

Max Richman: Good morning. Happy Thursday morning. For me, it's Bow Tie Thursday, so hence the bow tie. I'm really glad we're having this session on Bow Tie Thursday. My name is Max Richman. I'm a Chief Data Scientist at a group called, GeoPoll. It does market research around the world in emerging markets developing countries, and they also lead a network of data scientists called DataKind who lend their skills, and their time to organizations looking to do more with data science.

[00:00:30]

I am extremely and feel extremely lucky to be here today with, who I really do believe is learning from the best. We've got experts. People who have started, innovated and created new businesses and new organizations on three different continents, and I'm really excited to welcome this panel here today to learn from them, hear from them about the projects that they're doing and how they're incorporating machine learning and new technology into the work that they're doing to advance global development. And I'm also really excited to have a discussion with them to talk a little bit about what's some of the opportunities that they see, what's some of the challenges they see, and also hear a bit from the audience about some comments, and some thoughts as well.

[00:01:00]

So, just to give people a rough rundown of what we're going to be looking to do for the next hour. I'm going to sort of introduce the topic and introduce the panelists. Then we're going to have our three panelists do sort of quick 5-minute lightning talks, so that you can get a real sense, deep dive, into what they're doing on each of these projects. After that, we're going to sort of have a discussion among the panelists here that I'll facilitate, covering a number of topics that come up during their presentation as well as looking at what other domains we could think to be using some of these tools and technologies, and what are some of the challenges that we're seeing in this space as well. And then finally, I'll grab a couple of questions from the audience, and we'll have the panel go through those questions and then we'll kind of do a little wrap-up with each of the panelists.

[00:01:30]

[00:02:00]

So, that's what we're looking to do for the next hour. If you were expecting something else, it's okay, you can leave. Seems like people are coming in, but I'm detecting some smiles from the audience, so I think people are in the right place. All right. Well, let me kick off. The title of this session is "Learning from the best: The potential for machine learning innovations in international development". There's a lot of buzzwords and loaded words in that sentence, so let me try to translate it from my understanding and for what we're trying to do, and then hand it over a little bit.

[00:02:30]

So, learning from the best. I take that to mean let's look at some case studies from around the world. We've got case studies from Kenya, some case studies from Mexico and case studies from South America. And I'm really curious to sort of see how the best in practice are using these new tools and technologies. So, what do we mean by the potential? It means there's still some prospect. We're not fully saturated with hype yet. There's still a lot of potential for these tools to do real good. What do we mean by machine learning? Well, we're going to be defining that more throughout the course of the day and through the course of this session, but I

[00:03:00]

take that to mean using certain amounts of data that you have and trying to predict things that you don't have in the data, so trying to forecast or trying to model or trying to come up with some best guess, with machine support, about what's going to happen and what that means.

[00:03:30] So, we're going to try to keep this at a level. We have very technical people here who are the CTO's of organizations and I'm going to buzz them if I hear buzzwords that we don't know, but I want to make sure this is accessible to people who might be machine learning experts themselves to people who are more policy or more subject matter experts as well. And then, what do we mean by international development? Well, I'll leave that for some of the other sessions, but I think we all have a sense of what that means. So, I guess without further ado, we'll move into the next section where I have the presenters speak.

[00:04:00] First, let me introduce them. To my left here, we have Peris Bosire, one of the co-founders of FarmDrive. FarmDrive is using machine learning to develop a credit assessment model, which will help financial institutions understand the risk profiles of smallholder farmers. Using alternative data such as agronomics, satellite and market data, they can assess the credit risk of farmers on their platform and develop tailored financial products to manage their individual risks. So, welcome Peris.

[00:04:30] To my left a bit further, we've got Karlo Valetin Rodriguez, the Chief Technical Officer of Grillo. Is it Grillo? Is that right? Got my double "L" going. All right, cool. Grillo has developed a network of seismic sensors in Mexico that collect seismic data and uses proprietary algorithms to identify earthquakes. That sounds pretty timely. Dedicated alarms devices and a free mobile application provide early warnings of incoming earthquakes. The existing Mexican government early warning system is extremely expensive in terms of infrastructure investment and has

[00:05:00] limited coverage. So, this device provides affordable access to early earthquake warning.

And third, to the left there, we have Dave Grenell. Dave is the co-founder of the Rainforest Connection or RFCx. Dave brings ten years of environmental policy leadership in creating national models, programs and high impact pilot projects with an emphasis on technology, crime reduction and environment. Dave leads the strategic initiatives that can scale through partnerships, programs, technology and green economics.

[00:05:30] That's enough blabbing around from me. I really want to hear, as I imagine you do, from the best here in the room. So, I'm really excited to welcome Karlo to present. So, Karlo let me hand you this clicker and you can talk a little bit about what you're doing.

Karlo Rodriguez: Hello. It's fine? Okay.

[00:06:00] Thank you for coming. Thank you for the invitation. I was afraid that nobody will show up because this is very, very geeky. So, what we do at Grillo, as Max said, we

work in Mexico City. We are looking for global expansion, but earthquakes is a global problem. Have you heard the news that last week we were hit by a very large earthquake? 40 buildings in Mexico City collapsed and 3,000 buildings are in the way to collapse. They're going to be demolish it. 350 people died. There's cities around Mexico City like are completely destroyed, but literally destroyed. There's no building that's standing up.

[00:06:30] So, the current early warning system in Mexico is really whacked. Three out of four times they fail to detect an earthquake and they fail to provide an alert and they have a lot of false alert, so people don't believe in that system. And it's really, really, really expensive. I'm talking about in the range of tens of millions of dollars per year and it's really bad. So, me and my panel saw that okay this is not impossible. So, we decide to work on that field. We have prior experience of internal things that the science had to relief. One of my partners worked with a client foundation in the IT earthquake five years ago, six years ago.

