EPISODE 1343

[INTRODUCTION]

[00:00:00] KP: Money laundering is not a new crime. However, the growth of digital communications has greatly expanded the opportunity for money launders to find innovative new ways to hide their true intent. Some estimates suggest that money laundering could be as high as 2% to 5% of the world's GDP. Unit21 is a customizable no-code platform for risk and compliance operations. They offer a simple API and dashboard for detecting and managing money laundering and fraud. Today in the show. I speak with Clarence Cho, cofounder and CTO of Unit21

[INTERVIEW]

[00:00:39] KP: Clarence, welcome to Software Engineering Daily.

[00:00:42] CC: Thank you. Thanks for having me here.

[00:00:45] KP: Tell me about your background and how you ended up starting Unit21?

[00:00:49] CC: I'm the CTO and cofounder of Unit21. We are a no-code platform for detecting suspicious events. My background is in security and machine learning. We do some similar work over here and help companies to detect anything suspicious to them, whether or not it's related to security. A lot of our companies use us for detecting money laundering, fraud and really anything that is suspicious to them.

[00:01:17] KP: That's an interesting use case. If I were in charge of one of those platforms, I definitely wouldn't want it used for fraud and money laundering and these types of crimes. But I also know, there are a big enough platform, just weird edge cases happening. How do you tell the difference between that an actual fraud?

[00:01:35] CC: For something, it's just a lot more clear. For example, when you're dealing with payments through the credit card network, you know something is fraud when you for example

receive a charge back request. This is when the credit card transaction has been canceled or disputed. For something that's a little bit less obvious, so money laundering for example is notoriously hard to get ground truth around. If you think something is money laundering, you frequently don't get a very clear feedback loop into whether something is truly an attempt at money laundering or not.

Now, a lot of this is obviously a challenge because when we're doing detection, using machine learning or using anything heuristics, then having a very clear feedback cycle is important. But especially when dealing with something that has incredibly long feedback cycle that has to go through, maybe the US judiciary system. Then it's a little bit harder. That's a lot of what we do. We try to build models, we build rule sets for customers to detect money laundering, detect things that are a little bit harder to get feedback cycles around. We help them to be more sure around the things they're finding so that they can be confident in their detection schemes.

[00:02:46] KP: When I think about fraud and just transactions online, it's e-commerce that first comes to mind. Is that a common industry that you can deploy to?

[00:02:55] CC: Actually, e-commerce is not the most common type that we have. A lot of FinTech, financial institutions care a lot about money laundering because there's real consequence for them. Now, money laundering is a little bit interesting, because companies don't necessarily lose money from their consumers laundering money in their platforms. It's a little bit unintuitive. When someone launders money on let's say your payments platform. Then you may not be losing money, in fact, you may stant to gain a bunch of money from this activity.

Now, what really causes companies to care a lot of the time is regulation, because money laundering is this event with huge negative externalities, every dollar laundered through the financial ecosystem causes ripple effects throughout the rest of society. Every dollar laundered goes into the bloodline of organized crime, human trafficking, et cetera. The government cares, even though private businesses may not stand to lose anything.

This is why if you were to operate any kind of financial institution, any money transmission activity, any broker-dealer activity within the U.S., you have to prove that you're doing all you can to prevent money laundering from happening in your systems. In the least severe cases of

enforcement, if your platform was found to be involved with money laundering, you'll be given a warning. In the worst cases, you maybe fined up to billions of dollars. The person in charge of the money laundering program may also be sent to jail.

There's real consequence here and the most typical kinds of companies that integrate with our platform is financial solutions, payments platforms, crypto exchanges, forex platforms, anyone that deals with the transmission, storage and movement of money.

[00:04:45] KP: It is the analysis typically in batch or do you do a real-time service?

[00:04:50] CC: Actually somewhat in between. We don't just do batch processing and we don't just do real-time processing as well. A lot of our customers sends us data in real time. When the transaction happens, they send us a transaction, but they don't actually wait for the response in real time. Now, we are building stuff that allows our customers to get results in real-time, but today, a lot of our customers don't really care about getting these results in real-time, because we're not in a critical path. Why this is important is because, most money laundering are sophisticated fraud isn't detected in real-time. In fact, you may not want to detect and give the feedback to the adversaries in real-time, because you want to observe patterns.

