LAWYERS ON FIRE

**Lawyers on Fire**: Hello I am here today in San Francisco for Lawyers on Fire. We are talking today with a very famous speaker of this conference Legal Tech West Coast and I think he is also quite somebody in the US, let me briefly introduce Dr Zev Eigen. He was honored as one of "40 under 40 rising legal stars" in the US. He is quite an interesting combination of multiple disciplines and an absolutely interesting speaker as I just learned. Thank you very much Zev for being with us today and giving us the chance to learn a bit from you.

**Zev Eigen**: It's my pleasure. I hope we can have a good conversation.

**Lawyers on Fire**: Cool. You are both a lawyer and a data scientist, how comes?

**Zev Eigen**: I started practicing law after law school in 1999, practicing labor and employment law. I went in-house with 20th Century Fox Film Corporation in Los Angeles working for them, doing their labor relations work. So I was their senior labor relations council and I thought I wanted to be an academic so I went back to get my Ph.D. from MIT. Following that I sort of fell into data science so I did MIT and after MIT I was a law professor for years and I was also an expert and consultant during work in the space in the private sector at the same time so all of my education and research and work as a consultant required data scientific expertise. So I was trained as a data scientist. As an undergrad I had a lot of training in statistics and math and economics and computer science. I sort of had the background and expertise and had the passion for doing it and it made sense to do both simultaneously and use law as an HR expertise and background as a lever to improve and maximize on the models we can write in those spaces.

**Lawyers on Fire**: You are performing predictive analysis. To those folks listening over in Germany and Europe, could you elaborate a bit on that? What exactly does this mean?

**Zev Eigen**: Yeah I mean in simplest form, we predict stuff for clients right so the idea is one of the things we are predicting, I would say there are two big buckets

of stuff and then those buckets get sub-divided. So we predict HR outcomes. So in the HR space, typically those things involve predicting "who is likely to leave the company in next 6 months" or "which employee is likely to be a good employee". You know there's a lot of work to be done in those two spaces alone but then there are other sources like policy optimization. A law changes – what is the optimal policy response given all the constraints of your business, the employees of your business, you know where should you put the water cooler, things like that. How much paid time off you should offer the employees etc. That's HR prediction modeling. And then the other big bucket is legal prediction modeling. In the legal prediction modeling that I would say is sub divisible into two categories:

-pre-litigation risk and

-post-litigation risk.

- Pre-litigation risk is the question of predicting where risk is going to materialize next. So if you are a big retailer – let's say you have many locations across the country – and you like to know are you going to have wage an hour class action problems or are you going to have sexual harassment problems or are you going to have race discrimination problems what kind of risk will it be, where will it be materialize, is it going to be on the east coast, the west coast, the north, the south and when is it going to be in the first quarter of the year, next quarter when is it likely to materialize, that's pre-litigation risks prediction modeling.
- Post-litigation risk prediction modeling, the last category, involves once risk has materialized, say an employer has sued let's say for wage an hour violation class action and the allegation is 20 million dollars worth of violations. How much should this case settle for, predicting how much it's worth, what's the likelihood that this case will be dismissed on a motion or, if it proceeds to trial, what's the likelihood of success on the merits, given

LAWYERS ON FIRE

all the constraints of the judge, opposing counsel etc should you propose settlement, strategically when should you propose settlement etc. So those are the sorts of categories big ones HR, legal and then within legal pre-litigation risks, post-litigation risks. That's the general very big way of describing the things we can predict.

**Lawyers on Fire**: One of your jobs is the global director of data analytics at big law firm, Here you are actually writing the algorithms yourself as I understood it. Could you explain a bit you know how do you come from all these inputs to a computer which then predicts an outcome?

**Zev Eigen**: Yeah and to be clear I am not actually a software engineer and I have no expertise, no experience developing software. Other people are good at that and part of working in this space is knowing what are you good at and then finding other people to do the other parts that you need but maybe not your expertise because It's too much to do everything. I don't write software. So the basic process is, first you have to figure out what are you predicting, then figure out what data you have available before you do anything else. Then figure out what additional data you're likely to need if you don't already have it, that takes a long time. Once you do that then the next step is the part at which I am pretty good at which is developing the model to go from point a to point b. Saying you care about predicting this outcome variable. I try to write the model to predict that outcome variable as optimally as possible. Now there are other things that happen after that, before that etc when other people do that. For instance if you want the output to be a part of software package, that's not really my expertise, getting enterprise level data into the model that's also not really my expertise. I am not a database manager or database technician. There are also data security issues that arise that we have people for sure to work on and that also not my area of expertise.

