PODCAST TRANSCRIPTION SESSION NO. 269-ARJUN KAKKAR

Welcome to the Lend Academy Podcast, Episode No. 269. This is your host, Peter Renton, Founder of Lend Academy and Co-Founder of LendIt Fintech.

(music)

Today's episode is sponsored by LendIt Fintech Digital, the new online community for financial services innovators. Today's challenges are extraordinary with the upheaval affecting all areas of finance. More than ever before, we need to come together as an industry to learn from each other and make sense of this new world. Join LendIt Fintech Digital to connect and learn all year long from your peers and from the fintech experts. Sign up today at digital.lendit.com

Peter Renton: Today on the show, I am delighted to welcome Arjun Kakkar, he is the VP of Strategy & Operations at Ekata. Now, Ekata is a super interesting company, they've been around for quite a long time and they've really focused on combating online fraud. They've developed these series of tools to help do that which we delve into some depth. We focus specifically on account opening fraud in this episode, we talked about how fraudsters are approaching it, the different ways they do it and, obviously, the different ways to combat it, talking about how reducing friction can lead to false positives in that and how companies need to balance that out.

We talk about their Identity Graph and Identity Network which is their sort of secret sauce in really bringing in all the data of really billions and billions of data points coming together to give people a good indication on what is actually real identity and what isn't. It was a fascinating interview, I hope you enjoy the show.

Welcome to the podcast, Arjun!

Arjun Kakkar: Glad to be here.

Peter: Okay. So, I'd like to get this thing started by giving the listeners a little bit of background....you've had a pretty interesting career to date so maybe give us some of the highlights before you got to Ekata.

Arjun: Great, yeah. So, I think start from right now, of course, I lead Ekata's Payments & Financial Services verticals, I've been in the company for over six years now. I started as the Head of Strategy and progressed on to become the VP of Strategy & Operations and in my current role, I finally progressed on to this and keeping worthy goals on the operations side.

Before Ekata, which was your question, I worked with Booz & Company and Booz Allen Hamilton for over six years and that was advising mostly Fortune 500 companies on problem solving and business strategies. Before that, before business school, I was Co-Founder of a startup and an engineer by training before that.

Peter: Right, right, okay, okay. So, maybe give the listeners a little bit of....describe Ekata for us. I know it was at one-stage in the past called White Pages Pro, I believe, so maybe explain a little bit about the history and how you describe it today.

Arjun: Sure, sure. So, Ekata today is .....I describe it as a company that provides global identity verification using APIs. We were a part of White Pages as a company before this and we were called White Pages Pro and we've been operating for over a decade. Throughout this time, we've been growing in double digit percentages, we have over 1,700 customers. We started White Pages Pro as a part of White Pages because we realized identity data and many people were coming to our consumer website looking for identity data or fraud detection. You know, it's just one of those things you realize oh, there's an opportunity here.

So, we started contacting these companies and you will be surprised, many huge companies out there that used to use our data on the consumer website and they became our first few customers. That was my point about double digit percentages, it started like, you know, a $100,000/$200,000 business to going over $50/$60 Million run rate now and still growing at double digits percentages. I like to think of Ekata as the value we provide, we create for our customers is.... the foundation of it lies in our identity data and there are two secret weapons that we have like two differentiators.

The first is our Engineering and Data Science team which are global teams, part of it is in Budapest, and these teams result in normalizing and keep the data quality high which is really required by identity data. It's a very hard problem, it's one of those problems, Peter, that, you know, on the surface looks like, oh, okay, I can do that, but once you go deeper, you need a big team working super hard.

Peter: Right, sure.

Arjun: And the foundation of that is it…we build risk signals and scores using machine learning such that our customers can make better risk decisions. You know, those engineering teams are a part of our power.

The second piece of major differentiators....now, as I mentioned, we have over 1,700 customers globally and this includes banks, fintechs, merchants, payments providers and they all are using us using APIs so, you know, they're paying our service on locating the data and now our customers on the backbone of our network, they further contribute value back to our customers. We can talk more about our network later. That is truly a differentiation because only we have our network and, you know, the risk signals coming out of it are really valuable.

Peter: Yeah. I do want to dig into that, but before we do, I want to just ask sort of a broader question. I'm talking about online fraud because, obviously, it's a major problem globally. Maybe we could just touch on what are the biggest challenges today in combating online fraud?