Observing patterns requires you to collect data and take in one, or two or three fraudulent or malicious events before you make a decision. This is how you capture the entire network as being greedy and capturing single events at once. This is a lot of what our system is good at. For customers that don't send us transactions real time, we also support this. We support this because a lot of our customers don't necessarily have a lot of engineering resources. They are lying around available for them to build API integrations into systems like us.

What we do is we allow them to export data in CSV formats and arbitrary formats from their own internal transaction processing systems, and then upload them into our systems so that they don't have to wait to get started. Our quickest customer that has caught and integrated has taken less than 24 hours because of disability of hours and just data without having to write any code at all.

[00:06:25] KP: Could you expand on that onboarding process? How do most users get their data into your system?

[00:06:32] CC: By far, the most common way that customers have integrated data into our systems is via API. This is because API is just the most customizable and the most flexible way of data ingestion. We don't have to worry about file formats changing and the **[inaudible 00:06:47]** that we have available no longer being relevant. Customers can change this at any time they want. Most of the time, this involves them writing some code to integrate into our APIs. The real challenging thing about building our software here isn't the fact that we have to ingest this data even though we do ingest a large quantity of data from all of our customers. The most challenging thing in fact is that, all of our customers deal of transactions, and events and users, but each of them actually has a quite different definition for what a transaction does, what a user can do in their system. This results in us having to build a somewhat more flexible system than we expected initially in order to support all the different use cases.

I can go into a few different examples about why transactions are different and why detecting fraud and money laundering across our customer base is interesting and also challenging. One of our customers coin base for example defines transactions as a trade-off currency. This currency is commonly U.S. dollar and Bitcoin for example. You're buying some Bitcoin with some U.S. dollars and you transact it yourself. Now, another of customers for example, cheaper cash, which does cross-border, people-to-people payments within the African continent, deals with transactions that are flowing from one person to another. Now, the kinds of fraud models, money laundering models that are relevant for each of these customers are very different.

Going one step deeper, the kinds of ways that they would model their transactions are also quite different. Just because of the different patterns of transactions and clusters that can form possibly from their own networks. We have to be flexible enough to encompass the different types of fraud models, money laundering models that exist within their systems and be able to do this in with this as little mental effort from us as possible. That was perhaps the most challenging things about building our system.

[00:08:50] KP: Is that the motivation for developing a no-code solution?

[00:08:54] CC: That is actually. A lot of customers, a lot of companies that were in financial services before really only had a couple different options when they wanted to build something to be compliant of regulation, and they needed to do this because they need to be compliant with the laws so that they can operate their businesses. The first option is of course to build it themselves. We see a lot of companies doing this, especially if they are tech first companies, if they're building a FinTech company in the truest sense. A lot of companies will still choose to do this. But most FinTech companies, most financial institutions are not fintech companies. They aren't technology first companies and don't have **[inaudible 00:09:38]** resources to spend on building up transaction monitoring systems to detect fraud and money laundering themselves. A lot of the time, it doesn't even make sense to.

What they do and what they used to do was to go to one of the off-the-shelf solutions out there. A lot of off-the-shelf transaction monitoring systems for detecting fraud and money laundering was built in the age of transactions when you could only transact with banks, mortgages home loans, things like that. We've realized that these systems were not nearly as flexible enough to be used in the age of FinTech, where there are literally hundreds and thousands of transaction models being invented, and every day, there are dozen new different ways of laundering money than there existed before.

This is a great, great thing that's happening for money launders out there, but it's not so great for the regulators, for people trying to keep the financial ecosystems clean. Because it is easy to present a new way of transacting, but it is really hard to find new ways of exploiting the system and to find ways that we can detect when someone is exploiting the systems. A lot of what we try to do was to lower the barrier off detection, and to make it such that people within companies that are in charge of detecting fraud and money laundering don't have to be the best engineers or don't have to be the best data scientist available. Frequently, they're not. Frequently, they're not even engineers. We wanted to give them the tools and the powers to engage with their data, the data that they're collecting anyway in ways that they previously weren't able to. Our tool allows us them to do this without having to write any code and knowing about databases, or tables or queries at all.

[00:11:29] KP: Let's put ourselves in the shoes of maybe some with the title, senior fraud analyst. They're going to utilize the tool. Is this something they log in and configure or do you have some maybe pre-developed recipes for them?