[00:07:00] So, we decided to built a new system. How this system works, we take a [inaudible 00:07:07] of the latest trends of internal things plus imaginary and communications. So, we deploy a network of sensors in the coast in Mexico. So, we detect ... One thing I want to make clear. Earthquakes? You can not predict earthquakes. If you read something of the prediction of earthquakes, it's a fake news and you block and report. Seriously, seriously. There's people working, analyzing many sources of data, your physical data, but right now you can not predict earthquakes. That's a lie, but you can detect earthquakes when they happen.

[00:07:30] So, when you detect an earthquake and you have these sensors of networks, you can analyze the data that is coming. Detect that there really is an earthquake. Detect the magnitude of the earthquake and you can project where this earthquake is going ... how far it's going to travel and if it's going to impact a population center. Based on that, you can admit an alert to that population center. In our case, we are able to detect the earthquake and admit an alert up to 90 seconds before the earthquake arrives to Mexico City. That's a lot of time, so you take cover, you can go to other building. That's the problem that happened in the latest earthquakes. The alert was submitted manually by the government when the earthquake already was in the Mexico City. People didn't have any time to run.

[00:08:00] So, our tech solution, because 1% of the current cost of the existing solutions and we provide alerts to the people in a free way through an app. It's free for the population. We got a social company with we need to find a way to keep the lights on. So, our recent model is we charge for that access to an API to the induce route to the enterprise, so they can have access to the alerts and they stop elevators, close gas bulbs. So, for the industry most of the application in that.

That's what we do at Grillo. Love to help, to be here.

[00:09:00] Max Richman: Thank you, Karlo. Great. Next, we'll welcome Peris. So, I can grab the microphone.

Thank you, Karlo. Welcome Peris. Caribou to Washington DC.

Peris Bosire: My name is Peris Bosire. I'm from Kenya and I'm one of the founders of a company that's called FarmDrive that's working at the intersection of agriculture, finance, and technology. The work that we do as FarmDrive is using data science and data analytics to drive more capital to smallholder farmers. As you might all be aware, there's a big funding gap to agriculture. Globally, it's about a \$450 billion funding gap and when you zero it down to Africa, most of the people are employed in agriculture, and most of them are actually smallholder farmers, but the traditional ways of assessing risk that banks use do not really work for this kind of population.

[00:09:30]

[00:10:00]

The banks look at credit history, they look at collateral, but most of the farmers are not banked, and they do not have collateral. They do not have anything to use as security to get a loan, but these farmers actually have identities. There's a lot of data out there about these farmers, especially because in countries like Kenya and in most of the countries in Sub-Saharan Africa, ABSA drove the use of mobile money. A lot of people are now starting to have digital identities because of the use of mobile money. When you look at agriculture in itself, this is a sector whose risk has to be looked at from, especially at production level, from the agronomical aspect that affect productivity of our farmer.

[00:10:30]

[00:11:00]

You have to look at things like weather, soil, market trends, how is that going to affect the risk, and when you bring all these data sets together, then it becomes easier to start extending credit to these smallholder farmers who need the loans. Most of the farmers where I come from, in my country Kenya and in other parts of the country, especially the small farmers, they have less than 2.5 acres of land. So, we're talking about very small microloans to help them buy inputs, and then moving them from a scale of A to B, to more precision agriculture in that level.

[00:11:30]

We're using, what Max has described as, machine learning to collect all these data that's being generated, data from the telcos, whether soil data. We're looking at automated Know-Your-Customer kind of identities from the farmers and bringing all this together to generate risk profiles that financial institutions can use to then extend credit to these farmers. We've been doing this for the last two years and advancing our model and we've seen the success of the work that we're doing.

[00:12:00]

The first micro finance institution that we worked with increased their agriculture portfolio itself by 20% when I started out working with them initially, and we're just getting into another phase where we'll be working with one of the largest telcos in Kenya to provide this at scale to all the smallholder farmers in Kenya and in Africa. The alternative credit scoring models have been tried and tested in different ways. In my country, Kenya, we have about 20, or more than 20, digital lenders that are looking at different kinds of data. Some of them are scraping things like social media data to extend credit to most of this excluded segments like the youth, most of the students who don't have credit histories yet.

[00:12:30]

So, we believe and we've seen the results that thinking different about ways of assessing risk for different sectors is the way to go to make finance more inclusive

and expand access to private financing to people that need it most. We'll talk more about how this works as we go on. Thank you.

[00:13:00]

Max Richman: Great. Thank you, Peris. And thank you Moses for helping switch to the next one. Next we have Dave.

Dave Grenell:

Hi. I'm Dave Grenell. I'm a co-founder at Rainforest Connection and what we do is we take ... We basically leverage cellphone technology to stream all the sounds from rainforests into the cloud where we can search for whatever you want to look for. So, we're starting with chainsaws, trucks, vehicles, things mostly associated with illegal logging, because that's our mission. To enable partners to get realtime data that they can receive alerts of a chainsaw in a remote place and show up to either stop the activity or to call someone who will.

[00:13:30]

[00:14:00]

So, these are the solar panels. This is what it looks like in a different format with a partner going up a tree, and this is basically the box, which can also operate like a camera trap and we've got maybe 45 of these deployed around the world. We're in currently Ecuador, Peru and Brazil where we're working with indigenous tribes. So, our view is that rainforests are the greatest libraries of life science on earth, not including the oceans. What we can do with all that data is we can make it available to the world's research communities, ecologists, land managers, climatologists, entomologists. Imagine what you could do with this remote resource that very few people have access to. It's very expensive to travel there. It's sometimes a difficult environment.

[00:14:30]

[00:15:00]

If you can create that into a data library that's incredibly shareable on an internet platform, which is what we're doing. We think the potentials for comparisons month-over-month, even if cricket populations year-over-year seasonally, can tell you incredible amounts of information about our planet, and there's a critical need at this moment as you know, because deforestation is a really big part of the climate change equation. We're experiencing one of the largest species extinction crisis's since the time of the dinosaurs. Not only that, but in terms of human populations, many poor communities rely on rainforests for their sustenance. Drought in São Paulo, which affects 20 million people is a direct result of destruction of the Amazon rainforest.

