**NETEVENTS**

**GLOBAL PRESS & ANALYST SUMMIT**

# *Day One Opening Keynote: Artificial Intelligence: Out of the futurists' lab, into the real world of network and cybersecurity*

**Kathryn Hume, President, Fast Forward Labs**

**& Stuart McClure, Author, Inventor & Entrepreneur**

**Interviewed on stage by:**

**Paul Jackson, Principal Analyst, Digital Media, Ovum**

MANEK DUBASH, NETEVENTS

Let me start off by saying a very warm welcome to all of you who have come along here. An especially warm welcome for those who have not been here before. Another especially warm welcome for those of you who have. It's nice to see so many friendly faces here again. Briefly I'll explain the format of what we're going to do. You probably gathered that from the agenda anyway. But basically today we have a plenary session. Tomorrow we have a plenary session in the morning as today and in the afternoon the meat of the agenda if you like is the meetings between press and analysts and the vendors.

Okay so that's the nuts and bolts dealt with. Basically we have a really packed agenda for you over the next couple of days. IOT and Cloud Innovation are our themes, along with Artificial Intelligence. We've also got a shark tank coming up later on this morning which I hope you'll find quite entertaining. Since we're in the heart of Silicon Valley it seems only right to do a kind of start-upy sort of thing really.

So that's it. That's really all I have to say apart from the fact that I'm looking forward to the awards dinner this evening where we'll be finding out the results of the prognostications of our sharks to see which one of our start-ups will win the IOT and the Cloud awards categories. Without further ado I'd like to invite our first keynote presentation from Kathryn Hume, President of Fast Forward Labs. Kathryn if you're ready, please come on down. Tell us how it is. Kathryn Hume.

KATHRYN HUME

Thanks very much Manek. Thanks for the warm introduction. It's really delightful to be here. It's 2016. We're here to talk about the future of artificial intelligence, the ways in which it's impacting enterprise operations like security. I wanted to - let us sit back and invite us to take a trip 2600 years back into the past to the Greek agrarian world where Aesop was writing the stories that would become his famous fables. I think it's incredible that 2600 years later what Aesop said still has relevance in a remarkably different time and place. That's not a coincidence.

Aesop's fables, to go back to grade school - give everybody a little bit of an introduction to these - he started always with a particular story. So in the image on the screen he starts with a lion who sees a statue of a man beating up a lion and in this particular narrative says that, had the builder of the statue been a lion artist he would have flipped things around. So it would have been the lion beating up the man.

So from that very - that unique particular - he elicits a general message or moral, in this instance that the story depends on the teller. That moral be it applied genuinely or ironically can then be shifted through context and applied to tell us a new story sometime in the future. So the story depends on the teller. Everybody I'm sure is familiar with the Microsoft Tay release this summer where a chatbot went slightly rogue when humans were using racial slurs and all sorts of uncomfortable language to train the chatbot. I want to say if you take nothing else from today's conference, know that you can have a mental model for how artificial intelligence algorithms work by thinking about fables.

Artificial intelligence algorithms always start with particular use cases, particular data sets from which we elicit up general algorithms that then may or may not be able to be applied to different other use cases but both the opportunity and the complexity of this space lies within that transition from particular to general.

More concretely for today I wanted to tell a fable about the future of artificial intelligence using one of Aesop's fables. He wrote one called *The Farmer's Treasure* which tells the story of a father who on his deathbed is worried that his sons aren't going to carry on his legacy and till the land. So he lies. He's a [conceit] - where he said that there is treasure in the backyard and dies before he's able to tell his sons where it's located. So they spend years and years digging up the land in order to find the treasure only to discover that the real gold lay not within the treasure that they were looking for but actually the land that they've tilled. So the value shifts from that gold to the yield and the crop.

What I'd like to argue today is that this is an excellent depiction for us to think about - a framework for us to think about how AI is going to evolve. I spend most of my time speaking with the leadership at Fortune 500 companies who

say we've seen IBM Jeopardy - IBM put out the Watson - the Watson tool that beat Jeopardy in 2013. Shift to 2016. We've seen Google DeepMind build AlphaGo which is a tool using a technique called reinforcement learning, a set of artificial intelligence algorithms that put in position a system of rewards to train systems to excel at a particular task - beat Lee Sedol who is the leading Go champion at this game.

Executives say, interesting that DeepMind can do this but what does this mean for my Churn analysis? What does this mean for my ability to optimise marketing programs? What does this mean for my ability to protect my networks? The answer is, it's a really hard question to solve. It's not trivial. It's very hard to go from that particular original fable use case of the gold to find the actual yield in the crops that will come from these applications.