[00:11:42] CC: It really depends on the customer and what they're looking for. About half of our customers come in and look at us for expertise, because we are involved with a large number of different industries, and different customers. We have a unique view into the type of rule sets, types of models that are relevant to different customers. For example, if you're starting a new crypto exchange today, and you wanted to find out how about other crypto exchanges dealing with this problem. Then you don't need to come to us with all the expertise already prepared. What we can do is to provide you from within our software the most popular kinds of models, the most popular kind of detections, scenarios that other crypto exchanges have also found useful.

We do this not only by looking at a number of models that different crypto exchangers using our platform are deploying, but also by telling you what the false positive and negative rates have typically been when other crypto exchangers have used this. Now, then you can get started without necessarily having to build all the expertise and hire the a most competent AML experts in house to do what you need to get done. A lot of our customers however also come to us with all this prepared. They've hired the best teams, they have operations teams that have been dealing with these problems for years and years, whether within their companies or within other companies.

Now, what we then allow them to do is to use our software and express what they want to express without having to describe this to a team of engineers, that may or may not understand exactly what they mean. This is really what's challenging here, in the detection piece. What we're purporting to allow our customers to do is not only to have a detection model that allows them to detect anything suspicious, but also to promise that whatever they want to find, they can express it within our system without writing code and we can productionize this, allow to back test this, validate this and deploy it without involving any engineers. This is challenging. We're not all the way there yet, but we've gotten so far along since we started and we're confident that this is not only a really interesting problem to solve, but also a never-ending problem that we can solve.

[00:13:59]: KP: Do you have any insight into the way your customers consume the data and feedback you're giving them? I could see it being something where there is just rough analysis and exploratory things or perhaps there's some split-second decision real-time automation use cases. How do people deploy it?

[00:14:16] CC: Again, this depends on how our customers use us. They way we've designed our system is dependent on how our customers are thinking about risk mitigation and the appetite for risk. Some of our customers have teams of people that look through queues of alerts that are flagged by our system. Whatever they designed to be suspicious in our platform, flags alerts and flows out to these queues. Now, these operations teams then go through these alerts one by one and make decisions on them. Because it involves some kind of human review effort and these reviews frequently take minutes, if not hours after the transaction occurs. The user, the customer frequently doesn't block transactions from happening before it flags an alert.

Some of our customers however choose to block these transactions and this is because, maybe they have a lower risk appetite or because every transaction presents them a much larger financial risk. Now, you can understand this if some of our transactions that we deal with are maybe on average $50. Well, some of the transactions that we deal with are an average $500,000. In the case of the $50 transactions, you're maybe, okay, letting a couple slip. But in the case of the $500,000 transactions, you may not be so willing. In the latter case, then what our customers frequently do is to put the transaction in a hold, until the agents, until the operators have a chance to review them. After the review, if the operators decide that this transaction looks okay, they'll let it through. If it doesn't look okay, then what they'll do is they will pause the user account, they will suspend the account, maybe place them under further scrutiny, lower the transaction operating limits, et cetera to prevent any further repercussions on their systems.

Now, some of our customers also use as entirely automatically. Whenever something is flagged, even if they don't have teams of people that review these alerts, they want to be able to automatically resolve these and automatically deny any of transactions going into any one of these cues. If the alerts that are flagged are known to be true positives, they are known to be bad and you know that you don't want to spend more time reviewing them, then this is what we can help them to do as well. We can help them to automate the denial of all these transactions,

so that they don't have to spend any of the human time on reviewing these. This saves them time.

[00:16:46] KP: My understanding of money laundering is you want to create a large number of transactions so it appears that they're coming from a diversity of sources when they're actually really one source behind the scenes. So that person doing that is by design trying to hide their behavior, leaving you varied clues for machine learning to find. Is it possible to achieve absolute digital anonymity? Could this be a cat and mouse game that doesn't end well?

[00:17:12] CC: Yes. This is I think what makes the problem so difficult and it's something that's not very specific to money laundering. But any type of machine learning applied to security problems or any adversarial contexts. This is what I spent a lot of time looking at previously, in my previous life when I look at how attackers can make use of systems to evade detection of "next-generation" or machine-learning driven detection systems. Now, I think the key to a lot of this frankly is to layer machine learning systems with heuristic decision engines, i.e., rule sets. When you know the patterns that you're looking for, the ways that attackers can evade you are limited. But when the patterns that you're looking for are broad and unpredictable, then it's a lot harder for an attacker to try to go and guess the patterns that caused them to be detected.