Arjun: Absolutely. So, it's helpful to think of it from the broader trends that impact the (inaudible). If you look back at the trends related or the mega trends related to fraud, the fraudsters are getting access to a lot of identity data, what lives in the dark web. I'm sure, Peter, your identity, my identity are all stolen and out there being from multiple databases and over 80% of stolen identity, stolen records online are actually identity records. So, that's a big risk, right, like the have the data.

The second major broad trend is that online banking and commerce is taking share from physical banking and commerce and we all know that, also today, right.

Peter: Particularly this year, yes.

Arjun: Exactly. Essentially, the next result of these things is that the fraud industry is growing, you know, if there's somebody doing an analysis of the fraud industry (laughs) like, you know fraudsters having their conference are like our industry is growing and the biggest challenge out there to fight fraud is to get access to good data so you know, fraudsters are getting identity data. We need like….the good side needs access to good identity data to fight fraud and not just identity data, other data elements too, but that's the foundation of it.

Peter: Right, right, okay, okay. So, let's maybe just start to dig in because one of the big challenges obviously is when a new customer comes along, they want to open an account, it could be anything, it could be a savings account or it could be getting a loan, it could be, you know, even just trying to… something so there's this account opening moment that happens. Maybe you can ....this is obviously a really key point for you guys and I know for the fraudsters so let's start off with say how are fraudsters thinking about that particular moment, the moment of account opening.

Arjun: So, it's helpful to...if I may step back and also give some context on the account opening perspective, right.

Peter: Sure.

Arjun: Particularly, you and I touched upon the whole COVID thing. I mentioned that online banking and commerce is taking share, but further there is this crazy surge right now. So, for example, in banks I think there is data of just post-COVID, the number of new account registrations doubled and nothing surprising, right, you would expect that. What is interesting is that about 40% of people doing that now like using online banking said that we won't go back to the branch.

Peter: Right.

Arjun: So, it's kind of ....it's not just a temporary thing, it's quite a shift. If you look at e-commerce, you know, about 15% of commerce is e-commerce, was e-commerce in 2019 and within two months it became 30%.

Peter: Right.

Arjun: (laughs) How much of that 30% will go back? Will it go back to 15? It might go back to 20/25, but still it's like a massive short. What usually would increase by 2% each year, now it's up to much bigger. You put that together, one thing that stands out is new account opening, right, and why new account opening, because it's not just an increase in transaction volume. Many people were not buying grocery, opening accounts or grocery shopping.....people who did not have online bank accounts are opening new bank accounts. The challenge in new account opening, particularly what's very hard, is that you do not have any historical data to go by.

Usually, if you are on a bank account, your normal behavior over there, they know your behavior and there's a change or anomaly in the behavior, they can detect that. Same with e-commerce companies that you see, it's normal shopping behavior and then there's some change, but that's not the case here, right, like new account opening, you need access to better data. So, zeroing in on your question how do the fraudsters think about it.

They do it in two possible ways, they do it using stolen identities, which we talked a bit about, and then they do it using synthetic identities and what are these things? Firstly, stolen identity fraud is essentially a fraudster stealing your identity, Peter, and saying that I'm going to.....they'll get your identity data from the dark web and they'll start impersonating you on any online transaction. They would usually move very fast because they want to do everything quickly and buy the products or transfer money from your accounts. They do it in multiple locations because the data is gathered from the dark web so multiple fraudsters may be trying to defraud you and before you realize that they want to do things very fast there's a certain behavior that you should know off of stolen identity fraud.

One interesting thing to note there is that if I'm a fraudster and I steal your identity, I'd still want the product delivered to my address or my email or my phone numbers, some delivery location, I have to change, tweak one little thing. That's the kind of thing that good data has to catch.

Synthetic identity, this one is very interesting because the fraudster behaves differently here. They will start by making a Frankenstein identity that is attached to a real, but somewhat of a less used government identifier or something like a Social Security number so they take an SSN or they might even sometimes create a random SSN and they will attach synthetic identity to that like all made up identity elements. So, they'll have your address with somebody else's phone number and somebody else's email or they might even not be real.

The interesting thing with these people is that they will start applying for credit and start building a history, It requires a lot of patience do that, they are very patient people because there's a big prize at the end of it. So, once the credit score rises over a time period, even years, they will secure larger credit extensions and finally, they'll do what is called "busting out." So, they'll bust out, in other words, they'll just max out on the credit and vanish. You can imagine like the first one is like very transactional fast, a steal and get a product or transfer some money and go

away, This other one is this very patient, build something like almost from the foundation and then bust out with like tens of thousands of dollars.