[00:15:30]

So, this is what a device looks like up in a tree and our role is really to assist. So, we want to get the data to the people who can use it and take the right actions. How we do that is we use AI and machine learning to develop the profiles for the different sounds we're looking for, and currently those are mostly related to illegal activity, although we've for the first time experimented successfully with detecting rare species of parrots that are being reintroduced to a particular ecosystem in Ecuador. This is kind of what it looks like. Basically the sound comes in. If it's a chainsaw or a vehicle, you might have a 1-kilometer radius. It streams up to the API. The software will basically begin to run a spectrogram analysis and we'll use a neural network to increase the probability of accurate detection, and then send an alert to partners on the ground.

[00:16:00]

[00:16:30] Another thing that happens is we create month-over-month reports of activities, so people can see the patterns of activity. You might find out that at 2:00 in the morning in a remote place in Africa, as we did, that's when hunters go out because there's no one else out there. So, after your ranger, lets say we're working with indigenous tribes in Brazil, you get an alert, and what you do is you go here and you can download the sound file to check is that really a chainsaw or is it a bug that sounds like a chainsaw? So, this is actually also a part of the machine learning process is over time human verification kind of helps us get closer and closer to higher accuracy rates.

[00:17:00] A lot of people getting an alert for an activity, lets say it's a chainsaw or a vehicle late at night, and they do not want to like jump on their horse and go into a dangerous situation. So, in many cases the best thing we can do is provide the overall pattern of activity, and that's also again we're in an advanced platform, we think there's a lot of analysis potential. So, how this actually works is you might have an indigenous person who is seeing ten large scale illegal logging trucks coming out of their reserve. They call someone. That information is easy to ignore in the same way that if a gunshot goes off in Oakland, California in a city I used to work at, you call the police. The police want to know well, where is it. So, people don't necessarily respond to information.

[00:17:30] The idea here is to create a report that creates pressure and transparency. It can be shared with different partners, NGO's, law enforcement, governments, where people say okay, we have the information, now what are we doing with it. This is the outcome that we're looking for and this actually happen sometimes where you'll get an alert and there'll be the three weeks it takes to built a road into a remote rainforest, so that you can clear the timber. That's a 3-week window where if you get the right kinds of activities, you can actually stop it before it happens. So, our partners in Brazil are working against really large scale illegal logging that looks like this.

[00:18:00] So, the other part of the first slide I didn't mention is how do you pay for scaling this kind of technology, and the answer is carbon funds. In fact, I think the best answer I've ever seen to how we save the world's forest is carbon funds. What I mean by that is REDD+ programs, which are coming online now in the voluntary market. The compliance market, which is mandatory starts in 2020 through the Paris Accords in the signatories. We are working in region number 8 in Brazil with this tribe and we are leading an initiative to create a carbon fund.

[00:18:30] Basically, step one protect the forest, step two go through the REDD+ verification process to sell carbon credits to companies and corporations and governments around the world. Our partners, Wildlife Works, are some of the number one carbon product developers in the world and they have sold \$23 million of carbon credits since 2015, and so that's our end game for how we scale this technology. Thank you very much.

[00:19:00] Max Richman: Thank you so much. Well, you can see why we have really quite a range of good

geographies and also quite a range of different applications of machine learning intended technology. I want to ask a couple of questions here of our panelists to kind of kick off some of the discussion here.

[00:20:00] So, let me start first a little bit more, since machine learning is here in the title, a bit more on the technical side. I heard a term like neural networks and I'm hoping whoever used that term can define it a little bit more. I'm really curious since sort of you walking me through three things in a bit more detail than we showed in the slide. Number one is the data piece, where does it come from? Sort of walk me through the pipeline. We heard that it seems to come from some sensors, whether it's earth sensors sensing the earth or sensing people or sensing people's activities on the earth. Talk a little bit more about your decisions to ascribe data for some things.

[00:20:30] Talk to me a little bit more about the analysis. So, step two. We heard about how you're doing some modeling but clearly you're also doing some outlier detection. You're making sure that's not a cricket. You're making sure that's not just a bumpy train going by or something. Talk to me a little bit about, more technically, about sort of some of the analysis and the human verification. And then, talk to me a little bit of sort of, I think what Peris had mentioned as well, how do you make it useful to someone. So, talk to me a little bit about how you then take, after you have that black box of mystery, and how do you make that useful.

[00:21:00] So, maybe we'll first start with the data piece. Talk to me a little bit more about that. Then we'll talk about the analysis. Then we'll talk about making it useful. So, I guess, maybe Peris, I'll pass it first over to you. Talk to me a little bit about we saw your chart with all the different data sources. What does it mean to really tap into those data sources?

Peris Bosire: The most important aspect of our work is using alternative data. So, looking at risk assessment differently from how banks are doing it, which is what makes it not so

[00:21:30] easy for smallholder farmers to get finance. Beyond credit histories, beyond collateral, what else can you use to assess the risk? When you think about credit risk assessment, it still boils down to the unit economics of a bank has to make a loan and they have to make money out of that loan, when you look at it from a banking perspective. And for the smallholder farmer obviously, it's more I need to

[00:22:00] get a loan when I need it. I need to get the right product that I need at the right time.

[00:22:30] So, how does all this ... How do we look at alternative data and how does it all come together to predict the risk of the small farmers. The first thing that we look at is what are the components of credit that the financial institutions look at? They look at things first like they look at capacity, they look at character, they look at the conditions of making that credit available and the four C's of credit. Character, what, what ... I've forgotten the other one. So, this are the kind of different bands that financial institutions look at when trying to assess the risk. And we don't diverge differently from that because it's the fundamentals of extending credit.

[00:23:00] What we do differently is okay, we do not have credit history for smallholder farmers. What other data can we use for this process to understand their capacity and their conditions? The first thing that we look at is Know-Your-Customer information. We call it KYC in short form. The reason why this is important is because when financial institutions are making loans, they're required by law globally, and in my country as well, to verify the identities of the people that you're making loans to. When you're doing it digitally, when you're using mobile phones to make these loans, then you have to figure out a way to identify the people that you're giving loans to in a digital way. How we get that data, especially we need to know the national identities of these farmers. We need to know their names. We need to at least know their location to a very accurate level before we can help these financial institutions give them loans.

