanut butter jar has been knocked off the counter. What happened?

Robert Ness: 00:55:29 So an example of the dog thing, if we were going to kind of throw a technical word that we might think of that as say root cause analysis, right? And there are algorithms for doing root cause analysis. Let's say you have a bunch of data about events that happened on a network and there was a data breach and you're trying to do some root cause analysis.

00:55:48 Then there's you, the guy who's staring at all these logs trying to figure out what's going on, or you can take those logs and paste them into an LLM and say, "Tell me what happened." Or you can take those logs and paste them into an LLM and say, "Take this data, isolate the key variables and apply this root cause analysis inference algorithm to it, or at least give me the code to implement it," right? And so in the same way you could say like, "Well, here's the problem that I'm thinking about. What's the right DAG to use here? Okay, giving the right DAG. How would I write this up in du AI for example?" In the same way you can say, "Give me this problem. How would I write this up in Stan?" Because it seen a lot of Stan code. So that's one aspect.

00:56:32 But I think to me from a research standpoint, the more interesting aspect is to say, well, how can I use causality to make even cooler generative AI, right? So there's an example of that in a book, kind of a basic example where I'm saying, let's imagine that we can kind of have a separate, say, generative model for each node in the DAG that's conditional on its parents in the DAG and then connect this all together and so you can still get, by implementing it as a graph that reflects causality, you still get all the benefits of the theory, but you can also generate like you would from a generative model.

00:57:13 But there are other really interesting ideas about where you could take this, say for example with multimodal models. So say if we have a model that's incorporating natural language and maybe also say video for example, there's some causal reasoning behind how you and I talk, right? There is some causal signal that's being extracted from natural language. And insofar as a video is a time series and causes precede effects in time, there's some causal signal hiding in that video data as well. And when we combine these into a multimodal model, what kinds of new interesting causal things is that model learning? And

can we extract it, can we constrain it so that we can get certain guarantees?

00:58:15 Now, one of the things that I'm working on is looking at the space of generative AI for video games and saying, "To what extent can we get this generative AI to understand the underlying game mechanics or the underlying game physics?" So rather than just generate really lifelike or really generate images of Minecraft that look a lot like real Minecraft, can we train this model in a way that it's going to learn something about the underlying mechanics of how Minecraft works? So maybe then it could generate things that honor those Minecraft mechanics, but still create environments that have never seen in the training data, or in a way that makes this model, this game kind of composable with other models. Let's say if I train a model on game A and I train a model on game B and I make it such that their mechanics are explicit in certain ways that they learn the right mechanics, can I then combine them to create a new game? So to me, those are kind of more interesting applications, but they're still kind of fresh.

Jon Krohn: 00:59:36 Tons of fascinating possibilities with generative AI in this causal AI space and so many other AI spaces. It's really fascinating to hear some of those examples. Let's jump now into some of the audience questions, we've got one here from Dr. Doug McLean, he's got a PhD in applied math and works as the lead data scientist at a food company in the UK called Greggs, which is famous for its delicious pastries, though they have lots of other foods as well. But anyway, so Doug, technical guy, has got a good technical question here. He says, "Could you make sure Robert comments on Judea Pearl's ladder of causation?" So rung one is association, rung two is intervention doing and rung three is counterfactual. And so there's some terms there that you mentioned counterfactual in some of

your other responses, but you might need to dig into the definition of that a bit for our audience.

01:00:31 And so the reason why he's asking you to comment on this ladder of causation, again, rung one is association, which I think is kind of like correlation. Rung two is intervention, so the kind of being able to make some assumptions about causal effect. And then rung three is counterfactual. He says, "To be blunt, I'm really stumped when it comes to causal modeling, it seems you need to know what to expect first before running any analysis." So I guess he's basically saying here that this ladder of causation or the way that you set up causal models, often to set up the model effectively, you already have to have some understanding of what's going on, and he finds that confusing.