So the question that I'd like to pose is, where is this land? How do we find it? Since we're in Silicon Valley one of my mentors is a man names Geoffrey Moore who in the 1990s wrote a book called *Crossing the Chasm* - bible of marketing theory and how technologies move across their adoption life cycle. Right now in the AI space we're in that very early part of the distribution curve. We're working with the early adopters. We're building out vertically specifically applications that are the early instances of what this might become but don't yet foretell large wide scale adoptions in the pragmatic larger marketplace that's actually going to have the larger impact of these technologies across the industry.

So early apps aren't always the killer apps. DeepMind winning Go isn't necessarily the application of reinforcement learning that's going to really transform inventory and supply management. But it's possible for us not to predict exactly what those killer apps are but to hone our skills to recognise and pivot our strategy when they occur. So for the rest of the talk I'd like to just give two examples of where we're already seeing this occur in the artificial intelligence space, to give some insight into where we think the space might evolve in the future.

Fast Forward Labs, at my company we study technologies that are on the threshold of shifting from the academic sphere into wide commercial adoption and applicability. We educate our customers as to how they can apply them in their own business processes and business environments by building out prototypes with reports that explain what these tools are and what one can practically do with them and then advising them on where they can be applied in their individual environments.

Our first work was on a technique called natural language generation. I'm sure everybody has heard of natural language processing which is starting with lots of unstructured text that humans write. It's messy. It's not amenable to the numbers and digits that computers work with. So these techniques are focused on putting structure into unstructured text. Natural language generation is the opposite. We start with structured text - not text, excuse me - structured fields

like an Excel spreadsheet that has rows and columns with data entries inside of it, and then automatically write articles that use qualitative language to communicate those quantitative insights.

When this application - when this set of technologies was first released to the market we thought that the killer app was going to be in automated journalism. I'm sure you've seen in Forbes and the Associated Press articles that are written by computers that describe sports performance, weather reports, company earnings reports, lots of descriptive, non-interpretive, relatively repetitive type reporting. That was the focus of this technology when it first came to the market. Since that time Automated Insights and Narrative Science who are two of the key players in this space have realised that the real value of this tool is not in journalism per se but rather in narrative applications of business intelligence.

Taking the mess of numbers that exist in data warehouses there was a first generation of tools like Click and Tableau that would provide nice visual interfaces to provide charts and graphs to executives. They said that's not clean enough. We want it even simpler. We want it in human language where we can very quickly have insights into how our business is performing. We started off with gold - automated journalism - and shifted into something relatively more prosaic, right, narrative business intelligence but that has really had a yield for business.

The second technique we've seen evolve is in the space of deep learning. It's very hot right now. Most of the time when we talk about artificial intelligence we are referring to these artificial neural networks that are powering a lot of new applications and capabilities. We did a report focused on using these techniques to automatically discern the objects that are in images. So the application you see there, hook up to your Instagram feed and it will classify and reorder your pictures according to the objects that are in them. Just as a funny side note these systems - we're going to talk a little bit later about supervised versus unsupervised learning. So they're supervised. They have to start with a training data set in order to get a vision - what their vision of the world might look like so that they can perform.

My colleague, Hilary Mason, likes to take pictures of the New York City subway system on her way to work. The training data set that we used for this tool had no images of subways in it. So it used to classify all of these gates as correctional institutions, as prisons, which both tells us about the limitations of the tools and also give some insight into why New Yorkers have their temperament when they go to work in the morning.

We start off with fun applications to classify images. Who cares? Where is the crop? A couple of cool things I've seen on the market come out since we worked on this project, one is applying - this is coming out of the artistic world - there's a technique called style transfers where artists start with famous paintings like the *Starry Night* from Van Gogh, abstract out the style and then

enable people to go through their Facebook page and turn their selfies into works of art imitating the grand masters.

This is actually using the same set of techniques that enable us to see objects and images also enable us to abstract out the style and graft it onto other images. So this is an app called [Picasso] from a start-up out of St Louis. With a little bit more gravitas and enterprise importance a set of start-ups, one based in Silicon Valley called Orbital Insight that's taking satellite data available from a whole host of new small satellite providers and enabling us to use these convolutional neural net deep learning techniques to gain insights into macro-economic activity where the data was not formerly available. The image on the right is a picture of shadows from buildings in Shanghai that hedge funds are using as a proxy to try to get an estimate of macro-economic activity going on in areas where there is not traditional market and pricing data.