LAWYERS ON FIRE

**Lawyers on Fire**: So are you actually designing the model then?

**Zev Eigen**: I design the model and write the code. So there's definitely a lot of writing of code to optimize models. Identifying which models to use, which to apply and then evaluating the process and I have written code that is a model optimization tool frankly so it's a way of using cognitive computer machine learning to optimize a machine learning model.

**Lawyers on Fire**: Well just to go back a little bit, coming to consumer understanding. What are the benefits clients receive from your algorithms? Can you summarize those?

**Zev Eigen**: So the benefits are pretty substantial and I think you can't look at the benefits without the idea of how much it cost to get you there so if I said this is going to cost 50 million dollars to predict an outcome it doesn't make any sense to do it that's great but the case is only worth 100 thousand dollars. It does makes sense right. So it's never the case, I never had the experience the case where the price was too expensive for the benefit. Never it's never the case, If I could anything, I think would often undercharging for the value because often there is an ongoing benefit that extends beyond the case at hand if its litigation risks. So typically the way to model the benefit let's talk about HR for a second. The way to model the benefit is okay let's say you have on the average your sales company a 100 people who walk out the door every year. And those 100 people you are asked to make a book of business if they walk out of the door let's say for each 100 people there is 100 thousand dollars for each person well simple math to figure out how much you are losing for them all walking out of the door. Let's say some of them you are not going to lose anyway like let's say 10 of them and you are left with 90 people that maybe you can retain maybe you can't. If I can tell you ahead of time with 95 percent certainty that here are the 100 people who are likely to leave, you could then claw back or keep perhaps let's say you can keep 50 percent of those people. So that's a simple mats question figuring out how

LAWYERS ON FIRE

much money you are going to stand to make just on this one six month period and multiply that out for years and you will see as long as we are charging less than that its sort of a no brainer that it makes sense and every time that's the case. It's never been a case when the cost exceeds the value.

**Lawyers on Fire**: How do these services fit into a law firm? You know you actually you work in a law firm and it sounds to me more like something different from classical I would say legal services. How do you find your role there with what you are currently doing?

**Zev Eigen**: Yeah I think it's very difficult for most law firms to take on a startup like the one I am running for Littler because this is what is called intra-preneurship. Intrapreneurship is where an organization develops a branch that is not necessarily completely in parallel with their existing business model. Right so it's not like I am orthogonal to the business model but I am definitely a supplemental piece of the model but to be clear I am generating business that is not the same as the billable hours charged by other attorneys. So if I charge a client to run a diversity order say that's business that Littler didn't have before now my team and I do it and Littler gets the benefit of the work performed. It's difficult for a law firm in general to do that. Littler was easier because it is sort of in the law firm's DNA to innovate. We have other innovative processes and procedures like Little Case Smart, our joint venture with Neota Logic called Compliance HR and historically in 2000 we had a spinoff company called ELT so we have this in our DNA to think innovatively and be ahead of the curve so the management at Littler, it was easier to do something like this with our management than with other law firms but make no mistake it's definitely a challenge for lawyers or law firms in general and Littler is no exception to that in some respects to do something that's different and ahead of the curve.

**Lawyers on Fire**: Do you think that every big law firm should innovate and have people like you?

LAWYERS ON FIRE

**Zev Eigen**: I don't. As you know I have said this many times. You know I don't think it makes sense for all law firms to do this. The reason is most law firms have a business model and that business model derives from a very linear approach like a wiz old old tree. You know it's the billable hour and we are professional lawyers and we bill by the hours and we have expertise in subject matter of law and that's what we do and it's very difficult for wiz old old trees to grow a branch that eventually bears fruit that are too heavy and kind of squeeze the whole tree little bit. Even if that's the right thing to do it doesn't make sense for firms that try to reinvent themselves. They may have not the experience, the expertise. They may not be able to sell the product and services to clients and they may not be able to hire the people necessarily to make those changes. Just like any other industry where there is tremendous change and upheaval, it doesn't make sense for every organization in that industry to become the new model. Like taxis, it doesn't make sense for every taxi company to try to reinvent itself as a version of Uber.