Robert Ness: 01:01:14 And so to answer that last point directly is that I'd say that, I think oftentimes the way that people are trained to think about data, particularly, again, in industry where we unfortunately often have less control over how the data is collected, so causality is asking you to think more about the data generating process than the data. And so when you don't actually have control over that process, it can get a little bit frustrating because it's kind out of your hands, right? So typically we get some dataset we're thinking about like, "Okay, well, what are the transforms I should apply to this? Maybe I should discretize this, maybe I should apply a long transform to that." So we're thinking directly in manipulating the data to make it more amenable to the models that we want to use. And causality is saying, data aside, you need to tell me what your assumptions are about the underlying data generating process.

01:02:11 So now going into the causal hierarchy, I go into it very extensively in the book, I'd say probably more so than any other book. Again, at level one of association, we're just

kind of in kind plain old vanilla statistics and correlation land. And then level two intervention, this is what we're talking about say with when we're trying to emulate an experiment, when we're asking what if questions, like what if I were to take this vaccine, would it prevent me from getting sick?

01:02:49 And then level three is counterfactuals, and here we're asking questions where we're imagining what might have been different. So say for example I didn't get vaccinated and I got sick, would I have gotten sick had I been vaccinated? So this is actually the same as kind of level two, which is to say, what if I had taken the vaccine? What if I take the vaccine, would I get sick? But now we're going to condition on two pieces of additional information, the fact that I didn't take the vaccine and the fact that I did get sick having not taken the vaccine. So in level two I'm saying what if I take the vaccine? Will I get sick? And level three, I'm saying, what if I take the vaccine, or putting aside tense, putting aside present tense, past tense, say, A, take the vaccine, B, get sick, yes or no? And so that's the same as the first case. But now we're going to condition on two extra things, the facts that given not A, that not B happens, right?

01:04:06 And the reason we call it counterfactual is because either the hypothetical condition, taking the vaccine, or the hypothetical outcome, getting sick or not is in conflict with some data that we've actually observed, which just makes it a bit more challenging to model. But at the same time, we're still just conditioning on some extra evidence. And so in terms of what is required to make those types of assumptions, to answer those types of questions requires very generally speaking some additional causal assumptions. Some of those assumptions can be specified entirely in the form of a DAC, and then some of them can't often, more frankly, the more interesting ones can't

rely on DAC-based assumptions alone, we need to make additional assumptions about mechanism.

01:05:06 But if him being a mathematician who understand things, for example, so not only does A cause B, which we represent with a graph, but maybe you make an additional assumption that B is monotonic in A, that for a change in A, there's a corresponding change in B, and that change is constant or that change is going one direction no matter where you're at with A. So it is just asking for more assumptions. It's not asking you to know the answer, it's just asking you to say for certain questions you can get away with lesser assumptions and for certain questions you need to make a few more. And so we can think about our assumptions as falling on different levels of your hierarchy, and so sometimes for some interesting questions, you need to make level three counterfactual assumptions.

Jon Krohn: 01:06:04 Nice, okay. That is super helpful and hopefully Doug McLean finds that answer helpful as well, as many of our listeners. We have time for one more question here. So this is from Adriana Salcedo, she is a flight attendant in Bavaria in Germany, but she is training to become a data scientist or an AI engineer. So she's been taking lots of courses online for over a year now, she's a regular listener and regular commenter. And yeah, the last time I talked about her on air I said maybe we will get her on air at some point because it's such an interesting journey. And so yeah, we haven't scheduled that yet, but maybe someday Adriana. So Adriana had three questions, I think you've actually answered two of them. The first was around what types of problems does causal AI give you a clear advantage of over non-causal machine learning approaches? And we've talked about that a lot in this episode already with examples of when you need to understand if one variable is just causing another not correlated.

01:07:05 And then she had questions around whether you need domain knowledge in a particular area to apply causal AI. I'm guessing that for the most part the answer is yes, but she has a cool follow-up, which is if it is essential, then could LLMs like GPT4.0 Help us automate that part of the process or understand some of the domain knowledge that is required? And so that ties into, you answered that earlier. You said it can help you suggest what the directed acyclic graph might look like in a particular domain or help you put the code together in Stan or in PyTorch. So I think that that is kind of already answered. The one question that we haven't really talked about and I am really fascinated by is what does a typical causal AI workflow look like in practice? So when you set out to tackle some causal AI problem, how do you go about it? What's your workflow like?