Peter: Right. So, it sounds like with synthetic identity, it must be....I imagine they're both challenging, but the thing about synthetic identity is that I always think....it's operating in a normal way, like you said they're opening a bank account, they're funding it, they're taking out a loan, they're paying it back so this look real. So, is synthetic identity actually harder to kind of fight against than stolen identities?

Arjun: Absolutely. It's a very hard problem to solve. Some of the things that make it hard, first of all, detection and labeling is very hard. So, every time I talk to folks about synthetic identity in the market, I ask for like hey, can you give us outcome data because, you know, that's how you usually test and some of them do, but they are always questioning even themselves because nobody is complaining about them, very rarely complaining about it and it looks so credible.

Peter: Right.

Arjun: How do you figure that out like and if you do.....one of the important things and anything related to machine learning, anything related to rule writing, pattern recognition, you need outcome data. You need to see like okay, whether this is a fraud, fraudulent synthetic identity or not so that's hard.

Another interesting thing about it is since the fraudsters are looking for the big prize and are longer term thinkers, if you will, many of them will pretend to be just false positives even if you catch them. So, you catch them and you say like hey, you are a fraudster and they say, no, why do you say that, I'm not and how do you prove that, right. This is all online, right, so they are very sophisticated, there are like real crime rings doing this.

To figure it out, the magic of having a web structure, something like what Ekata has of different identity elements tying together helps find data inconsistencies. So, if you are joining up a Frankenstein identity, there is a database of say the truth, you should be able to compare against the truth and find inconsistencies and see whether something is fraudulent or not. You know, that inconsistency thing you might recognize, Peter, it's like .....it is true even for stolen identities because people are changing one address or one email or usually something like that, but, particularly, a very strong indicator for synthetic identity fraud, it's one of the best ways to catch it.

Peter: Right, right, interesting, interesting, okay. So then, the challenge that often platforms have, particularly lending platforms, shall we say is that ....you know, there's always this kind of battle or balance, you've got to struggle between increasing more friction and encouraging more signups. You want more friction to reduce fraud, but you put too much friction and no one signs up so how do you recommend that online platforms balance this friction and sort of the openness of encouraging signups.

Arjun: Yeah. One of the pet things I always talk about is customer insult. Part of our customer insult is like you catch one element of fraud like you were to catch one fraudster, how many good customers are you declining, right. We've done estimates on this and across industries it is a very big delta. So, if you have to save a certain dollar value related to fraud like $100, it's usually a 10X factor of how much you're losing on revenue because of things like false declines or friction. As they add up….the problem is hundreds of billions of dollars costing as cost to the industry.

To your question on how do you think about that? Most of the people who are sophisticated have started thinking of it as a two-step process and I'm simplifying it, of course. The identity verification process should be two steps, the first one being passive probabilistic method so you know, you are not trying to ask right away, Peter, show me your ID when you join like we won't do that right away. But, based on your behavior, the elements of data I have collected, I do probabilistic risk assessment not a definite risk assessment. We can talk more about that. Add friction for only the bad customers, like the riskier customers so just eliminate friction for people you think are trustworthy and we're going to push it low risk and that's pretty much the insult.

The second step is you do put more friction for people you feel who are risky and on that point, I think, it's worth talking about this thing that I told you about, probabilistic risk assessment because, generally, you know, if you go online on banking, lending, things like that people think...I should say people think because the industry has wizened up over time on it. But, you almost always try to look for a definite answer, is this truly Peter, right, who's coming online and you can never get a definite answer so that's the historical school of thought that always tries to get definite answers.

It includes, you know, using government issues, static identifiers such as government issues, photo IDs or social security numbers and things like that, age, you know, things that are somewhat immutable, right, like static identifiers. You cannot change your age, of course it increments by one each year, you cannot try to change that. Social security takes a very ....it's a crazy hard process to change it, but you can change your phone number, you can change your email, those are dynamic identifiers. You know, these kinds of identifiers are usually country specific and credit bureau-driven so you cannot get a global solution. There are those elements like that and there are, of course, the problem of they've always been through some form of data breaches.

On the other hand, probabilistic risk assessment, you can think about dynamic identity elements, things like name, email, phone, IP address, all of these things. You know, it's not like your IP changes quite regularly, your address changes, at least every few years. Your device ID changes a lot, your email, you might use multiple emails, these are dynamic identifiers and using these for risk assessment for the frictionless side….when a customer comes in and you have these identity elements, can I use these identity elements to see what's the risk and to do that you can apply good data science behind it.