The third application, my personal favourite, comes from a San Francisco start-up called Enlitic that's using deep learning image technology to automate radiology to examine chest x-rays. Stunningly they are getting the speed of their system able to complete the work that a typical radiologist would complete in one month, in 104 seconds. So it really could have massive impact on work flows, the way in which hospitals are managing their work force in the future.

Final example - and its transition into the presentation from Cylance - comes from the world of text. We did a report using deep learning techniques, a slightly different style than those used to process images, to automatically summarise text. The application here we start with an input data set, say a relatively long article in the *New Yorker* or the *Atlantic* magazines here in the US. We build a model of the meaning of the article and then use that to select out the sentences we think best represent the meaning of the article as a whole.

So to shift around the goal here is not to absolutely revolutionise journalism but to shift around the reading experience so that users can start with the skimmed main points of the article and then read the entire article at will if it suits their purposes. What's interesting here is we shift - once again gold to yield and crop - is that we were focusing our models on aspects of text that are relevant to summarise, relevant to the human reader. Stuart and his team at Cylance are adding in additional features from text to discern whether or not incoming data is malicious or not. So same sort of technique, just shifted to a slightly different use case.

The moral of the story I think going back to *The Tortoise and the Hare* one of my favourite Aesop fables I'm sure everybody knows of, is that we're at the beginning. We really are at the beginning of this. We are not really sure where things are going to pan out and how they're going to pan out. But there is massive opportunity if we stop paying attention to the hype and really focus on the discrete applications and use cases that are going to be available on the early market.

Thanks very much.

[Aside discussion]

MANEK DUBASH

Thank you very much. That's fantastic. That's I think a really great scene-setter. I guess like a lot of people I've always been a bit sceptical about whether or not AI is real and whether it's actually going to work. Let's hear about some practical applications as Kathryn said. Let's hear from Stuart McClure who is an author, an entrepreneur, somebody who now heads up a 650 employee company which three years ago I'm told was just a couple of people sitting around a breakfast table. One of the fastest growing security companies in the world today, ladies and gentlemen Stuart McClure from Cylance.

STUART MCCLURE

Get the energy up around here. Alright. Thank you so much Manek. Welcome everybody. I'm Stuart McClure. I'm CEO of Cylance. What I'm going to walk you through today is sort of the applied application of AI into a totally different field, that of cyber security, it's relevance not just to cyber security but to a lot of different problems today. I go back to the beginnings of my career about 30 years ago believe it or not. I hate to think of it. I've got to go do my 30 year high school reunion next year. Besides wanting to get on a diet really fast and go for runs every morning I'm also thinking about those early days of learning how to program a computer.

Now back then I was fascinated with being able to control this machine with software and with my keystrokes. I could get it to do almost anything that I wanted. But I was effectively telling it to do things and it would respond. Today what we're able to do is actually teach the computer, get that computer to learn what we would be doing with it and programming with it. We've been able to do it in a lot of other industries, just not in cyber security so far. So of course Kathryn had a great example. A lot of other applications of this sort of self-learning or supervised learning but it really started I think quantitatively at least some 300 years ago with the insurance industry and being able to create actuarial tables.

A whole field of study came from it called actuarial science, being able to take large amounts of historical data and then being able to predict, in the case of life insurance for example, the average life expectancy of the insured. That became so successful that we have a $100 billion insurance industry out there in the world today. So we saw it early on in its germinations. Then it quickly turned into things like high frequency trading, being able to trade on decisions on information in almost microseconds if you will.

We've also seen the failures of that. For example - I was just talking to Kathryn - if anybody remembers the AP Twitter hack. Somebody hacked the AP Twitter feed, told the world that the White House had been bombed and the stock market went straight down for I think it was about five, six seconds. That was all because of high frequency trading that makes decisions on trades algorithmically.

Hedge fund managers love it because they make billions every year. But it can also false positive. It can make the wrong decision based on the wrong information. You have both of these sides of the equation. What I'm going to walk you through a little bit is how we've been able to apply this very very real science that's been advancing quite rapidly actually in the last - I would probably call it five to seven years - into a space that formerly had never heard of AI or self-learning or anything of that nature. That is protecting cyberspace. What we do is we make software that predicts and then blocks cyber text on the end point in real time. We do it using pre-execution AI algorithms.

What that means is - there's basically three ways to address security. You can prevent the attack. You can detect it only but you can't prevent it. You can just see it happen. Then you can respond. Typically what happens is you try to build a prevention program that's 99 per cent. With today's technology it's basically impossible. But let's say in theory you could get to 99 and prevent 99 per cent of all the attacks that occur. Of the one per cent that you cannot prevent, you want to at least detect 99 per cent of the one per cent. Then of the one per cent that you couldn't detect you're going to eventually catch years - months, years - later and have to respond to it. You want to be able to respond to 99 plus per cent of that.