Robert Ness: 01:08:06 There's theoretically different workflows, but I'll talk about the main workflow that most people get exposed to when they first get exposed to the space, which is to say you kind of start with a question that you want to answer, right? Does side quest engagement have an effect on in-game purchases, then you're going to specify your causal assumptions in some shape or form. In my book, I focus on doing so in the form of a graph or potentially a structural causal model, which is a graph plus extra assumptions about how variables relate to one another. But let's just stick to a graph for now. So you say like, "Okay, well I think that guild membership causes side quest engagement. I think that side quest engagement is a cause of in-game purchases. I also think guild membership is a cause of in-game purchases as well."

01:09:03 The next thing I want to do is make sure that I'm going to pick my library, let's say du AI, and I'm going to specify that DAG in code, say as a network X object and then pass it to some constructor in a model class in du AI, and it's going to give me a model object in Python.

01:09:23 The next thing that I'm going to do is say, "Okay, well I have this question and I'm, I'm going to now look at my data," so I have some data on certain variables in the system and then I'm going to do something that it's called identification. And so what's going to happen is say, identification means that given these assumptions here in the form of the DAG, given this data that covers certain variables that are in the DAG, and DAG has maybe some unobserved variables as some of them might be confounders, common causes between the cause and the effect of interest, can I answer this question? Let's suppose the answer is no then we say, okay, well can I get better data? Can I, for example, do something like not even better data, can I observe additional variables that would help me answer this question? Let's just say it like that. And so that could be, for example, getting an instrument, something that is a cause of side quest engagements but is an indirect cause of in-game purchases and side quest engagement mediates that relationship completely.

01:10:40 Maybe then I could say then in that case I can kind of do something called an instrumental variable analysis or maybe there's some kind of intermediate variable between side quest engagement and outcomes, like what we call a mediator and we could do something called front door analysis. But you might say like, "Okay, there's something else I can add to this data, I can collect to this data to give me that could help me get to answer?" Great. Now you have that thing, now you can answer the question.

01:11:04 The next step is given that you can answer the question, you want to pick some kind of statistical approach to do the estimation of the causal relationship. And those are going to have the same old statistical trade-offs we've seen elsewhere. So maybe you're using a linear regression type of model and maybe you're kind of leaning really

heavily on say the linearity assumptions or constant error or assumption, or maybe you do something like propensity scores, maybe you use something called double machine learning and one of them has really fat confidence intervals with the other one is a bit unstable, all kinds of stuff that you're thinking about. But these are all good old-fashioned stats questions, right? You've already done all the causal heavy lifting.

01:11:51 So you've done that and then maybe you do some sensitivity analysis at the end and du AI, we call it refutation, just to see how sensitive your results, your conclusions are to the assumptions in the that you're making in every step. So maybe you miss some variables in your DAG, maybe your data is a bit small or maybe you're leaning too hard on linearity this or that, and so you can kind of test all of those things to see how robust your results are to violations of those assumptions.

Jon Krohn: 01:12:21 Excellent. That was super helpful. I wish I had thought to ask that question earlier in the episode because that provides so much context around what causal AI is and how you do it. Super helpful. Yeah, great question, Adriana. So thank you so much for taking all this time with us, Robert, today. You've been very generous with your time, we've gone over our scheduled recording slots, so thank you for that. Very quickly before I let you go, something that I forgot to prepare you for is that at the end of every episode I ask our guests for a book recommendation. So yeah, this is something other than your own book, do you have anything for us?

Robert Ness: 01:13:00 Do you have a specific kind of domain in mind?

Jon Krohn: 01:13:03 Not necessarily. I mean, you have actually already mentioned some books, there was a Judea Pearl book that you mentioned earlier in the episode on causality, that could be a great one. But we also just sometimes

people give us a novel that they've enjoyed recently or whatever.

Robert Ness: 01:13:27 So off the top of my head, okay, so recently I've been reading, so yeah, if you're kind of new to causality and you just kind of want a light read, I think The Book of Why is a good call by Judea Pearl.

Jon Krohn: 01:13:38 And so it's pronounced Judea?