This is what a lot of machine learning can play a part in, in detection. When you have fuzzy rule sets that apply to detection engines, then the real key over here is that, we want to make these rule sets, they want to make the detection criteria as unpredictable as possible, so that the traditional model of attackers trying to evade detection can no longer work. The first thing you brought up about there being very few data points that an attacker has to be involved with is still relevant. Now, when we look at things like detecting spam or detecting DDoS attempts, this problem is not going to be prevalent. Because by definition, spam and DDoS attacks are large-scale attacks. It's very much suited to the machine learning paradigm, because you can collect examples of spam emails, you can collect examples of network traffic attempts that are belonging to the DDoS category and you can run classifications on all of these.

If you're look at very targeted attacks, like money laundering or sometimes types of fraud or APT threats in malware, then it becomes a lot harder for problem for machine learning, because there are much fewer examples off the positive samples or the samples that you want to detect.

Then you have examples of other types of samples, i.e., normal traffic. In some senses, it may be more healthy to think about this as an anomaly detection problem or an outlier detection problem, rather than a classification problem in its more traditional sense. What we do is to rely on heuristics primarily, and these heuristics, we try to mold recommendations to what the customers are going to be looking for. We don't apply machine learning to our product in the most traditional senses. We're applying it in a different way to what most people expect we do.

We don't do fraud detection by trying to build a binary supervised classifier and detect fraud and not fraud. Instead, we think of this as a recommendation system problem and we keep ourselves agnostic of the type of rules that people are running on our systems, so that we can make recommendations to how they can improve these rules. If we find that they are flagging certain users that are performing some malicious activity that have triggered their rule sets before, and we see that there are other kinds of users that are similar in nature, but maybe haven't or wouldn't have triggered any of the rule sets in their systems. Then you recommend us to them. We recommend new ways of writing these rules so that they can flag these transactions. This is a way that I think people can really uncover the unknown unknowns and flag the false negatives.

[00:21:15] KP: Fraud is going to have a class imbalance. It's kind of rare and even then, I don't imagine you get a lot of labelled data, or maybe you do from your clients. Did you get much labeled data and can you make use of it or make use without it?

[00:21:30] CC: Yes, absolutely. Whatever labeled data that we have, whatever our customers can provide us is going to be helpful and very much welcome. Because we're dealing of a space where labeled data is so rare. We are doing all that we can to help our customers generate labeled data. Now, the reason we can do this is because customers don't just use us as a detection system, but also use us a review system. Whenever alerts are flagged within our system using our rules, it goes in queues that operators review one by one in our system. Now, after every review of an alert, they make a decision on whether it's alert that's flagged as relevant or not relevant. If it's relevant, how relevant is it? If it's irrelevant, how irrelevant is it?

This forms a natural feedback loop that we can build in into our models, so that we can help them improve their own custom models. Now, this is frequently rare. Before using a system like

Unit21, customers were very frequently using tools like Salesforce or Jira, ZenDesk, the very typical ticketing systems. These systems weren't built to be labeling tools. In a lot of cases, customers even devolve down into using spreadsheets, and Google sheets to drive investigations efforts. They would export their investigation findings from Google sheets and have them be ingested by their own data science team so that it can use them to improve their models for detection.

Now, all of this doesn't have to be as brittle as it used to be. In our system, they can define new labels to define – to add onto alerts after investigation, and use this either internally within our system or export this from our system and use it to train their own models. Many of our customers don't just rely on us for detection and we welcome this. The more that they build their own internal models, their own data science teams, building detection sets that are customized to them and integrate the alerts into our system, the better our alerts get, the better our rule systems get. Because the more signal we get around how good our detection systems are in comparison to theirs and we can improve.

[00:23:47] KP: You're working with N number of customers and similar spaces, is there any opportunity to federate the data to look across all of those and learn something in aggregate?