Well, we are all about prevention. We don't necessarily care - I know this is contradictory I think to most in the room - who did the hack? We don't care. It could be Russians. It could be Chinese. It could be a grandma in Idaho. We don't care. All we care is that it's blocked. If it's not blocked why did it not get blocked? That's what we care about and pushing the real focus and value on prevention. I think most people would rather take a pill and actually prevent ever getting a cold or a disease like cancer or an autoimmune disease, rather than well I'm just going to respond to it as quickly as I can with decongestant and pain killers to manage the symptoms. So our focus is all about prevention.

When we started the company we said to ourselves there are major problems in our industry. The first and the biggest problem is that we cannot determine if something is bad unless we've already seen it before and it's bypassed all of our prevention capabilities. Now that is backwards. It's sort of like saying, well I can't tell that a burglar - or this person that's walking up to my house is going to burglarise - until they burgle. Right? Now you might just think well yeah that's pretty - that's normal Stu. Right? I mean you can't predict every single person.

But if you were to watch the video cameras from every home around the world, for every burglar that ever walked up to a house and burglarised it you'd create patterns in your mind. You would create connections between what they were dressed like, how they approached the house, how they interfaced with the locks. You would figure it out pretty quickly. So when a new person came up to your house you'd think, uh this person looks pretty sketchy. I'm not going to trust them. I'm going to watch them the whole time.

It's the same kind of concept that we're now being able to employ inside of software and these learning algorithms. We've basically broken down the entire process of determining whether something is good or bad and are now allowing our machines to do it. We don't have the need for 2000 analysts which is what I did. In my former company I was global CTO for a multibillion dollar anti-virus company. We had over 2000 people that sat there all day long and just determined what next signature to write. So what bypassed these other controls and now - what do we need to look for now that's new that will catch them tomorrow and the next day and the next day? Over 2000 people doing that all day long. We said we want to eliminate 99 per cent of that effort because we can't possibly scale like that. The threats come out way too fast.

So by applying this science into the problem of good or bad in cyber security where we were able to actually replicate and improve the detection capabilities even in our first model from the industry 30 per cent up to 75 per cent with seven data scientists. So seven people versus 2000. We tripled the efficacy. It was nothing short of staggering. So we knew we were on to something. If you look at the industry and where it has come from in terms of just detection - forget about AV - just any kind of cyber prevention capability. It always starts with signatures, trying to detect after the fact why was this bad?

The problem is of course the speed of these things that come out are just way too fast, way too voluminous. What we've done is we've created these sort of generic signatures or generic ways of determining good from bad. But ultimately we've all done it using human beings. We haven't trained a computer yet to look at these things. That's what we've been able to do. We talk about two parts of AI quite a bit - supervised and unsupervised learning. There are two parts to what we do. The first part is we automatically look for features that are going to be potentially indicative of good or bad. For example I would look in this room and I'd say, the first feature might be to determine between a male and a female in this room I'm going to measure hair length. That's a feature.

I'm looking around the room. I might get pretty good here. Pretty good. Maybe not 100 per cent, maybe not 95, but probably about 90 per cent. But if I then took other features that were very indicative and I had a computer go through and look at all the images of everybody in the room and be able to tell you that, well somebody that has a beard, facial hair, is more indicative of a male than a female. There's another feature. Now imagine, those are two features and you

could probably get to 99 per cent determination of male or female in this room. Now if I told you that we are over five million features that are indicatively defined as malicious or safe you probably wouldn't believe me. Right? Five million? That's insane.

I communicated two to you, but to get to 100 would be hard, 1000 impossible and five million is off the charts. But that's what the unsupervised nature of what we're doing in AI does. It actually determines which features are most predictive of good and bad in binaries and files. The second part of it is really a very supervised part which is - okay this feature is definitively good or bad - that requires human intervention still to this day. So we can get a listing of all new features that seem to cluster. But we don't know if it's truly good or bad. So that supervised part has to become a part of it. How we do it is we collect as many files as humanly possible. It would be like collecting as many pictures of human beings in the world in our earlier example of gender, as many as we possibly can. Then we extract as many features as we possibly can that we've already mapped or learned are potentially useful. Then we transform those. We vectorise them then we train - using neural networks for us - we train on what is actually going to cluster to good and what's going to cluster to bad. Then we classify it. If it's bad we block it. If it's good we allow it. It's that simple.