[00:24:00] We capture this data from the farmers themselves through an SMS kind of survey that we have developed. Most of the farmers do not have access to smartphones, but they have access to phones. There's a lot more phones in most of the developing markets than even toothbrushes and toilets you could say. So, they have friends and they can type on SMS. So, we have a simple survey, but that asks them: What's your name? What's your ID number and where are you located?
[00:24:30] We capture that at the lowest administrative level.

And we are connected to the government databases that has the registries of this person. So, immediately we get that ID number and the name of this farmer, we are able to verify their identity and actually confirm that yes, this person exists in our government registry. This is their name. This is their ID number and from that you can even start picking out things like their sex, I don't know if you understand sex but I think it's sex, that's the appropriate one, their sex and the age of this person. When you get the location of this farmer, at least to the lowest administrative level, in my country it's at ward level, which is a couple of kilometers. We then ask them a few questions because they don't have GPS-enabled phones.

[00:25:00] What's the nearest primary school to your place? Most of the farmers know the primary schools. It's a point of interest for most of these farmers. Once you get that location, then you're able to start mapping their nearest macro climate zone or ecological zone that they belong to. When you have that data, you can then start making inferences. If this farmer is located in this area, what kind of soil do they have? What are the weather patterns in this kind of location? So, already with those different kinds of data sets, you've identified the identity of this person, you've done that already. You understand the risks as associated to the environment where this farmer is located.

[00:26:00] We then plug into the telco databases and the reason why that's important is because most of the low income and also most of the people working in the informal sector in the African markets now have access to mobile funds. They have a mobile money account. At every single point, every single day in my country there's at least 5 million loans outstanding in form of digital loans from different providers. So, these farmers have some kind of data that's being generated on this.

We then plug it into those databases as well to try and see how are these farmers transacting on mobile money? How much money do they receive? How much money is going out? That kind of helps you to start making inferences on the different cash flows and the risk bonds of these smallholder farmers.

Max Richman: [00:27:00] Can I ask you a question about that one? Are the telcos pretty well established? Is this at a service or have you had to really teach me them to get what you want from them? How is it sort of you pass, obviously safely and securely, information there and you already know what you're going to get back? Or is it more of a prolonging engagement with a secondary or third party groups that you're working with?

Peris Bosire: [00:27:30] Actually, I mean, Kenya is pretty much advanced when it comes to mobile money and the other services that can be derived at from mobile money. So, we have telco provider like Safaricom, who already have developed this credit scoring kind of mobile usage statistics and they make them available to developers of different applications, financial services. So, that's already existing in the market, and it even goes farther that Safaricom themselves have tried to use data, this kind of data, in cooperation with other financial institutions before to create other digital products. [00:28:00] In Kenya, we have products between Safaricom and other banks like M-Shwari, for those who know the Kenyan market, and KCB M-PESA, which is another financial product that's been already proven and been up taken by about 15 million clients who didn't have access to credit before.

[00:28:30] So, the mobile usage statistics are kind of proven predictive data when it comes to providing loans, and the good thing about the Kenyan market is detail, because of made these data available to developers or companies like FarmDrive to plug into and make it valuable. What makes FarmDrive different is that beyond the layer of mobile money data, for agriculture itself, things like soil, things like weather become very critical. Behavior analysis of this farmer. How long have they been farming? What's their size of land? That becomes very important when you're trying to design a loan product that's targeted for smallholder agriculture, and not just a payday loan, which will be consumed by the farmer in different ways.

[00:29:00] Max Richman: Great. Well, thank you, Peris. As you can see, especially on the fintech side, on the social side, there's a tremendous amount of different data sources that are available and increasingly able to be integrated together.

[00:29:30] To move just briefly to the other piece. I want to ask more of the internet of things people on the other side. We've got earthquakes. We've got rainforests. Internet of things. Talk to me a little bit more about that second level, the analysis, the modeling and the human verification. What goes into your black box? How are you detecting these outliers? Tell me a little more on the technical side about what's happening there? Maybe we can start with you first, Karlo.

Karlo Rodriguez: Okay. Well, in for as we're working in earthquakes, earthquakes kill people. That's our main problem, and the thing is that we can not wait until the earthquake have

[00:30:00] finished to say, "Ah, there was an earthquake." We need to do it in realtime, really, really, really fast. So, our problem is that when we get the data, we can not wait too much to say ah, it's an earthquake. So, all our sensor network, all our algorithms are designed to have some kind of flags. When a sensor detect some kind of a special movement, the other ones around are on standby, and if they detect some kind of patterns, they already know what type of patterns they should detect, then we're raising those flags. When we achieve three or four flags and certain conditions, we say oh, there's something going on.

[00:30:30] We up to a second level, then we analyze the data and we analyze windows of data, really short, like two or three seconds of data. We can expand it. If we expand it, we're going to be more accurate, but we can not wait because each second is counting. Each second is a life that somebody can go out.

Max Richman: You had mentioned that the governments were failing three out of four times. How is it looking on your devices? Obviously, nothing's perfect, but how are you measuring progress?

[00:31:00] Karlo Rodriguez: The problem is that the system in the government is like ... We take a multi-sensor approach and the government is taking a single sensor approach because also, they are using really, really old technology. So, the difference is mostly that, that we use a multi-sensor approach, and the way is that in the way we do it, we analyze the data, not in the device, in the cloud. So, that's how we are able to be more accurate and also, we record that data, not in an analog way, in a digital way. And after the earthquake, we include that data to improve our algorithms.

[00:31:30] Max Richman: Thanks, Karlo. Maybe a question over to our friend with South-American rainforests, Mr. Dave. You'd mentioned it also goes in the cloud and some neural network stuff happens. What does that mean and then how does that translate into action in the real world?