[00:23:59] CC: This is something that we've been ask for so many times for by our customers. It's really funny that even when we got our third or fourth customers, they were asking us this. Of course, back then, when we did the study into how meaningful it would be to federate our learning, we didn't find too much, because there needs to be some meaningful overlap or whatever the definition of the overlap means. Within your customer base, before you can build any meaningful consortiums. Now, today, we're at a much later stage compared to them when we have customers that span several of our use cases. When you see a single user logging into multiple of our different customers, many of whom have flagged them to be suspicious users. Then, we can notify our customers on this.

This is something that we're uniquely in the position to do so. No company can really build something like that without having to build the consortium themselves by going to all these other companies in the industry. We see this being most relevant in crypto. Of course, when people using crypto exchanges to launder money, to do surreptitious activity, they're only just going to

use a single crypto exchange. They're going to want to, by definition, if they're a good money launderer, use multiple of these systems. The broader view we have over all of these, the better we get at this.

Now, we haven't really started to do a lot of this yet because we think there's still a lot more room for us to grow before we can truly release a product that people will be willing to pay for here. But a lot of what we do really is on the detection modelling layer, when we find a crypto exchange that's using a specific rule set really successfully, then what we do within our product is to propose this as a rule set to other customers in the same industry. They can then leverage this same rule set and most likely be able to capture the same user who is going to be exhibiting the same behavior across the different crypto exchanges. This we find to be the most robust way to capture and share any of the insights that different exchangers, different industries of customers are learning from their customer base.

[00:26:13] KP: It stands to reason that if you do a new software release of some new innovation and detection could even be some prevention step that gets put in place that there could be a steer step event here where a bunch of fraud is suddenly shut down and those criminals need to go scramble and try and evade your new tactics. Are you able to observe any interesting dynamics when you do a release and see people respond to it now that the problem has gotten harder for them?

[00:26:40] CC: Yeah. This is really interesting. The economics of fraud, the economics of money laundering, I think deserves more study. There are some really interesting studies on this, where criminals that try to launder money online or commit fraud online frequently try to go after the lowest hanging fruit. If they're facing some kind of setbacks for the services they are used to, then most of the time, they won't give up. It costs more for them to innovate around our solutions than to go to another platform that maybe has a lower barrier for entry.

This is interesting because it means that it's an arms race. But because the platform of all different ways that money launderers can launder money or a fraudster can commit fraud is so, so broad. It would be highly unlikely for fraudsters to want to innovate because they are blocked on some platform out there. One example is, if someone is trying to launder money through a well-known crypto exchange that is doing a lot to block these attempts at laundering money,

then what they'll do is probably not to go a little bit deeper and to innovate around our solution. Because most fraudsters aren't the most technically enabled people. They have downloaded scripts, they have found ways of committing this kind of fraud and money laundering online, .net forums and you name it. They aren't necessary the ones who do too much more to avoid detection, to spoof anything that they needed to spoof to avoid being characterized. Instead, are going to go to a less well-known crypto exchange where they can hide their tracks a lot more effectively. The cat and mouse game is occurring, but because of the broad surface of the different ways that fraudsters and moneylenders can operate, this cat and mouse game is going to last a long, long time.

[00:28:44] KP: There's a public perception that cryptocurrencies are kind of the Wild West that they're haven for criminals to do illicit transactions. I guess that's plausible, but I haven't really looked into it. You're a lot closer to the data. Do you have any intuitive sense of the degree to which there is fraud and money laundering compared to typical commerce?

[00:29:06] CC: Yeah, crypto is definitely a really interesting space for money laundering. The thing that's interesting though, is that many fraudsters at this point I think still don't know what crypto really can bring to them. There is obviously a very clear confusion between privacy and anonymity, and crypto only provides for one of them. When we talk to some of the leading crypt exchangers around how they're attaining success in identifying money launderers and identifying surreptitious individuals that they are reporting to the governments, then we find that the general aptitude of money launderers and criminals using these platforms is still quite low.

Now, of course, there is a self-selecting element in there because there is the high possibility of the real high-aptitude money launderers and criminals not been detected. The crypto exchanges, the crypto networks do present a very fertile ground for them to operate. Because if they're able to efficiently mix and effectively hide their tracks using especially one of the privacy-preserving networks out there, then this is a real problem to the realm of money laundering. Even though there are techniques to probabilistically define if an account even in a privacy-preserving network is involved with any surreptitious activity. this is still much more difficult to do than in a public network. Like Bitcoin or Ethereum.