We might be one of the early players in cyber security to apply AI and machine learning into the space but we will not be the last. In fact I think it will probably be the only thing that saves our industry ultimately is this kind of technology and this way of learning. We just can't possibly scale to meet the demand. So with that I hope I'm close on time. I wasn't doing my check. But thank you so much Manek.

MANEK DUBASH

Thank you Stuart. Now I'd like to invite Paul Jackson from Ovum to come down and give them - no - to invite them to explain more details about their technology and thoughts about AI. Paul Jackson.

PAUL JACKSON

Thanks Kathryn, thanks Stuart. I've got a couple of questions here for our speakers but starting [unclear] questions I really want to get a couple of questions from the audience towards the end of our slot here.

Very interesting insight. I think - to both of you AI has been around for a number of years. But I think to a lot of us it seems to have really come to the fore over the last 12, 15 months. Why now? Kathryn why is this suddenly at the forefront of a lot of companies' thinking?

KATHRYN HUME

Yeah sure. There are a couple of reasons. You're right in saying it's been around for a while. So the back end techniques these artificial neural networks

that Stuart was talking about that I mentioned under the umbrella term deep learning - one of the first instances of a - there actually used to be hardware as opposed to software. There were a bunch of wires connected together. In 1948 there was a thing called a perceptron that was just a bunch of messy wires. It was attempting to mimic the structure of the human brain so as to develop artificially intelligent systems. That's where the lingo comes from.

But there were some problems, (a) there wasn't a lot of data to work with. We didn't have the big data area - I use the term big data to refer to storing and processing data, not doing stuff with it. So 10 years ago it became really cheap to store a lot of data, keep it up in the cloud and then do stuff with it. In 2011 is I think was when Google had a first coup using artificial neural networks to automatically identify cats in videos across the internet. As Stuart mentioned there was part of this that was unsupervised. So figuring out the computers, figuring out that there was something about cats that made them similar [unclear] could cluster together all these patterns. Then the supervised part was humans coming in and saying, oh yeah that thing you see that looks kind of like a blob of something, this amoeba thing, that's called a cat.

So we put our human label on top of that. There have been some changes since then that have just propelled this forward. The one is increasing compute power. So the best - I was talking to Paul about this before - the best description I've heard about the power of the cloud is that today for the same price you can rent 1000 servers for a day or you can use one server for 1000 days. So if we think about that shift, if you can do 1000 times the stuff in one day that's a big shift from an economic perspective.

The second is if everybody's heard of graphical processing units - GPUs - coming out of Nvidia - back to my gold versus yield - some kid playing video games realised that the structure of GPUs to process images was pretty good at matrix multiplication which just so happens to be the type of math that's powering these deep learning algorithms. So they said, the gaming industry is huge but gosh this other thing might be a lot bigger if we can actually apply these things to enterprise artificial intelligence needs. So there was that pivot which is opening up all sorts of things because we can train these things faster.

Then the third of course is the data. There's just more - it takes a neural network and in Stuart's case - actually let's go to the cat case - we as humans can see two or three cats. The first time we've seen these animals and if somebody says oh yes this is a cat we can then see our fourth cat and recognise it as such. It takes a neural network probably 50,000 examples in order to gain that ability to recognise things. So you can imagine if we're going to go through all of the types of objects we might want to identify to build a visual recognition system we need a lot of training examples. So that data has also propelled the transition.

STUART MCCLURE

Yeah, I'm going to just echo it but emphasise it a bit. The first is the computational availability today. We never could have started this company and done what we've done without the cloud, without Amazon in particular and AWS. So for us to build a model - in the early days - so I guess two years ago - three years ago - when we were building models it would literally take about six months to build a model. Now that was with as many CPUs as we could possibly hire from AWS. Today our models take about a day and a half to build. But we have to spin up over 10,000 CPUs to do that day and a half. Without that flexible compute fabric there's no way we could be doing what we're doing. It's just that simple.

The second, which I can't emphasise enough, is just the data. So it's sort of like - many of you probably have children and you remember the early days of having to explain to your child and associate the word ball to a round object that bounces. You introduce more and more of these different kinds of balls. You say ball, ball, ball. The eventually start to get it. But it probably takes dozens if not hundreds of times for them to associate the word with the visual image. Then you bring in a cantaloupe one day and you say what is this? They're going to probably - if they've only seen one ball ever and they've never seen a cantaloupe they would probably say ball. Most often.

It's because they've tried to identify certain features of that object and then apply it to another form. Well of course they'd be wrong but once you clarify and say no, this is a cantaloupe. If you showed them more cantaloupes and more balls and they were able to identify the individualistic features, you're able to create a distinction. That's the same thing that we see in today's AI is the more data you give a learning algorithm the smarter it gets. But like Kathryn says it's not like 10 balls or 10 cantaloupes. It's millions of these things. That's the only way that the algorithms today can create a real cluster or a separation between one or another class. So both of those I think are just critical parts to a successful application of AI.

