Pioneers and Pathfinders: Charlotte Alexander

(This transcript was generated through AI technology.)

Zeynep Ersin: Welcome to Pioneers and Pathfinders, Seyfarth's podcast about the people and ideas reshaping the legal industry.

Kevin Young: I'm Kevin Young, an employment litigator and counselor, and I've spent my career putting new tools and ideas to work in client service and helping colleagues and clients to do the same.

Zeynep Ersin: And I'm Zeynep Ersin. I lead innovation and design thinking at Seyfarth, focusing on reimagining how legal services are built, delivered, and experienced.

Kevin Young: Each episode, we sit down with someone on the front lines of legal innovation to unpack what they're doing and what it means for where our industry is headed.

Zeynep Ersin: We'll keep it fresh, we'll keep it fun, and we'll keep it to 30 minutes. Let's jump in.

Kevin Young: Today's guest is our good friend, Charlotte Alexander. Charlotte is a Harvard Law-trained scholar whose research has been published in some of the most prestigious journals in the world, including *Science*, the *NYU Law Review*, and the *Texas Law Review*.

Zeynep Ersin: Charlotte leads the Law, Data, and Design Lab at Georgia Tech, where she and her team use AI and machine learning to process massive amounts of court data, which leads to surfacing of patterns and disparities that have long been buried in millions of pages of legal text.

Kevin Young: Charlotte's work has attracted funding from the National Science Foundation, the U.S. Department of Labor, and Google. She serves on the AI Committee for the Georgia Judiciary. And in 2023, she took her research global as a Fulbright U.S. Scholar studying court innovation in the Dominican Republic.

Zeynep Ersin: Simply put, Charlotte has been doing this work long before AI became a buzzword, and she brings a perspective that is both deeply technical and deeply human. So in today's conversation, we'll dig into the court data problem most people don't know exists, what AI can and can't do for the justice system, and the ethics of deploying these tools at scale. So please join us as we welcome Charlotte.

You know, just to kick things off, Charlotte, I'd love to hear a little bit about you, your background, what you've been up to in your career journey, and we can go from there.

Charlotte Alexander: Yeah, so I'm a professor of law and ethics at Georgia Tech. Many of your listeners may know Georgia Tech does not have a law school. And so what I do is I teach business students, both at the graduate and undergraduate level, and then a lot of my work also involves collaborations with our computer scientists and along with students from public policy and business. So that's what I do now. Before this, I was at Georgia State University jointly at the business school and the law school for over a decade. And then before that, for a little bit of time, I was a practicing employment lawyer. So in the same space as Kevin.

Kevin Young: Thanks so much for being on with us. I had to dig back in my emails to remember when we met and it was in 2017. And I think I emailed you cold because you had put

out from your post at Georgia State's Law School at the time—I think it was a study on employment misclassification litigation and issues. And it just caught my interest. I was dealing with a lot of misclassification litigation at the time. I've told people before—I think I told Zeynep this—that you were one of those people who was focused on AI and machine learning and big data before it was cool, before it became really in vogue. And I'm curious if you can share with us and with our listeners how you came to find yourself at that intersection in the first place?

Charlotte Alexander: Yeah, sure. So in the research part of my job, I've been motivated for over a decade now by trying to answer basic questions about how litigation works in this country. And when I first shifted over from practicing into academia, I thought, well, gosh, I have questions about what the most common types of claims are and what the outcomes are and sort of basic things that I thought we should be able to look up about how the courts operate. And surprise, surprise, it turns out that lawyers and judges like to write a lot of words, but don't like to produce a lot of structured data. So there's not a lot that you can really look up.

And so this was really about the time that we first got in touch, Kevin, about a decade ago, started hearing whispers about large language models—so the distant ancestors of today's GPTs and Geminis, et cetera. And I started working with computer scientists on the faculty to try to extract key information from court documents to create the data to answer the questions that I had.

Kevin Young: In those early days, what surprised you about what you were able to get your hands on and then what you found as you started looking at it?

Charlotte Alexander: So just on the method side, I think I initially naively had this idea that computers were magic, you know, and that you could just say, "Look through all these judges' misclassification opinions and tell me the percentage in which the worker was deemed to be an employee and the percentage in which the person was in a contract." So, of course, right, that is not what happened. And so I think one revelation or one surprise on the methods side was computer science and data science are called science, but they're part science and part art. And there's so much judgment that goes into using these tools. And I think that holds true today as well.

And I think that on the substantive side, what has surprised me over time is we have a lot of kind of lore around litigation and what happens in litigation. And I did a study on settlement to try to understand what percentage of cases settle. And you see a lot of percentages thrown around, but it turns out when you actually dig in and try to measure it, it's a much more difficult question to answer. Anyway, lots of more complexity, I think, was my overall surprise on both the substantive side and the method side.