[00:30:48] KP: Money laundering is something that happens when people decide to start laundering money. I don't know if there's any trends or patterns there, but certainly, there could be spikes and unexpected events that you can't predict. To what degree do the operations of your company ever have to assemble them like a war-room type scenario?

[00:31:08] CC: The COVID situation was actually really interesting for many of our customers that had to deal with different attempts at fraud. When dealing with disbursements of PPP loans. or any kind of COVID-related loan programs, obviously, we saw tons of cases of identity theft. We saw cases where people were cleaning unemployment benefits for other people, that they were not. This was interesting. There were situations where we had to jump on board room calls of our customers to deal with this, because of course, these programs are great, but when they are hastily implemented and where there's not a lot of guidelines around responsibility and blowback, then this is troublesome. We had to work for our customers to ensure that they were giving out the loans in a timely fashion, but weren't giving them out to the wrong people, because of course, this would be catastrophic.

There are situations like that, COVID being one example. But in different times of the year, there is also deferring ways of fraud. Payment fraud for example becomes a lot more rampant at times when shopping is a lot more rampant. Towards the end of the year, Thanksgiving, Christmas, you see lots of online transactional activity. You also correspondingly see a lot of waves of secondary financial events online where people transfer money and give people money. This causes a lot of fraud. A lot of criminals know about these patterns and they commit fraud at times when there is much more of an increased volume within normal accounts. They use this to hide their tracks.

If you wanted to launder a larger amount of money, laundering it around Christmas time would be a great time to do so, because it's just a lot harder to predict the behavior of an account at times like this compared to normal times.

[00:33:16] KP: Does that impact your ability to take a vacation?

[00:33:21] CC: Yeah, in security and fraud, these are waves that people are used to. I used to work in security and also similarly, Christmas Day and Boxing Day was the time where there'll

be the most attacks. Of course, this is something that is very unwelcomed by security professionals everywhere but it happens. If you look back at many of the APT attacks, many of the ransomware attacks, they happen during holiday season. It's not fun, but it's also something that we know is going to be a pattern and being prepared for that allows us to at least remediate some of the work that we need to do during those times.

[00:34:01] KP: If mostly ask you about the transaction monitoring components, how do you help with case management?

[00:34:10] CC: Yeah. Case management is the investigation piece. Now, we use the term case management because a lot of our customers know what case management means. But what case management means to us is a little bit different from what the old school systems refer to by case management. When we talk about case management, we're talking about what to do after something is found, instead of just being, what to do to release yourself off the regulatory burden after you found something.

We find that a lot of our customers are not just interested in performing anti-money laundering detection for the sake of compliance. They actually care about their systems being clean. We believe that's the farsighted way of doing things. The more that you can guarantee to your users, and the regulators and the public sphere that your platform is a clean and legitimate way of transacting, the more you can gain traction and foothold off the system. A lot of what we help them to do really is to get to the bottom of everything that we flag and try to make sense of what's going on. We do so by presenting them with information that they need to make their decisions beyond just giving them all the transactions in a CSV format or table format. We give them visualizations, help them to make sense of data without having to dive in all the way.

A lot of these decisions are going to be snap judgments. They're aren't going to be situations by you can afford to spend hours and hours looking at transactions going on. This is naturally a challenging event because, I don't know about you, but when I look at my credit card statements at the end of the month, it's challenging to find transactions that I find anomalous. Even if it's my own traffic spending pattern. But if you're looking at someone else's credit card transactions or looking at someone else's transaction history, how do you make sense of this? How can you

quickly within 30 seconds figure out if something suspicious is going on here? That's challenging.

Our job over here with case management is to really help investigators get to a decision as quickly as possible. Some of this involves intelligent UI design, but some of it also involves some prediction. If we see that there are some patterns in here that we can help them to uncover before they can uncover themselves, then we'll do it. We'll give them suggestions on what they should maybe look at. Sometimes it's relevant, sometimes it's irrelevant. But from their dispositions on the alerts, we can get better.

[00:36:50] KP: It seems like there would be an urgency to take an action when things are detected. Are users fighting the clock?