**Lawyers on Fire**: You also started your own startup Cherry Tree. What are you doing there?

**Zev Eigen**: I am very excited about Cherry Tree Data Science. Cherry Tree Data Science is a very innovative application of neural models to empower employers to safely hire individuals who have criminal records. There are millions upon millions of Americans with criminal records. A third of Americans have some criminal record. In Germany the criminal justice system is very different. Many Americans have criminal records and unfortunately a disproportionate number of minorities, particularly black and Hispanic men have criminal records in United States. They are heretofore almost uniformly discarded from employment. So whether you stole a candy bar or shoplifted or had possessed drugs five years ago that might stop you from getting a job at a fast food place or a call center. You know you are answering phone calls because of that criminal conviction in

your background. Our startup uses data science to identify those people that you are safe to hire and enables employers to hire those people without any problem or risk of being sued for negligent retention. It's a very powerful thing and I am really excited about it.

**Lawyers on Fire**: This morning you did a presentation about machine learning and about artificial intelligence and the ethical implications of it. Could you summarize just you know: where is the danger? How far can it go and what do you think is happening currently in the legal tech or legal scene there?

**Zev Eigen**: Sure, the real risk in this space that makes it a bigger challenge than it would be otherwise for lawyers in this space I consider two big challenges. One is understanding what it is that's happening. If there is a new technology and a lawyer doesn't understand it, ethically, the lawyer has an obligation to understand that technology if it's being used by him or in the court in the context of litigation. So that's challenge number one – getting the lawyers to understand what is going on in data science and machine learning. The other challenge which I think may be harder – because I can help lawyers understand the basics, I can explain things to lawyers, lawyers can understand this, I mean that's the hard part but that's a challenge – the other challenge which I think is much trickier is the fact that neural models involve what is literally called a hidden layer which is very difficult as advertised for even a data scientist who is working on the model to tease out and identify the weightings of different features that are affecting the outcomes of the model. Right so the idea is how much is someone's gender playing a role in a model? Say you don't want gender to play a factor in a prediction model, how much is it playing a role in that model you can evaluate that for sure but it might update overtime, it might change over time and there might be proxies for gender that are very hard to tease out of that model given the huge features. Let's say you have 5000 parameters going into a model, maybe 10000 parameters that are being used in this machine learning tool. How do I

know that one or maybe more of them interact in a way that is a proxy for gender? Right, so exposed it is very easy to identify that because if it has an effect and it's harmful then you can say well this is biased. But the ethics involved – what you have to do to get that information or how you can get that information, maybe you can't just use it at all – are really not answered, we don't have good answers from courts or experts or anyone in the space and it presents significant challenges going forward. First learning what it is and then secondly even once you know what it is, ethically speaking it's very difficult, even if you wanted to do the right thing, to do the right thing in the space because that's the black box nature of it. I would say that as we go forward, the government in the US at least is certainly paying attention to these issues obviously the legal issues and some of the ethical issues as well that arise out of using data science in making any organizational decisions. The FTC (Federal Trade Commission) issued a report in January of this year advising employers of the risks associated with using data science typic methods like machine learning to render business decisions identifying several laws including but not limited to the Fair Credit Reporting Act, Title VII which is a big discrimination law in this country, the federal law that prohibits employers from using protected categories like race, religion and national origin as a basis for rendering employment decisions. There are other laws too but just to give you an example. So I think government is paying attention to this. It will become a bigger issue going forward as well as privacy issues that are going to be very big in the next 5 years

**Lawyers on Fire**: Hmm. You are also teaching at Yale?

**Zev Eigen**: No I am not teaching now (laughing).

**Lawyers on Fire**: You did. What would be your advice to young lawyers or law or business students today just coming on to the legal market. How should they prepare themselves for tomorrow's legal market?