You will never say this is surely the answer, there is never a definite answer, but you'll get a probabilistic answer of like okay, I'm confident enough to let this person go forward and that's the key of it, that's the kind of thing that Ekata enables. Things that I'm telling you, Peter, this is becoming accepted standards in the industry, right. There's something that regulations might require you to do, but these are things you do on top of that to make sure that your customers have a good customer experience.

Peter: Right, right, okay. So, I want to talk now about the Network that you just touched on earlier and the Identity Graph. Maybe you can explain what they are and how these help identify fraudulent account openings.

Arjun: Yeah. So, in the early introduction to Ekata I mentioned about our super powers (laughs), this is like our ingenuous build. The Identity Graph and the Identity Network are two distinct data assets. The Identity Graph is essentially built from over a hundred data sources globally and this is licensed data. We take care of compliance to global security regulations, we have a data sourcing team that's very ...you know, especially things like GDPR, there are so many regulations around the world, it's a hard thing to do so we've made a discipline of that.

Our Graph, essentially, help validate linkages between identity elements and we have over 8 billion identity elements globally and the linkages could be between things like emails, phones, addresses and they are linked to names off of each other. You can think of it an as interconnected graph, right, Peter, like if you cohabitate with someone, you will be linked to them through your address, for example, and the graph goes very deep. It also gives you what we call metadata or like additional information related to any element so for example, with the phone you'll get information like phone carrier or line type or email, you can get email validity and other such signal and these are all risk signals because we focus on risk signal. That's our Identity Graph, again, like based on source data.

Our Identity Network is the other super power I talked about like the part about our customer access and it essentially derives insights from billions of anonymized customer queries and it finds usage patterns between these queries. So, you know, all our customers use our APIs so they ping on to our servers and we anonymize those queries and we look for patterns that will find anything that's abnormal or assess the risk or other things that are related to usage patterns that will give us more information about risk. There is no question about this, this is only unique to Ekata because, you know, these customers calling identity elements is very unique to us.

These two combined together....I always think the biggest value...I mean, these are individual assets, the biggest value comes in the combination of these. So, if you put these linkages, the additional meta data I talked about and usage patterns from the Identity Network. It becomes very powerful because it becomes harder and harder for a fraudster to penetrate like they can make up a part of your profile, but they cannot make all of these up, you know, there's some signal that will give you away.

Peter: Right, right, okay. So what you're doing is you're getting....like in real-time, obviously, probably many, many time everyday you're monitoring transactions, real transactions around the world and then...I imagine, I mean, as you're building up I can see how it would get...it build on itself, right, because the more data you get, the more accurate it is, the more data and it's just kind of ...people continually feed into it to bring it more data....

Arjun: Flywheel effect, right.

Peter: Yeah, exactly. Let's talk about the network score. I noticed that you released this network score earlier this year, maybe explain exactly what it is and how this is useful.

Arjun: Yeah. I will talk about the Network Score, it's a part of our Identity Network. So, you know, the Identity Network that I just described gets all of these signals from the network. What we decided was to simplify it for our customers, not all of our customers have machine learning shops, but we do. So, we can build a score that indicates the risk in our network using a single score so we put our machine learning team. You know, we have a machine learning team because we are already building other scores, for example, based on our Identity Graph. We build this new one from the Network's site particularly because there was a lot demand from the market that we want to use something related to your Network.

So, what the team did was using our customers' outcome data....so we have multiple customers that give us outcome data which we use for modeling, we found what is the...we built our machine learning models to give out a Network Score that helps them and then we test it back with them. And we realized.... this is a very interesting thing we realized, Peter, that the Network Score is powerful by itself like the Network Score tells you the risk online based on behavioral patterns or based on usage patterns, but it's particularly helpful when it is combined with the Identity Graph because when you put them together, it is the additional data of somewhat orthogonal data sets like this is reality sitting here and this is behavior and you combine the two.

There's an interesting study that ....one of our customers. they actually found that by combining the two they were able to find about 20% of the transactions that they were declining were actually false declines. These were good customers they were declining and these are the customers which have got low risk in our scores related to the Graph, our confidence score, and low risk related to Network Score so that's what our Network Score enables.