[00:36:56] CC: There is. They are either fighting the clock because of the economic cost of reviews or because of the transactional flow. The former is the most common, because fraud detection teams and money laundering detection teams are frequently viewed as cost centers in the business. When you spend more time investigating a single alert, you have less time to investigate other alerts. None of our customers really have the problem of having too many resources. They're all dighting against the clock. They're all trying to get to a decision as quickly as they can, so that they can focus their time on the things that they should be focusing on, i.e., the more sophisticated types of fraud that require a little bit more analysis, a little bit more investigation. Instead of the ones that are false positives or very clearly fraudulent and they can block.

In the latter case where we're dealing of transactions, there's a much more clear financial outcome to this. When they have to make decisions to unlock transactions for example for the users, then they frequently only have one to two hours to find out if a transaction should be blocked or not. When they have a queue of 100 transactions in their backlog, then this is something that they have to visit the site on. Companies have to play around with de minimis thresholds for how large the transaction has to be before they spend analyst time on reviewing these. This is a very much operational economic decision that companies have to make.

[00:38:40] KP: It seems that in order to be thorough at investigating fraud and money laundering you have to get to the point where you're looking at at least a few false positives. Do you have to consider that in the design of your product, not overwhelming the user as you explore where the line is?

[00:38:55] CC: Of course, alert fatigue is real. We all know that you can effectively detect all that activity if you were to flag everything and lower the bar for what you consider to be suspicious. But of course, this just means that your operators have to spend a lot of time to look through events and to take pick out the needle in the haystack. What we want to try to do is to both reduce alert fatigue and maximize the rate of true positives. This is challenging because we're being pushed in from both ends here. We need to make sure that we're not over flagging some types of events and we're addressing in as agile a way as we can, so that customers will not have to deal with things that they already flagged away before, that are going to be false positives, are clear to us that it's going to be false positives and we can constantly change the thresholds in a way that allows them to inspect and audit the way that we're changing threshold. These manifest our product as recommendations on these rule sets, so our customers don't have to go through mental exercises to adjust thresholds here and can rely on us to do so.

[00:40:09] KP: Do you have any sense of how your customers utilize your services and their operations? Is there a particular KPI that you're helping them improve?

[00:40:19] CC: Yeah, absolutely. A lot of this is really interesting to us, because even though we're building the solution and we provided this solution to our customers, we're not involved to the day-to-day operations that our customers are using our tool for. It's only when we go back to them and run a user studies or run case studies of our users do we have any kind of visibility into the kinds of fraud they're catching, and how they measure the effectiveness of their program.

Now, with some customers, we found that just a couple of months of using our system, they can uncover hundreds of thousands of dollars in fraud, which more than pays for the solution itself. The agility that it will bring to them is a true winner here. Previously, because operations teams were disconnected from the detection systems, and they were only in charge of investigations, operations teams weren't able to make decisions like change thresholds, and change detection

patterns and do things like that. They have to rely on going to engineering teams, describe to them the problem and having this secondary thing that are involved with. When using our system, operations teams that know the problem of fraud and know the way that money laundering is changing within their system, can be the ones that directly run experiments and change detection logic, so that they own the problem. They own the problem from start to finish.

This is a gamechanger to them. It means that the barrier for running experiments and their ability to understand the data and play around with the data is hugely amplified. They don't longer have to rely on engineering teams, multiply week long sprint cycles, getting them to a queue of competing priorities and get what they needed to get done, done and can do it themselves. Not only do we help them with reducing resolution times of alerts, we also happened to be a lot more agile with alerts and do things that they would previously require SQL knowledge, database knowledge or data engineering effort to achieve.

[00:42:34] KP: Can you speak to some of the interesting things on your roadmap or maybe that you're considering for it?

[00:42:42] CC: That's a good question. There are a couple different angles to this problem. One is, what we learn from our own experience working with customers, and non-technical operations teams, who know their problem of fraud so well, but may not be the ones to write software to catch fraud. The second would be, what we expect to be some things within our pipeline for implantation in the future. For the first, what we've learned is that, when we first started operating, when we first started serving customers, we've realized that the hypothesis we had around how to enable operations teams was not entirely right. Building rule sets, building logic on data not only requires operators to be able to express logic, but also requires operators to understand data. It's understanding their own data, not our data because all the data ultimately comes from them.