Robert Ness: 01:13:41 I think it's pronounced Judea?

Jon Krohn: 01:13:43 Judea? Okay. Oh man, I've been butchering that for years, for a decade. Okay. But yeah, The Book of Why.

Robert Ness: 01:13:52 When I talk to him, I just call him Dr. Pearl. He's never corrected.

Jon Krohn: 01:13:55 That's a good way to go. That's what I'll do from now on too.

Robert Ness: 01:14:00 But I think so recently I've been reading, I think it's called Surely You Jest, Dr. Feynman, or Mr. Feynman, the book by Richard Feynman.

Jon Krohn: 01:14:07 It's Surely You're Joking. Mr. Feynman.

Robert Ness: 01:14:09 Surely You're Joking. Mr. Feynman. Is it Mr.?

Jon Krohn: 01:14:12 Yeah. Even though he had a PhD, the book is Mr. Surely You're Joking. Mr. Feynman, yeah

Robert Ness: 01:14:17 Yeah, I mean, to me his ability to kind of boil complex concepts down into really clear explanations is something that I aspire to. And so I've been trying to reread that book and, I mean, that book is not so much about explanations, it was kind of more about his way of thinking, but I say that would be my recommendation for folks. Light read, it is not a technical book.

Jon Krohn: 01:14:49 Yeah, he's a lot of fun. I've been watching some lectures by Richard Feynman recently and it's just great. Fantastic. Thank you so much, Robert, for taking all this time with us today. For people who want to follow you after today's episode or connect with you, what's the best place to do that?

Robert Ness: 01:15:05 Good question. I occasionally lurk on LinkedIn, but I actually have been dialing back at social media recently. It's one of those times where social media can get a bit much. But LinkedIn is good.

Jon Krohn: 01:15:23 Yeah, and it's hard to be doing things like writing books and writing papers if you are too absorbed by social media. It's a tricky balance. So yeah, thanks again Robert and hopefully catch you again in the future. Thank you for an illuminating episode.

Robert Ness: 01:15:37 Thanks for having me.

Jon Krohn: 01:15:43 Great. Thanks to Dr. Robert Osazuwa Ness for coming on the SuperDataScience podcast and teaching us so much. In today's episode, he covered while how well humans and animals intuitively understand causality, AI systems developed using correlation-based learning because traditional causal methods were rooted in classical statistics and didn't scale well to high dimensional data. He talked about how libraries like Pyro, du AI, and PyMC now enable causal AI by handling statistical inference automatically letting practitioners focus on specifying causal assumptions rather than implementation details. We talked about Pearl's causal hierarchy. Level one is correlation or association, level two is intervention, asking what if questions, and level three involves counterfactuals, asking what would've happened given the observed evidence. He also talked about how LLMs can propose causal graphs generate analysis code and synthesize causal knowledge from training data, though

they still hallucinate and require careful validation. And finally, he provided us with a typical causal problem workflow where he starts with a causal question, specifies assumptions, often as a directed acyclic graph, he checks if your data can answer the question, chooses a statistical method for estimation, and then performs sensitivity analysis to test robustness.

01:17:01 Yeah, so 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 Robert's social media profiles, as well as my own at superdataescience.com/909. Thanks of course to everyone on the SuperDataScience Podcast team, our Podcast Manager, Sonja Brajovic, Media Editor, Mario Pombo, our Partnerships Team, which is Nathan Daly and Natalie Ziajski, our Researcher, Serg Masís, Writer Dr. Zara Karschay, and yes, our great Founder, Mr. Kirill Eremenko, thanks to all of them for producing another deep episode for us today 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 sponsors' links, which are in the show notes. And if you'd ever like to sponsor the podcast yourself, you can see how to do that at jonkrohn.com/podcast. You've got that link in the show notes, of course.

01:17:54 But yeah, lots of other ways to support us, we really appreciate it. You can share this episode with someone who might enjoy it, review the episode wherever you listen to it or watch it, subscribe if you're not a subscriber. 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 love for years and years to come. Until next time, keep on rocking it out there, and I'm looking forward to enjoying another round of the SuperDataScience Podcast with you very soon.