[00:32:00] Dave Grenell: What's most interesting to me is basically the predictive modeling ability of taking large sets of data and finding patterns, which meres a predictive ability of human beings. So, an artificial intelligence capability to do that. Like Karlo, we're putting all the information, in our case data sound, in the cloud. Then we ... Things we're looking for, we turn into pictures, spectrograms, which are frequency time graphs. Each sound has a fingerprint, right? And you create that fingerprint and now you're looking for that fingerprint amongst all your training data, but one of the problems with machine learning that lots of people find is to get good accurate probabilities, you need a lot of data, and so, actually this is one reason I think that financial markets do really well with this stuff in quints, looking for patterns of activities because there's a lot of financial data that people can make available.

[00:32:30] Max Richman: Thank you. I'm going to stay with you, Dave, and we're going to go back around here a little bit.

I heard, when it was great, because I didn't have to prompt it, each of the panelists mentioned during their presentation, touched a little bit on the business side. And

[00:33:00] you really can't ignore that when it comes to cloud-based solutions, data science projects, machine learning because you can do a lot more with a smaller team, but at the end of the day somebody's still got to pay for that Amazon web services. Someone's still got to pay to keep the lights on, as you said, Karlo. So, I think it was really interesting to hear the three different models and I want to go a little bit more into that because it is an important consideration for machine learning work.

So Dave, you had mentioned that you guys are tapping into the REDD scheme. How quickly does that sort of pick up? And then I'll ask similar questions of the other panelists. Or how does that work in the interim or is that more of a longterm solution?

Dave Grenell: [00:33:30] It doesn't work quickly enough. It is working today. By 2020, when the mandatory signatories of the Paris Accords ... There's five years to kind of people are able to buy credits to fulfill certain mandates that they've set. By 2020, it becomes mandatory. Countries like Japan, Germany, China will have to do certain things, but the deeper question, I think, is if data has value and data has monetary value and people are willing to pay for data, how do you use that, in our case, to translate data from rainforests into a value that competes with the short term destructive values of illegal logging and mining? If the value of the data can be shared and valued by universities, the public, other people, that creates attention awareness, which drives behavior and all of a sudden you have something called the spotlight effect.

[00:34:00] Where we want to go is how do we make the data more valuable than the destructive practices, and once you've valued the data, we're already seeing this. Our partners, often indigenous tribes, take a renewed interest in the values that they have, which are protecting the rainforest because all of a sudden there's a market or someone else is valuing that, whereas before people only valued it in a destructive sense. Can I take this tree and turn it into a house?

[00:34:30] Max Richman: Thanks, Dave. So Karlo, similar question. You had mentioned that the government has sort of been doing it their way. It's a bit slower. It might not be as fast. And that you're finding some of the early adopters are industries, people who manage elevators, people who have industrial production who are at real risk without shutting things down, hitting an emergency switch. Tell me more about how you're finding that product market fit with the industry.

[00:35:00] Karlo Rodriguez: Well first, we tried to approach the government, but we found out that there was not a way. They spend like 20 millions last year just trying to expand the network and they found out that a couple of guys in the basement are making all data in basement obsolete. So, that's why at the moment we approached, they rejected because we were making them look stupid, and they are stupid in that some way.

Max Richman: This is being recorded by the way.

Karlo Rodriguez: I know, I know. I don't care. I don't care. They were responsible for a loss of life of

our people, so I don't care.

Max Richman: No, fair point.

[00:35:30]

Karlo Rodriguez: So, in that way the Mexican government is not going to be a good customer in our case, and also the level of people is like an insurance. You're buying an insurance, always hoping to never use it, but how much value you will pay or ... What will be the value of alerts of an earthquake? So, we found that the way to do it is people adapting our system, like giving it for free to regular folks, but the people that is losing money, in this case the corporations, are like sponsoring. Well, we charge them, but the development of our network and they are going to be paying to have all this around. That's the way we fund it. It's really easy to a person to a corporation because you can add a value. Okay, if an earthquake happens and you lost elevators, you lost machinery, you lost anything, you can add a value to that.

[00:36:00]

So you say, if you don't pay this alert that cost this monthly fee, you are going to maybe be in the risk of losing this. That's our way.

[00:36:30]

Max Richman: Thanks Karlo. So, maybe we can pass the microphone. I can help here. Oh, you got it? And so, I guess, yeah-

Peris Bosire: Yeah, I have a slide on that.

Max Richman: Great. I was just going to yeah, prompt you as well because we did hear you mentioned you were working with obviously a lot of stakeholders, people who would generally want to give a loan to somebody who they don't have a lot of information about. Is that your customer or who is your customer with this information?

Peris Bosire: Yeah, so our clients are financial service providers. They're the ones who have the intention to extend credit, and the biggest pain of most financial institutions is how do I reduce my cost of acquisition of clients of providing my service to customers, and how do I improve my bottom line? How do I get a quality portfolio? That's the work that FarmDrive does. So, it's a very direct value proposition to the financial institutions. With a digital way of acquiring clients, you're bound to reduce your costs tremendously. When you look at some of these figures in the market, as I mentioned, there's 5 million mobile loans in Kenya every single day at any single point, there's 5 million loans out there.

[00:37:00]

[00:37:30]

Some of the financial institutions, from the initial work that we did, I mentioned that we increased the portfolio of that microfinance institution by 20%. Just very few numbers. That's one of the biggest values that financial institutions derive from our platform. The other one is the improved portfolio, the bottom line looking good. Better ways of assessing risk means more profits for this financial institutions because any loan that they have to write off is a loss in their books. So, if you provide them with an accurate way of making sure that every loan that I'm giving

[00:38:00]

out, I'm getting my money back, they're bound to buy into that.