LAWYERS ON FIRE

**Zev Eigen**: I think it makes sense to first figure out, I think the market is changing. So the market for legal service provision is definitely evolving. At least in the United States. I think in other countries as well but I won't pretend to know enough to be an expert in other areas outside the US. But in the US, you definitely see a shift. Here is a shift you see very clearly from the market and this will answer your question. There are whole law firms developing within companies. So the model used to be small in-house roles like may be 1 or 2 lawyers in-house for a company like Wal-Mart or Target and then they would rely on hiring lots of law firms. So when you are coming out of law school, you would go get trained at a law firm that would pay you a big salary and you work there for several years, learn how to practice your trade and then maybe some of you, some of those people would get hired in-house and they get a job in-house with a big company or a small company. Now the model is sort of shifted where we have many many lawyers and not enough jobs for them at law firms because of the way market has shrunk and the services have changed and a bulging or growing of the in-house staff. So now there are companies that have huge in-house staff. So the need for flextime attorneys which is what Littler uses on its Case Smart platform and other atraditional staffing models is very very clear. That model is very important and needs to be more robust. I think that's true in other markets as well not just in the US. So if you are a lawyer coming out, I think the way you and I had young lawyers many students ask me what do, I do given the market. I think you have to do two things. One is contemplate alternative skills that adds value for you and your clients like I am a labor employment lawyer so I speak to labor employment. I think understanding HR and understanding things that are important in the HR prediction space are a value add for lawyers. So if I were a young lawyer coming out I would focus on learning first of all about HR and then gain expertise in HR and decision making in HR and then secondarily learning what tools If I can learn one or two or maybe some programming some basic understanding of statistics, some additional corollary skills that would help me improve HR decision making. Right this is an example in my space. And I think

anyone can do this. There is so many resources available online that makes this a really available, accessible thing. The other thing lawyers can do who are coming out is contemplate entrepreneurial ventures. Right they don't have to, you don't have to take a traditional path of a policy job or a law firm job to "practice law". There are other ways you can practice law. I give you two examples that make a lot of sense to me. I know someone who came out of law school, couldn't get a job, traditional law practicing job and he started his own company and the company involves expungement of records for people as a minor they had some criminal record and his company helps expunge those records, process the expungement, doesn't always work but the idea is its low dollar value paying like 50 dollars I don't remember the price point and his company would help expunge the criminal records. Does that count as practicing law? I don't know. I mean it's not really arguing cases in the Supreme Court but it's a service, it's an important one it helps people get jobs. I am a huge proponent obviously from Cherry Tree Data Science for helping more people to get work so I love things like this. That's a clever entrepreneurial application of a legal service that he provided never practices, never officially practiced law in his daily life in a law school I that's a huge value add for the community. That's the one example. The other example is a friend who came out and said you know what I see that if you work in contracts and drafting there's a lot of inefficiency here because people spend so much time reviewing these contracts. There's got be a way to write a code that does a better job that more optimally does this service. And he went out and worked with some other people and learnt how to code himself and then work with the people to design a system that enables companies to do a quick document review and that's a piece of software that he now licenses to other people so again not everyone is going to entrepreneurial not everyone who go to law school has entrepreneurial spirit but if you have that coming out of law school and you see a need and we are right now in a space but there is a lot of needs that need filling and it makes perfect sense to me that that's a way you can add value. Third thing, I know I am speaking too much, but third thing I mention that advice for students coming out

LAWYERS ON FIRE

think about other ways in which law degree or legal expertise can help optimize business decision making. So you might be in a position where you can add value to other markets in atraditional ways. An example of that would be, you know I routinely come across the clients who say well I wish we had someone with legal knowledge to work in operations because then they will understand these risks like as an HR manager. Now again you are not practicing law and maybe you know the price point for law schools is frankly silly because now you are now overpaying for this education and expecting – you have been overpaying anyway – but now you are expecting to get a job that pays you more than that job that your subscribe pays but the truth is that might be in the long run more secure if you are not going to be entrepreneur, a more secure option for adding more value and you are harder to be replaced.