PAUL JACKSON

Stuart you were talking about several examples in the security space. We are increasingly facing many more sophisticated types of attack. You're looking specifically sort of the end point protection?

STUART MCCLURE

Yeah.

PAUL JACKSON

A number of us spent an afternoon with NetScout yesterday looking at [their denial] of service tools. Is security one of those areas where AI is particularly

well suited or is it just because we're wanting that instantaneous response that a lot of these trained tools can do far better than people?

STUART MCCLURE

Yeah. No I think cyber security is the perfect place to apply AI and machine learning. Quite honestly I don't know why it hasn't been done before. I mean it's sort of one of those things that - once you tell somebody about what we've been able to do they're like wait that's never been done before? That seems pretty easy, straightforward. That would be a natural assumption to apply. So when you think about the other problems though that AI could solve inside cyber security - there are really three core problems in cyber security. That's it. There are three core ways attacker gets in.

There's number one, denial of service which you talked about. Okay? That's just starving the resources of the target. So you starve memory. You starve network bandwidth. You starve a CPU or a disc or something and the system falls down. It breaks. Denial of service attacks. That's the first.

The second is what we do. Execution based attacks. An attacker gets something sent to you or gets you to click on something that executes something in memory to do bad things on your box. Okay that's the second.

The third one is authentication based attacks. So being able to steal your password and pretend to be you on your computer when you're not there, or bypassing authentication or brute forcing your password or any of those things. Like the AP Twitter hack - there's only one way to hack Twitter. It's the password. So somebody guessed or brute forced or something the AP handler guy that's supposed to type in the AP news reel - their password for the Twitter account. It's that simple. There's nothing more exotic than that. So those are the three. All three of those spaces can apply AI in a very very meaningful way. You just need the data.

PAUL JACKSON

Okay thanks. Final question from me before we go to the audience. It's one that you tend to get at every conversation about AI. We've talked about unsupervised and supervised learning. The whole fear of wholly unsupervised, sort of ghost in the machine, Skynet like, growth of AIs is something that's been discussed a lot in the press and by some names that you would have expected to be wholly in favour of technical process. Comments on that? Is it unfounded? Is it realistic? Is it something we have to keep an eye on?

STUART MCCLURE

Well I'll throw my two cents on it. Kathryn could probably talk actually a little more in depth on it. But my feeling is we're a long way away from having unsupervised classification, of being able to classify one thing as a ball or a cantaloupe all totally unclassified. In other words you give the - so it learns what a ball is. It learns what a cantaloupe is. You just expect it to know that this is a car, this next object. We don't - we're nowhere near that today. You still have to tell that well all these features are car-like and so now treat all things that are relative to this feature set as car. That's where we are today. Maybe in 10, 15, 20 years we'll have a different discussion. But at least my viewpoint, we're a little ways away from being able to realise the whole Skynet making decisions for yourself and for human population or whatnot. So still a bit away.

KATHRYN HUME

My take is actually that the thing to be more concerned about in the near term is supervised learning and not unsupervised learning. It's not because computers are dangerous but because people are dangerous. Here's why. A lot of these discussions coming out of - Nick Bostrom [at Oxford] who wrote, superintelligence is one of the key thinkers thinking about the long term speculative risks of systems that become truly intelligent. There's a place for that. This technology is changing faster than we can keep up with. It's worth our time to potentially imagine potential risks so that we can be prepared.

But today humans - there's all sorts of stuff that we do as people in society. We leave traces of that in our data. As Stuart mentioned supervised learning algorithms start with human input. So we train systems based upon the decisions that humans have made in the past. So let's take an example of using algorithms to try to automatically hire somebody into your company or recruit students to your school or even give a loan for a credit application. If we try to automate that, the systems aren't that smart. They go out and they look in data sets. If in the past - let's take an example of - we're right near Stanford - Stanford has tended to recruit a certain type of candidate. They've tended to recruit relatively wealthy white males let's say for sake of the example.

We go into the system and we say here is a model for the type of candidate we're looking for. These are the decisions that humans have made in the past. The algorithm will go and they'll say okay cool. I'll go find candidates that look like that. I'll base my decisions based upon what the humans did. Then the algorithms come back and they say here is a pool of 95 per cent rich white males that we suggest you recruit to your school, precisely because if we think about a normal distribution this is where the bulk of the features tend to lie. So if we relegate our decisions to the algorithms they tend to propagate and amplify the stupid decisions we as humans have made.