[00:38:30] So, our revenue model is split into different ways that we've seen financial institutions in the market adopt these solutions. The first one is a credit scoring fee. So, they pay patron such and by a score that they look at or proforma that they have to run through our model, then they pay a fee for that. The other revenue stream, that again has been tested and tried in the market, is a revenue share on profit. So, for every profit or every money that comes back from the clients, you split a revenue share, either 50/50, 80/20 depending on the amount of work that goes on there to extend that credit. And the other one is more of a subscription of this software as a service. This means we provide the FarmDrive scoring model as a white label service to the financial institutions. They plug it into their existing operations and they pay a fee for every time or for how long, the duration that they'll be using the model.

[00:39:30] So, these are the different kinds of potential or possible ways of charging financial institutions for this service. We've already tested the transaction fee, the credit scoring fee, to the financial institutions. It's what financial institutions actually do currently because they have to pay fees to the credit reference bureau. So, it's very straight forward. The revenue share on profit, this has been tested again in Kenya, especially when you bring in partnerships between fintech's and banks and I think that's one of the things that is going to accelerate the financial markets, the partnerships between the fintech's, because of their technical capabilities, and the banks who have their lines of credit and their spending power. This is one of the most favorable models when it comes to charging or making revenues. Yeah.

[00:40:00] Max Richman: Great. Thank you Peris. That kind of wraps the first section here where I really wanted to get a better sense of the case studies. I wanted to get a better sense of how their data to analysis to impact pipelines look like, and wanted to get a little bit of sense of the business side.

[00:40:30] We're going to shift now into a second section here, which will be I want to talk a little bit more about these domains that machine learning and this technologies are being applied and other domains that we can think about. And then, the third section we'll get into afterwards are some of the things that I'm sure are on a lot of people's minds and we'll come back to that.

[00:41:00] But what are some of the challenges, particularly around the ethics and security and privacy and these things as well? So, going to section two. Again, maybe starting here with you, Peris. You've got your hands in a lot of different sectors. You're looking at environmental data. You're looking at agricultural data. You're looking at telco data. What are some of the other sectors and the other areas where you're seeing machine learning and solutions like that FarmDrive is bringing to the table? Where are some of the other places in other domains where this is happening in sort of the global development space that you're seeing?

Peris Bosire: Thank you. So, our domain as we keep reiterating and telling people, our domain is really agriculture because at the end of the day, what we are targeting or the

[00:41:30] motivation behind the work that we do, is smallholder farms, increase their productivity, increase their incomes. So, we've seen the application of this in helping financial service providers solve the agriculture domain best. That's where FarmDrive works. We've seen a lot of work again in machine learning and the use of data in education, especially in my country. In countries where you have a very big gap between the students and teacher population, then the use of things like mobile phones to provide content to pupils, to provide content to learners becomes one of the ways of reaching more of these students and also reducing the costs of government spending on education.

[00:42:00]

[00:42:30] We have a lot of startups that I've seen in my country that are using machine learning, that are using data science in diluting the content that's in textbooks into simple SMS surveys that are sent to students and they can use that to revise and prepare for the examinations, and there have been a lot of success stories, not just in my country, but in Africa. The reason why these tools have been successful is because of number one, the power of the mobile phones because when you have a mobile phone, it provides you a wider channel to access different kinds of services and information. And number two, it's because of the partnerships that have been enabled by startups like ourselves and the mobile network operators that have access to these distribution channels.

[00:43:00]

Max Richman: Thanks Peris. That's really great. I think you touched on two really important themes and I would like kind of your guys' reactions to those themes as well as you answer this question, which is: Do you find that mobile technology, like how is that enabling this revolution in machine learning, and also how critical are partnerships to expanding the usage in your domain and in other domains that you see?

[00:43:30]

Karlo Rodriguez: Well, mobile has revolutionized many industries, not just the data science and services. In our case it's an enabler, not in the way as a cellphone, in the way of how we can connect our sensors to the cloud because the same technology that provides the communications through cellphones and many other things, this is the same technology that we are using. Without that, we will not be able to have even the company.

[00:44:00]

Max Richman: Even in the very far inland as the coast? Tell me about how you guys get data from the coast?

Karlo Rodriguez: Okay. I'm going to get a bit technical about how an earthquake works. So, when an earthquake happens, usually it's in a fault place when they intersect. Sometimes in the middle of the plate, but the thing is 95% of the earthquakes happen in this intersection. In the case of Mexico it's in the coast and in the case of many countries that are susceptible to earthquakes, the famous ring of fire. They close all around the Pacific. So, we put our sensors in these places. So, we have to have a huge network, I mean huge by a lot of sensors, so you can detect the earthquake when it happens. They're really, really, really far away in the coast. The creative way we found to solve the problem of the location of the sensors is that in the Mexican coast, the beaches are beautiful. You should go.

[00:44:30]

[00:45:00] There's many places, many, many, many, many hostels, hotels. They're really far away from the civilization, but the tourists go to these places and the tourists are addicted to Wifi. So, they have to have cellular connections. So, we make a deal with these hotels, with this place, and say okay, please put up one of our sensors and we provide to you the alert for free, so add up one of our sensors. That's the way we are able to deploy, that's why we are so cheap. Instead of creating our own network and do some other technologies.

[00:45:30] We also have deployments in really, really far away areas and we use some other technologies that also are internet things based.

Max Richman: And what are some of the other domains then, as you're working in this, that you see that these tools and technologies can also have some benefit in Mexico?

[00:46:00] Karlo Rodriguez: Well, not just in the alert, in the industry that we can provide the alerts to the industry, we also have applications in mining and real estate. The real estate is really interesting because our sensors are really cheap and easy to deploy. We are not only measuring how the earthquake happens at the origin, we also measure how the earthquakes arrives for example to Mexico City. We are also deploying our sensor network in the city, so we can analyze how the earthquake arrives and we can create maps. We call it shake maps. A street by a street, we classify by colors and for example, you live in that red district. It means that when the earthquake arrives, you're going to feel it really strong in difference with a green street, which you are not going to feel it so strong.

[00:46:30] So, the real estate is also beneficiary of the that because they can see if the building should be more stronger than in other way and also the insurance, the insurance to know about the fee they should charge.