Zeynep Ersin: And in terms of the method, what was the reaction or the perception with folks who you were trying to work with to navigate this? Were people excited? Were they reserved? What was the general sentiment?

Charlotte Alexander: Well, I think it depends who I was talking to. So I found myself a lot of times kind of pitching law as a worthy use case to developers, computer science types, data science types. And law is so seen, I think, in other fields as such a foreign complex thing that I think it sometimes takes a lot of work to say, "Hey, no, this is this huge system, there are thousands of lawyers in this country, the courts affect everybody. Let's bring our collective skills to bear on this." So I think there's that.

But then also as I started kind of getting out there and talking to lawyers and judges about what I was doing, I think in the early days, I think there was a lot of excitement on the part of lawyers and law firms and legal departments to say, "Gosh, we can really leverage these tools and these advancements to increase efficiency." But a lot of sort of questions about how to implement that and a lot of worries about confidentiality and data provenance and how to manage all of that. I think judges have another set—similar set of reactions, which is there's a lot that could be done here, but also kind of worries about large-scale impacts on the legal system and what all this means for justice or the lack thereof.

Kevin Young: Charlotte, on that note, in terms of the lack of transparency, is that, in your view, an issue that has persisted at the same level, even as machine learning has grown more advanced and the ability to really analyze big data sets has grown more advanced?

Charlotte Alexander: Yeah, I think that the actual nuts and bolts of taking, say, thousands of briefs or judges' opinions and analyzing them, that has become dramatically easier. I think that the problem that remains is access to a large enough corpus of documents to study. And so we have on the federal side, we have the PACER system that gives access to filings in federal court, but charges 10 cents a page. And it's pretty difficult to easily assemble a large set of records. On the state side, of course, it's completely fragmented and it depends on the jurisdiction.

So I think that the tools that we have now to make sense of massive amounts of legal text, as I said, are super sophisticated compared to when I started in this area. But the access to the underlying raw material problem—the ball has not moved very much there. Now, of course, firms that sit on their own big troves of briefs, et cetera, there's not an access problem there necessarily. So there's a lot that can be done internally, but I'm sort of more thinking about public-facing court record availability.

Zeynep Ersin: Can you share an example or two of any current projects you're working on—essentially what the goal is and what you and your students are doing?

Charlotte Alexander: Yeah, so I've been saying this to as many people as possible as a way to hold myself accountable for this. So I'm starting on a project which I think is going to be a book about court delay. And the reason this ties into everything we've been talking about is what I really want to do is make the pitch that better data can help courts themselves understand where the choke points are, where the bottlenecks are. And if you have good enough litigation data that tracks every event that happens and the associated date, you can actually start to build a simulated court. And this is using AI not just to create the data on the front end, but then using AI to simulate case flows and then do things like simulate an additional judge. And if you add another judge, what's the downstream effect going to be on speed and outcomes?

And so I'm super excited about this. It's inspired by a bunch of work that's done in hospitals to track patient flows. And I'm really excited about sort of taking all of those learnings from that other field and importing them over into law and seeing what we can do.

Zeynep Ersin: Charlotte, you're touching on something that's so critical as it relates to AI, generative AI, is the quality of the data and what it's actually being taught from. And this goes back certainly to knowledge management and quality control and how data is captured and stored and leveraged. And so in a legal environment, whether it's a law firm or an in-house team, do you have any suggestions or guidance or advice you would give for how to kind of, as

best practices in this moment that we're in, kind of ensure the quality of data that they're using for their own purposes in tandem with AI?

Charlotte Alexander: Yeah, sure. I mean, it's definitely a difficult problem because the output of your tool is only going to be as good as the inputs on the front end, which we all understand by this point, I think. Yeah, so I think that there needs to be—and this is what I talk to my students about all the time—is you have to understand your data. And so if you're going to train a model, you have to understand what data you're using, what you're leaving out, what the process was that generated the data in the first place, what the incentives were for the people who generated the data. Because if they were incentivized, you know, to choose a certain option in a dropdown menu and punished for choosing the other one, then you're going to see a skew potentially that does not reflect reality, but rather reflects the incentives of the data generators.

So I think that there's almost a sort of separate phase that needs to happen before implementing any model that's trained on data, which is just this deep dive, interrogating the data and sort of adopting a really critical perspective on what's there and what's missing.