**Lawyers on Fire**: Is there any tools you would recommend to train yourself to train yourself up, any tool boxes?

**Zev Eigen**: I did this today when I went online before the presentation that I gave just to see if I just did a Google search for data science like I just want to learn I search for entry level or new online courses data science beginner something like that and you are flooded with advice. If you just add on any news aggregator the term data science, almost half the things in my flipboard were showing me news, clippings or articles about ways to learn data science. It's like oh ten new courses offered on YouTube, you know these online course systems for 40 dollars, I just saw like a few days ago. 40 dollars for a whole sequence of courses that teaches you code in python, a little bit of math, linear algebra and like those things are for free or 40 dollars you know and it's not, you have to pay tuition for some fancy school for 180 thousand dollars to learn how to do this stuff, you have to be disciplined, you have to be focused and you can't be intimidated by it and you are going to encounter things you don't remember from school or in the

back of your head, you remember calculus maybe and you have to go back and learn a little bit but that's how it is.

**Lawyers on Fire**: Any books you recommend, the old-fashioned way you like particularly?

**Zev Eigen**: I will have to get back to you, I will give you a list, actually I have a list that I keep that are pretty good but they are all titles that I can't remember, they are like really boring titles and it depends on the level what you want because some people I like Chris Surdak's book "Data Crush" and there's a new book called "Jerk". I love Chris, I love his work. His work focuses on sort of the organizational level understanding how data are important and how data science is important for driving decision making. But it won't teach you anything about the underlying technology. There are also books that are really focused on getting really from zero to one in understanding the technology so it kind of depends on for that one is more kind of text booky, that one's hard because you know now recommendation like, it depends if you have math's or not, so basically the skill set needed to do this is you have to understand calculus, linear algebra, matrix algebra and then some programming. So if you have no programming and no math it's harder to recommend stuff if you say yeah I am pretty good at math and I understand some calculus notation.

**Lawyers on Fire**: Last question, just to paraphrase an article you wrote yourself a couple of years ago with the wonderful title "Do lawyers really believe their own hype and should they?", just to go back to this whole conference you have seen so far. Should lawyers really believe their own hype or should legal tech people really believe their own hype?

**Zev Eigen:** So that's a paper I am proud of. I actually had a slide today of that paper with Yair Listokin who is a colleague friend at Yale. This was an experiment where we demonstrated that lawyers' over optimism bias serves them poorly. The

more overconfident lawyers are in their ability to argue acase, the worst they perform in our experiment. So that is definitely a sensationalized title which is publiczed in the Journal of Legal Studies, a prestigious general in the US. I am proud of that paper, it's a good paper. But the point you're making is should we believe our hype? As I said today too, I think we have to balance the evangelical side of data science with the doom and gloom side. Neither is correct. You know it's neither something that is terrible and horrible and lawyers should stay away which is some reaction you get nor it is something that's going to fix every problem all the time accurate a 100 percent of time without any issues legal or otherwise. It's a tool.

Machine learning represents a fundamental shift in the paradigm for using models and information to optimize decision making and it's important that I think everyone understands why that's a fundamental shift from our statistical models that we had before to this machine learning. That's important but that aside there is absolutely be some hype, absolutely going to be a lot of smoke and mirrors, yeah I also mentioned today that I am reading this book that describes when fraud is most common like in confidence games, schemes are most common, like with diet pills and things of that nature historically – you know "snake ware" they call it – and peddling false products false claims for things like diet pills or anything are always most likely to succeed when brand new technologies emerges. That's where we are right now. I agree this book I think is dead on and they actually interview people in the book who are con artists. They say oh yeah technology makes it easier for us to do this con stuff to trick people because people don't know what it is and absolutely we are going to see this in the HR space, in the legal space where people sell things that are snake ware. They are not really doing what are they saying, they are not great or they are held out as proprietary but they are not. I want to go back to one thing. The real key thing that is an important fundamental shift is going from statistical predictions to machine learning predictions. And that's something we all should understand. So