Now, I wanted to add an additional point there that what we are doing now is a lot of network-related innovation so we first release the Network Score. We are releasing a additional attributes related to the Network. I'll give you one example which offer network attribute that is always there with us and which is one of our most powerful signals, its our email first seen. Now, the way we think about email first seen is like when did you create your primary email, Peter?

Peter: Sixteen years ago.

Arjun: There you go (laughs), yeah, that's me too. As soon as Gmail came.....

Peter: As soon as Gmail came out exactly, I registered my name@gmail.com.

Arjun: Yeah, yeah, same, my last name because it needed six characters, Arjun wouldn't do. So, anyway, I got that and I retained that ever since. What's interesting to know is that 90%+ of consumers retain the same email for over three years and only 3% of users use emails that were created in less than a year back. What's interesting in online transactions is the spike in the number of recently created emails, why is that? It's fraud, because the fraudsters are creating these email IDs to get delivery of things or take over accounts or things like that. That is the key that.....you know, we give the signal based on our Network and other sources of email first seen which is one of the most predictive signals.

You'd take…. they think of another signal like when did....if you put your phone number and email together on a website or in a banking app and during account opening Ekata can check in our Network that these two are seen together recently. Now, there is a risk signal associated with that and we discover this all the time, Peter. I don't want to speak too much more of it because I don't want to reveal too many secrets, but we see this all the time based on our experimentation, our data science team on new signals and we are starting to expose these signals to our customers and new APIs such as our account opening API that's coming up.

Peter: Right, okay. So, we're going overtime here, but just before I let you go, one more question. You touched on it just there, but maybe expand on how do you see this going down the track. I mean, like it seems like it's a constant battle with the fraudsters. They're innovating just like you guys are so what's coming down the track that we can.....without giving away too many secrets, that we can expect from Ekata.

Arjun: I can tell you one thing that is....on the surface it's seems like the fraudsters are winning, right, like they are a growing industry...just like…and that's unfortunate, of course, just like online commerce and online banking are growing industries, if you actually drill down deeper, you go one level deeper and you see how much fraud you have per unit volume, it's actually flattened and it's going down a bit.

Peter: Really, interesting.

Arjun: Which is a good feeling and it makes me feel like we are winning, It's still there, it's still a growing industry, but per unit I think we started to slow them down a bit. The part to keep in mind there is that, you know, this year, we don't have data for this year yet. This year, the fraudsters would be winning.

Peter: Right.

Arjun: There will be two things that will be happening. One is either the fraudsters will be winning because we have not seen ....you know, machine learning requires previous data for analysis, we don't have previous data so fraudsters are either winning or we are causing a lot

more false declines like declining good customers. So, what Ekata would do, Ekata is...we realize that we are in the space wherein we are market leaders globally, but we are not done yet, there is so much we need to do to keep improving. The examples of stuff, like I told you on Network, are focused on Network to give something truly unique to every customer that they can learn from our customer network or the risk detection.

The second thing is our global expansion, you know, we have got offices in Budapest, in Amsterdam and Singapore now and we are continuing to grow in other geographies and expanding on those markets. Throughout this thing, our focus is on providing high efficacy risk signals so we will just continue in that direction. I think it's a long path, we think long term and we can easily see our path out for a decade.

Peter: Right. You have job security there because I'm sure there's...fraudsters are going to keep battling against you. So anyway, Arjun, I really appreciate your coming on the show today. It was a really fascinating conversation.

Arjun: I enjoyed it a lot, Peter, thanks for all of those questions.

Peter: Okay, see you.

Arjun: See you.

Peter: You know, I was just chatting with Arjun there after we stopped recording and he made a comment that ...he said, these fraudsters are smart. I mean, they have all the latest tools, the latest technology, smart people working for them so we certainly can't underestimate them and, you know, makes sense that this year, with the increase in online behavior and sudden changes in behavior, that fraudsters are taking advantage and they have really made some gains this year.

We need to be doing everything we can to really create an accurate picture of each person coming through our website or our mobile app and making sure that they are who they say they are. That's obviously what Ekata are doing, really, you know, helping give companies more confidence that help them manage their risk in this way.

Anyway on that note, I will sign off. I very much appreciate your listening and I'll catch you next time. Bye.

Today's episode was sponsored by LendIt Fintech Digital, a new online community for financial services innovators. Today's challenges are extraordinary with upheaval affecting all areas of finance. More than ever before, we need to come together as an industry to learn from each other and make sense of this new world. Join LendIt Fintech Digital to connect and learn all year long from your peers and from the fintech experts. Sign up today at digital.lendit.co

(closing music)