This is challenging because if you wanted to teach someone to write SQL, that would be maybe a little bit of a lighter lift than teaching someone to understand databases. But operators have to understand databases, they had to understand the way that data was modeled within their systems, before they were able to very effectively leverage the data that they had and understand it so that they can be effective at uncovering data patterns for fraud and money

laundering. This is challenging. So recently, we're starting to work on projects that can help our users understand their own data. We're calling this data discovery operations. There is a lot of opportunity here where we can allow non-technical people to go up the stack and to really understand what it means for data engineering to occur.

Data engineering doesn't just involve detection signs. It also involves understanding data slicing and dicing the information you've collected from users and understanding how data is modeled. This is a difficult task and we see this to be even a bigger mountain to climb than they want to be originally started on. Because data is so much more sophisticated and so much more complicated, especially when we aren't the ones that designed the data schemes. When we have to go in and help our users understand the data that some other people in their company had designed, then we have to make the entire workflow, the entire user journey a lot more intuitive than it would be, and it has been. Because all of the data discovery, all of data platform systems that are designed and built today are built for engineers, they're built for data engineers, not really for people that may not even understand what a database is.

We look to some inspiration out there that have done some interesting work over here. Look around Tableau for example, have done some really interesting work with helping business analysts understand data sets. They both have different approaches to the problem, but we're leveraging some learning from those bases and helping our operators, helping our customers understand their data better through user interfaces that visualize patterns for them that use graph analysis to uncover anything that is potentially suspicious without necessarily having them understand every single thing around what a column or index means.

On the other side of things for detecting suspicious activity within the fraud and AML realm, what is going to be really interesting I think is around the crypto space, as there is more and more traction within privacy preserving coins. I see this as an inevitable thing that happened, that's going to happen. How do we deal with cases whereby the fundamental principles of crypto of anonymity really goes against what is needed to detect money laundering, detect fraudulent activity and to clean financial ecosystems? This is a problem that I think the entire crypto system, crypto community is also grappling with.

There are of course privacy maximalist that are going forward with the very liberal standpoint that crypto is – private crypto should be anonymous and the reason why crypto exists is for privacy. But then there's also practitioners that realize that without the blessings of governments, without being able to ensure, improve from first principles ideally without sacrificing privacy or anonymity that people transacting on the networks are not using it for anything that is harmful and bad for social systems. I think this is uncontroversial and this has to be addressed at some point in the future. There's been tons of research papers on the subject and we're watching closely, we're also working with partners like Genalysis, or just crypto forensics and working with them to make sure that our customers that are using our product to detect transactions in crypto are able to do a good job and this is something that we're keeping a close eye on.

[00:47:59] KP: What types of organization should be looking at your solution?

[00:48:05] CC: The type of company that should be looking at Unit21, when we first started out, was only FinTech companies. Our customers like Chime and Intuit, Coinbase, Gusto, Flywire, for example have attained a lot of success using our product. Honestly, to a lot of our surprise, we were able to provide them with enough flexibility and customizability to user software to detect things that we didn't even think of supporting when we first started. Now over time, we've also realized that customers are starting to use us for different types of use cases. Because what we've built essentially is not a fraud and money laundering detection software, it's general activity monitoring software. Because we don't tell customers what to detect, and we don't build fraud models per customer. Our customers really can use our system to build models that they want to build for detecting the things that they want to detect.

That's why we see companies like Twitter using our software for powering their blue check mark verification program. Anyone signing up for a blue check mark beside their Twitter username is now being verified by our systems. Twitter has hundreds of agents using our system going through alert queues and making decisions on them. The entire reason why this is possible is because none of what we've built is very customized to transactions. What we're really excited about for the next 12 to 18 months is to go into mini use cases. We think that instead of building an anti-money laundering solution, where instead building a modern infrastructure for activity monitoring, whether this activity is transactional, or whether it is user profiles online, or whether

its insurance claims, or accounts, or anything that you might like to detect any network system where no users are explicitly operating.

We think there's a use case here and we're starting to work with some design partners to double down in these use cases with the goal of being much broader here. Giving the tool that was given to compliance and risk teams in FinTechs to other operational teams that face the same problems, we see that there are dozens of use cases at least where operations teams that sometimes are in the dozens or hundreds within companies have to look through events happening within a stream of data and try to find suspicious events, try to find **[inaudible 00:50:50]** events, anything or anywhere where there's a pattern like this is a use case for us and we're really excited about it.

[END]