So the ethics in my opinion lie much more - it's not about systems being intelligent. It's about our mixing together the corporate values with social

values to as data scientists take an ethical position with regards to potentially having to add a little bit of - hack the algorithm so that we can create the future that we want as opposed to the one that just perpetuates our biases from the past.

PAUL JACKSON

Interesting. I think one of the nice examples from the Go playing that came through was that after being beaten the Go player went back and looked at the moves that the AI had made and actually learned new things about how to play the game.

KATHRYN HUME

The Go player also found that the system - he was Korean - the system tended to play like a Japanese person. He had a prejudice against the Japanese playing style. So he could tell that it was trained on Japanese Go players.

PAUL JACKSON

A few minutes for questions from the audience. Anybody? Show of hands? Mic on the way. At the front.

HANS STEEMAN, JOURNALIST

A question regarding the DNA. Once you recognise a certain DNA [in] the zero attack what about an industry wide data bank where you store these signatures and share that will all the vendors. Because now every vendor is struggling - fighting against the world. By uniting all these vendors you are now more powerful and you can faster deliver solutions to the end user.

STUART MCCLURE

If I understand the question is, why don't we share what we've learned? Is that basically - okay? Well for now the short answer is we want to be able to sell something [laughs]. So that's the short capitalistic view. But I will tell you that just like almost everything eventually the value and the importance of what you're doing transcends the need to make yet one more dollar. So I think when that starts to occur - we're a few years away from it but - when that starts to occur I think there will be more sharing about it.

We've spent - $177 million is how much that has been invested in the company. You've got to pay that back first and then start to build on top of it. But I think

you can. I think the platform is there or it will be there in the next few years with more and more cyber security companies trying to do similar stuff as us. The problem I find is that no one is really doing what we're doing yet. I see a lot of marketing language but when I talk to them, my data scientists talk to their data scientists, they're nowhere close to us yet. As soon as they kind of start to crack the code on their own and they buy in to this new approach, I think a community will get started then created. But we have to almost purge ourselves of the legacy first before they're going to be open to a new way of doing it. So we're still in that early stages but I think eventually it's possible.

KATHRYN HUME

That's not really my space [laughs].

PAUL JACKSON

We have another question down here.

HECTOR PIZARRO, DIARIOTI

My name is Hector Pizarro from DiarioTI, covers Spain and Latin America. I have a question for Stuart regarding the authentication based attacks.

STUART MCCLURE

Yes.

HECTOR PIZARRO, DIARIOTI

You described the case of a burglar that will reveal him or herself by his own behaviour. So what would you do with a digital version of the burglar ransomware for instance that would simulate being the owner of the equipment and such have legitimate access by forging the password for instance? How would you identify that behaviour of the malware and differentiate that from the behaviour of the owner?

STUART MCCLURE

It's a common issue. As you retrain - so you've trained and built a model. Then you deploy the model and you find out which things bypassed the model. Like the owner that dresses up in a dark hoodie - black hoodie with a flashlight and all that stuff. So then you have to go back and you have to increase the number of features. You have to figure out what more features are present on the

system that distinguishes the home owner in that outfit from a bad guy in the outfit?

Let me give you a real world example. In our early models we trained on malware, APTs and zero days and adware, spyware, Trojans, viruses, worms all these things that we knew and we had a lot of data on. Well it started to catch things like WebEx and GoToMeeting as malicious. Now you might thing oh Stu that technology is crap. Like it's catching GoToMeeting as malware? But we said well hold on a second. Let's look at what GoToMeeting and WebEx can actually do. It can capture your screen. Somebody can take over your keyboard and mouse. It can exfiltrate files on your system directly. We can capture video of you doing everything that you do. That's malware.

So to create that further separation we had to train it with tons of GoToMeeting and tons of WebExes. The non-supervised part of the feature extraction started to find features that were truly distinctive between malware and WebEx and malware and GoToMeeting. So again it's just more and more features. That's where a lot of people stop. People stop and 1000, 3000, 10,000 features. You're going to just have severely high false positive rates just like that. Again it goes back to the data. You've got to extract more and more data and just not give up.

PAUL JACKSON

Probably the final question and then we're going to get thrown off the stage I think are we?

We're okay.

PAUL JACKSON

Excellent. Let's press on.

SIBHALE MALINGA, IT WEB

Good morning. My name is Sibahle Malinga. I'm from ITWeb in South Africa. Thank you for an insightful discussion. Stuart my question goes to you. I completely understand everything you've explained about choosing AI methods to fight and to predict cyber security. But how are organisations currently - or should I say traditionally - using methods which detect cyber-attacks before they actually occur? What are those popular methods that are currently being adopted? Where are we seeing the gaps or the loopholes there?