Kevin Young: Charlotte, you were talking about your current work. I will hold you to the book. I'm going to go ahead and pre-order it whenever we can. But looking at choke points and courts, which we all know are bogged down both at federal and state levels, have more cases than most courts were probably meant to handle, and things can take a while, including things that probably shouldn't under normal circumstances—I'm curious if you've had occasion to see whether federal courts, state courts, or even individual judges are generally receptive to kind of holding the mirror up against themselves in this moment, given the availability of better, more cutting-edge data sets and the ability to really explore what they mean or what they show us.

Charlotte Alexander: So I sit on the AI Committee for the Georgia Judiciary here. And so I can speak firsthand to say that—and I think those judges are emblematic of judges generally—which is there's a lot of interest in leveraging AI tools to figure out how judges can do their jobs more efficiently, but also how the system can work better. I think that some of the initial hesitations that were there when I first started this work a decade ago have started to diminish as judges have come to realize just really what the enormous potential of these tools are.

I do think though there have been some reports that have come out by people who actually aren't really attuned to the intricacies of litigation. And so there was a report a couple of years ago by some folks who had done a really kind of rudimentary analysis of federal sentencing practices and just really kind of got the data wrong. And this is back to understanding the data. And so I think there's a concern not only among judges, but among academics who focus on this area, that as these tools get better, that people sort of just jump in and say, "Oh, let's analyze all the data" without really understanding what the data is showing.

And so I will say that judges in my experience are super excited about leveraging what they can do with their data, but have this worry that, you know, there's going to be sort of half-baked stuff out there, analyses, et cetera, that may end up kind of setting back the whole enterprise. And so, yeah, I worry about that as well, because I think there's a lot of potential, but I don't want bad work to kind of poison the well.

Zeynep Ersin: As a follow-on to that, if and when we continue and your work is incredible and you are able to solve the data problem, you know, of course systems become searchable, understandable, and analyzable really at scale—how will that look? You know, what changes?

Charlotte Alexander: I think there's a variety of both positive and negative potential outcomes here. So of course I would love all that world in which there's tons of transparency and I have a super long list of questions that I want to answer. So I think that in the good column, there's more access for researchers, for court administrators, for lawyers to understand what the patterns are and be able to increase efficiency and get better at our jobs. And I think courts themselves can learn—again, learn their own patterns and increase their own efficiency and fairness.

I think there are good potential implications for access to justice, right? So for people who can't afford a lawyer and so who now can potentially use some of the data to better inform what they do as pro se litigants, or legal aid, legal services organizations use data as sort of a force multiplier and be able to become more efficient themselves.

I think on the flip side though, in that same category of access to justice and unrepresented litigants—like that's where I'm a little worried about negative impacts, which is, I think already we've seen examples of litigants who just use ChatGPT to write briefs and are not well served by what's there. So we all have heard about hallucinated cites, but I think the dangers of not having a lawyer's guidance are amplified when we're talking about hallucinated content and just using these large language models to generate stuff to file in court. And so I think that on the one hand, transparency is great and amazing, but transparency that then flows into tools that are themselves not vetted is problematic potentially for access to justice.

Kevin Young: Yeah, it reminds me of how, Charlotte, our relationship first started, because I was very curious at the time about how we could, at our firm and our litigation practice, use big data to help make litigation decisions. Is it worth the client's time and dime to file X motion or make this argument? Beyond just like the normal, "Well, can we find a precedential opinion? Can we find something that's a good guidepost?" Could we use data to say, "Is this the right decision, the right investment of time or not?"

And at the time, there weren't a lot of great research pieces out there, like what we found when we saw your studies, including on the misclassification of employees and contractors paper that you mentioned. It seems like the sandbox has gotten fuller of folks kind of playing around with these ideas and really investing in this space. I'm curious if you've seen that too, and if so, kind of what's the net impact then and what do you think it'll continue to be as we move forward?

Charlotte Alexander: If a company or a startup or whoever it is pitching some litigation prediction tool, my first question is, "What did you train it on? Where'd you get the data? What did you do to clean and standardize the data?" Because I will tell you that a motion to dismiss—you think a motion to dismiss is a motion to dismiss, but turns out a motion to dismiss is described 94 different ways in each of the 94 different federal district courts times many more variations. So that's an exaggeration, but not too far from the truth.

So I think that, again, we've moved closer to litigation prediction being functional, but I think it depends on what you're trying to predict. Some things are a lot easier to find in the data than other things. And so I think we do have the ability today to take into account judge, claim type, stage in litigation, and do some prediction, but we're still not there in terms of being able to model the whole system and understand the different branching paths in litigation.

Kevin Young: What components would you say have you found to be easier and a little bit more reliable versus what are the stickier, more complicated data sets?