Max Richman: Thanks. Similar question to Dave then as well. So, I guess, it's now a two part question, but number one talk a little bit about the themes that I'm hearing from both of the speakers about mobile or maybe more micro computing, let's call it, and the importance of partnerships and then talk about some of the other domains that you see these tools and technologies enabling in your areas of expertise?

Dave Grenell: Well, on the upside because I think there's lots of downsides. On the upside-

Max Richman: We're going to get there, don't worry.

[00:47:00]

Dave Grenell: Okay. But I think the upside potentials are generally wherever there's the possibility of ground trithing. As the world becomes more and more connected, people are making decisions that impact people farther and farther away, and if the information is bad, the decision is likely to have poor outcomes. So, an example would be if you could get better information on the conditions in a far away place where the conditions are showing higher probability of droughts, that would allow a local community to get ahead of that or a development agency to prepare for the negative effects, migration, food. That would be a pretty fantastic thing.

[00:47:30]

[00:48:00] It rolls into your next question. Data by itself doesn't really do anything. The question is what kind of partnerships are in place? What does it enable and who's driving the decision-making process? In our context, for Rainforest Connection, we don't save rainforests. We provide information that hopefully enables our partners on the ground to do that, who include Conservation International, local communities and that's our goal.

Max Richman: Thanks Dave. So, we heard then again the criticality of these new technologies enabling a lot of new connections. Creative partnering, whether it's with hotels or it's with international organizations or it's with big banks, coming up with innovative business models in that space. So, I think what I want to pivot to here is the third and last stage before we open it up for some wider conversations, which is sort of the natural conversation. We wanted to start on a high, like what are all the really exciting things about these tools and technologies, but we also need to be very realistic and acknowledge their challenges, their limitations and their risks.

[00:48:30]

[00:49:00] So, my question to the panelists, now in reverse order here, starting with you, Dave. Number one: Is there a harm story that you want to share? And by that I mean, it doesn't have to be your organization, it could be, but just something where you felt that the application of machine learning had the potential to do some real harm, and sharing these stories I know are difficult, but it's important to make sure we kind of get some of the stories out. Number two, what keeps you awake at night? So, what is something you've heard about or you thought about or you've seen that just, you'd be worried about the exploitation of this data in the wrong hands or privacy information or what is it that keeps you awake at night in this topic area?

Dave Grenell: Well, one of the things, I think, will keep everyone up at night at some point, is when there are more financial institutions that are investing money using these techniques, which has become very popular, but I have seen cases where an investment firm using these kinds of advanced modeling techniques because if you're trusting human beings to invest your pension money, human beings are kind of delusional. So, now there's well, maybe computers are a little bit better because they don't have the emotional piece to it.

[00:49:30]

[00:50:00] Well, if the coder made a mistake in the way that they're approaching that problem, which happens, suddenly investments are being made in things where people can lose a lot of money, and that's happened many times. I think it's under-reported. I think that's something that we're going to see a lot more of.

Karlo Rodriguez: Well, maybe you have noted that now, artificial intelligence, data science, machine learning is a buzzword. A couple of years ago everyone was doing big data or internal things. Now, the buzzword is artificial intelligence. We see artificial intelligence not as a product, and it should be seen not as a product. It's a tool. It's like a pickax. So, in our case, you can ... Well, you can use this tool for the good or for the bad. In our case it's trying to do good, trying to provide a service. There's many cases where people turn this into bad, but it depends of the ethics of each

[00:50:30]

application, of the ethics of each industry, that is doing this.

[00:51:00] What is keeping me awake at night is the aftershocks of the latest earthquake. Last night the only night in two weeks I have proper sleep, because I'm in an area that is not prone to earthquakes. You know what give me a night is that we can not predict the industry where I'm working on because of natural effect, natural disaster. It's how we can better be responsible in the way we comply our service and how we can improve this, but even if we fail, it means that we don't alert somebody. Somebody could lose their life. If we do it wrong, we admit a false alert, somebody could be running out of the building, fall and get killed, something like that.

Max Richman: You're right, the stakes are high. Thanks Karlo. So, Peris, you know we heard from Dave some concerns about some of the machine learning leading to predicting investments in stocks, but also it could predict either including somebody financially or un-including somebody financially. What keeps you up awake at night as you think about the financial opportunities for very vulnerable populations?

[00:52:00] Peris Bosire: Thanks Dave for ... One of the things that keeps us awake at night currently is how do we get financial institutions to actually trust this model. So, how do you keep on improving the accuracy of the models that you're building? And for the people who are deep into the credit scoring technologies, your credit scoring model gets better and better with more data, with more loans that you give out, with more repayment that you get, then you're able to start seeing the patterns and picking out what are the most predictive parameters in assessing the risk. In order to do that, you need to make loans, and if you don't have the money to make out those loans, then your model is kind of stuck. It's not progressing, so I'm kept awake by I need to convince financial institutions that you have to think alternatively about how you make loans to smallholder farmers, and you need to start somewhere.

[00:52:30]

[00:53:00] You need to start making out loans with a model that has this level of accuracy and then we need to keep pushing the needle on those repayments and making it much better. There's a lot of risk when it comes to handling of data. Data privacy issues, because we get a lot of sensitive data from the farmers. They're bearing their souls out there because they need the loans. Here's my phone number, here's my ID. So, how do you keep, and when you're working with different parties, how do you make sure that wherever that data is starting or passing when you're doing the verifications, it's kept private and you're not releasing the identities of these farmers to unauthorized people?

[00:53:30]

[00:54:00] Some of the other things that keep us awake at night is the credit market, the credit industry, the criteria information is still not unified in the markets that we are working in. So, there's a risk of indebtedness where our model is predicting the potential of this farmer as very a low risk client, but the reason why that's happening is because you haven't captured some kind of credit that they got from some other informal lender, which was never captured and it was defaulted and that was never captured. So, there's a lot more work that needs to happen, even at

regulatory level, to make sure that there's a centralized way of sharing all kinds of credit information to ensure that you're not giving loans to already over indebted clients.