statistical predictions derive from statistics and these statistics mostly come from hard sciences. And hard sciences are great for predicting things in hard science like the decay rate of a radioactive isotope. So you want to know how quickly a radioactive isotope decays, calculus and statistics are the exactly the tools you want. What we have done because back in the day that's all we had, statistics from that sort of thing, we took those things and used them to predict stuff like human behavior. Whether someone has a higher or lower risk for coming out of jail and is a good person to be hired or you know whether someone is going to leave the company or what the likelihood of litigation is. We use those models. Those are ill-fitting models because humans are not radioactive isotopes. We behave unpredictably, we do things that take into tons and tons of information not just 1 or 2 or 3 or 4 things and we are not linear. We behave in ways that are not predictable. Just because we behaved vastly at time 0 doesn't mean we'll behave the same way or in some incremental additional amount at time 1 or time 2. We are not radioactive isotopes. The real fundamental shift from statistics to what we are using now machine learning is we discard or we don't rely as heavily on those models and predictions. And we are looking at patterns or training models in artificial intelligence to recognize patterns in voluminous matrices and overlapping matrices of data. We are identifying and pulling out information in those – I won't say three dimensions that's wrong – 5000 or 10000 dimensional arrays to predict stuff. And that's better suited for predicting human behavior than the linear model or not even linear but statistical model of predicting stuff derived from hard sciences. That is a fundamental shift that is a very very big deal and I think everyone should understand that part of it but there is a lot of hype a lot of smoke about it.

**Lawyers on Fire**: Isn't it scary I mean if a machine learn system predicts what I a going to do?

**Zev Eigen**: Why is it scary? What are you afraid of?

LAWYERS ON FIRE

**Lawyers on Fire**: Maybe they will have some consequences linked to that one if it goes into wrong hands?

**Zev Eigen**: So I mean like anything else there are issues with how humans behave, right, and my view of humanity is I believe we are scared herd animals that's my personal view of humanity. I am more worried about proliferating guns like having guns in the hands of scared herd animals than something like this because I think that's more dangerous. But we already have plenty of things to harm people with. We have bombs, we have guns, we have all these other things and there absolutely is an application where you could see this being used for wrong purposes. You can say the same things for cars or for you know anything else you know like electricity or anything. You know I have a friend who is always worried about e-collars for dogs. You know what an e-collar is? So I have a great Dane, a German master, very big dog 160 pounds, and I am a fan of training dogs with a static e-collar. It gives – it's not a shock, its static kind of thing. He says it's very dangerous. Well, if you use it correctly it's actually really good, it helps train them and it's a very well trained dog, a service animal now, very well trained. And I said I want to remind you that you can use a leash, a standard leash, to terribly harm a dog and people misuse leashes to choke dogs and do things that are terrible to dogs. My dog is fine, he is a spoiled, lazy baby but this is a tool. Data science is a tool and absolutely you can use a tool for ill purposes and I don't think we should do that. The other thing that is really important to stress here is – and this is why I am doing Cherry Tree – data science should be use for good. We should have more people focused in HR space and in every space for using data science and showing the applications for how we can use it for good. Cherry Tree uses data science the same tool as a way of reducing bias as helping understood communities, helping people go back to work, who need to go to work, improving the economy, improving outcomes for minorities in this country or anywhere and helping employers too at the same time. Everyone is benefitted, the employers, the economy, tax payers, individuals, there is less

crime because instead of being on the streets people are working and so many benefits so it's absolutely true that things like data sciences can be used for bad proposes but our goal and particularly my goal as a data scientist is to find the good purposes and there are plenty of data science for good that's why it's called data science for good applications in optimizing charitable giving, identifying need and optimizing giving – I was just talking to a friend – we do this all the time, you know there is a food waste especially in this country that is voluminous. They have vegetables that are out in the supermarket and then they pass a day and then they throw them out. They are perfectly good to eat but they throw them out or there is bread that is day old they can't sell it – they throw it out. What if we had optimization tools that looked across huge vast spaces of geographies and found all that food waste and found a way to optimize it and then also identify the places where people need it the most and optimize the gathering and distribution of those scarce resources? That is great thing that would help a lot of people and that's one example. Cherry Tree is another example. There are many examples. We need to do more of that and I think as we move forward I am very proud of Cherry Tree because it is something that is a for profit venture that also is doing good and that's why I did it. I don't I mean charity is terrific but I like things that motivate people to do it and win, win, win is the thing that makes everyone do it more.

**Lawyers on Fire**: Thanks very much Zev for talking to us today.

**Zev Eigen**: Thank you so much. I really appreciate the opportunity today.