Charlotte Alexander: Sure. So I mean, I've done a bunch of work with docket sheets from federal district courts trying to experiment with using AI essentially to apply standardized labels to every docket entry on a docket sheet so we can understand at a sort of 10,000-foot level, "This case went from complaint to a motion to dismiss that was partially granted and then, et cetera, et cetera." And so I was exaggerating a little when I was using the motion to dismiss example, but dispositive motions generally tend to be pretty easy to find. Decisions by judges are pretty easy to identify.

Settlement is actually really difficult to identify because it's something that happens out of court. And then the parties choose a million different ways to report it on the docket sheet. And so that actually turns out to be quite difficult to figure out when a case is settled and how much of it has settled. Discovery, as you also know, happens out of court mostly. And so it can be hard to pinpoint exactly where on the docket sheet discovery has begun and when it's ended. If you're trying to model—trying to answer basic questions like "What percentage of cases settled before or after discovery?"—well, to do that, you have to be able to identify settlement and discovery, both of which are difficult to figure out just from the face of a docket sheet.

Kevin Young: Charlotte, to switch gears for a moment, you're at a place that we care a lot about, which is an educational setting. We have incoming lawyers every year and others entering the legal industry who are going to have firsthand exposure to a lot of the things you're talking about—AI, machine learning, big data—and they're entering a profession that no doubt seems to be changing by the week is what it feels like, I think, to a lot of us who are in it. And I'm curious, from what you've seen working on the front lines at Georgia Tech and then Georgia State's School of Law before that, what do you see from the next generation of folks who are coming into the industry? What can we expect from them?

Charlotte Alexander: Sure. So I think right now there is a lot of fear. Even though I don't teach law students directly anymore, I hire law students every semester as research assistants and so I'm talking to them. And I think there's a lot of fear about whether there's a viable career path for some of them. They're worried that what used to be the bread and butter of your associate's job is now being taken over by AI tools. So I think they, like the rest of us, are apprehensive and just sort of uncertain about how all this is going to shake out in terms of the labor market impacts.

I think though that this new wave of students coming out into the workplace are not just digital natives, but AI natives. And that's a huge asset for employers. The ideal combination is someone who's an AI native, but also is a critical thinker and has all of that set of skills that some people call soft skills. But I think they're even more important these days than hard skills—which is critical thinking, teamwork, ability to synthesize information across a variety of different sources and analyze. So I think that as a professor, I'm really attuned these days to building up that set of skills in my students while also encouraging their interest in and facility with AI tools. As the students and the graduates who can combine the two, I think are awesome and are great and y'all should snap them up as soon as you find them.

Kevin Young: We're trying. Yeah, I very much agree with that, Charlotte, in terms of the importance of those skills that you touched on, and as part of that, the problem-solving capability, because I think as we're well aware, the tools themselves are going to continue to evolve and change, same with the UIs, but it's much more important to be thinking through what are you doing with the tools and the respective outputs, how are you incorporating them into what you're doing as an attorney and how you're serving your clients, whether it's part of an in-house team or for a law firm. And that is just mission critical.

It's been interesting to see the perspectives from law students who've been coming in for us as summer associates and how those conversations have evolved over the course of the last few summers in particular. There's still certainly some fear. And I think that that kind of permeates across a number of law schools, but it has started to shift a little bit where I think that there's some more willingness and eagerness to embrace the technology and use it as a tool and not as a replacement. Have you seen a little bit of that shift at all over the last few years or would you say it's been pretty consistent?

Charlotte Alexander: No, I think I agree with you. And I think maybe I also hear a lot from undergrads and MBAs who are facing a really tough job market. So maybe my fear feeling is also influenced by those students as well. But I do think in law in general, and sort of talking with other law professors as well, that we're getting toward a level of sort of comfort and understanding that we can exploit these tools in a good way and maybe actually make the life of junior associates a little bit better even.

And so, I don't know if you all know the Gartner Hype Cycle, which I'm making a hand gesture, which is like a roller coaster where there are different phases of the adoption of technology and then you eventually reach the plateau of productivity. And so, I recently looked at that Gartner Hype Cycle for AI and we're barely on the plateau of productivity, at least according to Gartner, for various types of AI tools. So this is all to say that I think that we're gradually creeping towards some level of acceptance and less fear. But again, it's just things move so fast, as you have said, that I don't think we're there yet. And I think there's still this combination of excitement and apprehension for sure.

Kevin Young: Zeynep and I have talked about this quite a bit. I think that the ideal associate, ideal attorney, I think is exactly what you describe. It's someone who's got the substantive expertise, but also has the ability to harness these tools in a really powerful way. One thing we've talked about before is if you take a 20-year IP lawyer, they are as capable as anyone at using whatever AI, generative AI tool their firm puts in their hands—they can be really powerful. But if you have that same tool, write them a brief on employment law or real estate, they probably don't know what's why. They don't know what's right, what's wrong.