[00:54:30] That's a very, very big risk. Especially in the kind of market that we are working in, we have this a lot. There's a lot of loans from different providers, a lot of consumer loans, and you're coming here with your productive loan, but this farmer is already indebted with other consumer loans that they've taken up.

Max Richman: Thank you, Peris. All right, so we heard quite a few things that keep our panelists up at night. We heard about concerns around automatic investment and generally machines making decisions that are built on fallible human logic. We have the failure of systems to work can lead to loss of life or inaccurately working or falsely alarming causing other harms, and we heard about credit scoring risks that the adoption risk by organizations is really high and that there's also privacy risks and identification in an over-indebted risks as well. Obviously quite a few risks in these spaces, but hopefully counterbalanced by some of the opportunities that we also heard earlier in the conversation.

[00:55:00]

[00:55:30] So, at this time I'd like to take, we've only got a few minutes left, maybe two or three questions from the audience, and then I'll combine those together and then we'll put it to the panel and then we'll wrap.

Audience: Max, I think you did a really good job at focusing the conversation a little bit on the business side of things, which I think is like a really sort of interesting conversation on this. I was wondering, you talked a little bit about what are the plans to keep the lights on, but I was curious if you could talk a little bit about how the lights got turned on, like where did the seed capital come from? Was it a grant from a big international donor or was it venture capital? And yeah, just to all of you.

[00:56:00]

Max Richman: Great, thank you. Next question.

Audience: Yes, thank you, it is really interesting. My question is mostly for Peris. I'm thinking about who the beneficiaries are and how we kind of get things to beneficiaries. So, my question is: Do you have an estimate for how much of the value of this is being passed through to the consumers in terms of lower rates vs. accruing to the banks in terms of just larger margins by being able to better identify who's going to default or not?

[00:56:30]

Max Richman: Tough question. I like it.

Audience: I know with the Rainforest Connection, you touched on this a little bit about leveraging the data itself that you're collecting and making that valuable. I was wondering if the others could speak too what you're doing with the actual data that you've collected. Is that becoming an open resource that other people can use? Are you keeping it internal? What are the ways that you think about the actual?

[00:57:00]

Peris Bosire: Right, thank you. We started FarmDrive in 2014 with my business partner who was my classmate in college in Computer Science at the University of Nairobi. That's where we met and took different projects for four years in mobile phone development, and when we were finishing school, we were inspired by the amount of technical skills that we had acquired over the four years, and at the same time, really thinking deeply about our-

Max Richman: Microphone.

[00:57:30]

Peris Bosire: Sorry. About how we grew up. Both my co-founder, Rita Kimani, and myself grew up in smallholder farming communities, so understanding those kinds of problems that farmers face and the technical skills was a very sweet spot to come up with something. We started out initially with our own savings from college. Because of the background in Computer Science, our costs were much lower because we were able to build the technologies and the tools that we needed ourselves. And then, after some time, our lights went off. So, we had to start talking to different kinds of investors, that this is a solution that has potential for about 50 million clients. The lending market less than 1% of commercial loans going to agriculture. How do you open up this lending market?

[00:58:00]

[00:58:30]

It has a lot of potential on both parts, financial institutions and serving the smallholder farmers a very big market. We knocked on a lot of doors and got angel investors initially who kind of had a focus on agriculture and impact investing because this is an impact kind of company in itself. Then we moved on and talked to a couple of venture funds as well. One of our investors is Mercy Co Social Venture Fund. We have Engineers Without Borders and we looked at it more from capital that comes from people who really understand the market that we are serving and who have a special focus on trying to expand credit to agriculture.

[00:59:00]

Max Richman: Thanks Peris. And just to jump to the other question?

Peris Bosire: Yes, sorry. So, to jump to the other question-

Max Richman: Just cognizant of time. We've got about two minutes left.

Peris Bosire: Yeah, sorry. The impact on financial institutions, it's a very clear the reduction in costs. The direct impact on the smallholder farmers themselves is also very direct. Access to credit is a direct, especially for farmers who need it to buy input, for farmers who need it to expand their farming activities and buy assets, it has a direct impact on their productivity. There's increase in income, there's increase in yields and that translates into other kind of ripple effects in their livelihoods. We still need a lot more data to start quantifying the kind of increase per difference, kind of loan sizes. What does a microloan do when you push that loan to an asset loan that's above \$ 500? What's the kind of impact it has on our farmer?

[01:00:00]

I'll start publishing those statistics on my website once I have more data from my

farmers.

Max Richman: Thanks for answering the open data question too there then. All right, about 30 seconds each, just to make sure we wrap on time.

Karlo Rodriguez: So, it is earthquakes, because I hate earthquakes. One of my partners worked with the Clinton Foundation in the IT reconstruction. I wasn't born in Mexico City, I was born in the north of the country in a place with no earthquakes. When I moved to Mexico City for a job that I have there, a false alert in the middle of the night, at midnight, I ran out the building in shorts and nothing happen. Nothing happened and I really hated it. I told my friend about it, we could do something about it. That's how we started Grillo.

[01:00:30]

[01:01:00] About the open source. The data we have is really good quality data about earthquakes. So, some of these data, we are co-working with institutions, with education institutions and researchers, to work on that data. Some of this data is really valuable, so it's part of our business model, but all data we can share with general people.

Max Richman: Thanks Karlo. And Dave, to wrap us up.

Dave Grenell: Yeah, Rainforest Connection got started with a crowdfunding campaign, which was on Kickstarter and raised \$167,000. And then leveraged that to do more. Our main strategy has been matching funds, and that's mostly grants. I have to say the USAED support we've gotten through Dave has been kind of a game changing breakthrough for us. Thank you very much. And we're still trying to develop revenue streams. We share the data with vetted partners only at the moment.

[01:01:30]

Max Richman: Thanks Dave. Well, we are at time. I will not give a long-winded closing because we heard tremendous projects from the folks here. So, I just want to close with a big round of applause thanking them for [inaudible 01:01:53].