And I think we have to think about the incoming generation that way. We want them to be as productive and innovative as possible with the tools we afford them, but we have to make sure also that the substantive expertise comes along so they can become that 20-year IP lawyer or whatever the case is going to be. And it feels like a tough needle to thread, but it feels like it would be so worth it if you can stick the landing.

Charlotte Alexander: And I will say from my perspective as a professor, the students still need to develop their own independent human expertise, their subject matter knowledge, so they can evaluate the output of these tools. And what I really worry about in terms of what's happening in universities and law schools is that if students are using AI in their learning process, they're not developing that external expertise, and therefore they don't know if that employment or real estate brief is right or wrong. How do we make sure that they're developing not only those critical thinking and problem solving skills, but also actually know something substantive about the law? So it's tough and I don't think we in academia have quite figured it out yet, but it is something that I know I think about a lot.

Kevin Young: Same here. It's tough. It is a problem worth solving because the future state, I feel like if we can get it right, both at the university level and future employers, be it a law firm or

anyone else—to me, it is a lawyer of the future who is just so well-enabled and even more powerful than the best attorneys of today.

Charlotte Alexander: Yeah, I agree.

Zeynep Ersin: So Charlotte, most people remember or have an AI Eureka moment. And I know you've been dabbling and working in this space for some time, but do you have one that stands out for you?

Charlotte Alexander: Well, I think my initial AI Eureka moment goes back to when I first started experimenting with these tools. And I was working—again, it was an employee misclassification project—and essentially judges making those decisions have to use a multifactor test, as Kevin well knows. And what I was interested in is, yes, they recite these multiple factors, but maybe in real life, only a couple of the factors really matter. So I wanted to do an analysis to try to understand how much weight judges were giving to the different factors in the multifactor test.

And so I went through about a year of hiring 20 law students each semester, having them read opinions and hand code which factors judges were talking about. And then in chatting with a fellow professor who's from the computer science department about this project, I realized that maybe computers could do that. And so that was the eureka moment—was, again, those law students were out of a job. So back to the labor market impact question, but that was what launched me, you know, sort of that was one of the things that launched me onto this path of, you know, what sort of efficiencies, what sort of gains, what could we do here? What can I replace this super laborious human process with an AI-enabled process?

Zeynep Ersin: For me personally too, it's been—after having a couple of those, it's almost like a mindset shift of, like, when I'm going through my day-to-day, of thinking about, "Wait, can AI help me do this differently or better this time around?" where I'm also just stopping and taking a moment to ask that question directly. But yeah, it started from a couple of big ones that now has permeated into how I tackle my day.

Kevin Young: So, 10 years from now, 2036, what's something that you hope lawyers or judges, courts will seem outdated for them by that time? What's going to just die on the vine?

Charlotte Alexander: I guess in terms of—I'll think about this in terms of just everybody's daily life, like, Zeynep, you were saying, how do we think about our daily lives and what we do? I think repetitive tasks where we find ourselves doing the same thing over and over again—those are going to be gone, right? We're going to be using AI tools just to do those sort of rote work. And that could be drafting—you know, you can imagine just brief writing, if it hasn't already been revolutionized, will be revolutionized—to the extent that a lot of what we write is repetitive.

I think, you know, apart from the day-to-day, I think at a system level, something that will be badly outdated is the failure to capture and leverage all the data that's created in the way the legal system operates. And we've been talking a lot about text as a source of data, but there's just masses of, or what I call—not me, but data theorists call—organic data, which is just the byproduct of what we do in our daily lives, to capture as organically created data sets and really understand more about how the system works.

That could be things like understanding how long it takes people to physically navigate a courthouse, right? How long does it take folks who are incarcerated to be brought back and forth to the courthouse? So parts of how the courts operate that we could think of as data-

generating sites, and then being able to leverage that to make our justice system more just, more fair—without being too Big Brother. So that's the other piece of it, right?

Kevin Young: That's awesome. Thank you so much for taking time to be with us today. It's always enlightening and fun just to hear what you're working on and to get your perspective on everything that's going on right now, which is a lot.

Zeynep Ersin: Thank you, Charlotte. We appreciate it.

Charlotte Alexander: Well, thanks so much. This was really fun. I really enjoyed it.

Zeynep Ersin: Thanks for spending time with us today on Pioneers and Pathfinders. I'm Zeynep Ersin.

Kevin Young: And I'm Kevin Young. We'll see you next time.