STUART MCCLURE

The traditional techniques for detecting something malicious and trying to prevent it is signature based technologies like anti-virus or any other, quite frankly any other technology that's currently in the market today. What it does is it says, okay I'm watching this behaviour or this activity. This behaviour or activity is exactly what we've seen a year ago on a computer in South Africa. They got hacked with this sequence of activity. So now on my computers and on every computer around me I know that that activity is malicious. Before it gets too far I can stop it.

But it required a sacrificial lamb, somebody to actually get hacked first. What this allows you to do is eliminate the sacrificial lamb. You no longer need to have a victim to know that something is malicious. So that antiquated version of signatures which has evolved into things like generic signatures and heuristic signatures or behavioural signatures, these are all still signatures because of one thing. You can actually get a definition of why something is called bad. Here's the big trick to demystify and separate those that use signatures and non-signatures. You ask a vendor or a company or anybody - you ask okay so you called this thing bad. This is malicious to you. Why did you call it malicious? If anybody can actually tell you why, it's a signature.

I can't tell you why we call - we've mapped to five million features. They're all mathematically calced now. So I could tell you well okay 120,000 features were found in this one file. That's why it was called bad. But these features are abstracted from the actual original names and words and things like that. So I can't actually explain to you why we call it bad. It's just mathematically it aligns feature by feature more to malicious. That's the real differentiation that you can quickly find in the market.

PAUL JACKSON

It will be our final question now.

STUART MCCLURE

Yes - oh two more questions. Okay.

PAUL JACKSON

Two more questions sorry.

GABRIELA CHAVEZ, CNN EXPANSION

My name is Gabriela Chavez. I'm from CNN Expansion in Mexico City. You mentioned in your presentation that you think that AI, the application of AI as cyber security is the only way to save that industry. My question is if you

apply AI in a bigger scale, in a massive way, how far can you minimise the rapid increasing of threats or ransomware? Could really AI stop [a little] the trend?

STUART MCCLURE

The answer is yes. We're already seeing it today. With the AI technology that we have out - we have it installed on over three million end points, three million computers today. We've only been - it's only been available for two years. So it's just taking off like crazy. We're already seeing the ability to have all of that technology truly truly predict and get to the ninety-ninth percentile. So eliminating all of these - today - every day that we power on a computer and take a look we get hit with about - in the world - about 350,000 to 400,000 new attacks that come out.

The define attacks by binaries or files that are malicious or have malintent of some sort. Those are the ones that we just know about. So 350,000 to 400,000 every single 24 hours. As soon as that starts to reduce, in other words if that number starts to go down in terms of every day, so maybe it's only 300,000 today and then 250,000 and then 200,000, were going to start to see that the technologies like ours that are truly predicting all these new fancy attacks are no longer bypassing the systems that they're going to target. They're now getting caught. They're getting prevented. So there will be a natural desperation for the attacker of like well I've got to create even more. I've got to get even more sophisticated or more samples of the same technique.

When they realise that doesn't work then what will inevitably happen is they'll get more sophisticated. Now when they get more sophisticated that's where we start to worry. Because there's always a way to bypass anything. So if anybody is going to do it, it's going to be a nation state of some sort, the US, Russia, China, whoever that can spend a lot of money on it to try and bypass us. I don't mind them bypassing us actually. I would actually love it. Because every single bypass makes us Stuart McClurearter. We just have to know about it.

So if they hide it from the world then yes we can't possibly improve. You'll have the same problem all over again. But if we can be very responsive - catch it after the fact, build it back into our model, retrain - then a whole new class of sophisticated techniques will now be stopped again. We can do that very quickly new. We can do that within minutes today if we get a sample.

KATHRYN HUME

One of the things I'd add is that, Stuart mentioned there are so many - there are five million features they're looking at so as to classify something as malicious or safe. The world of artificial intelligence calls that interpretability. These deep learning algorithms are uninterpretable. So we don't really know why. This is his system. He says I don't know why it said it was good or bad but it

did, and it did reliably. So that lack of interpretability can be a problem for human oriented problems like is someone sick or not? Imagine you use AI for a diagnosis of somebody's cancer illness. We say it says they have cancer. We don't know why. We don't know if it's their heart. We don't know if it's their blood. We don't know what the features are. That's problematic there. It's great in security because it's harder for the adversary to identify exactly what they should shift so that they can pass the gates the next time. So what is a problem for some new cases is great here because it's harder for the - there's no rational human explanation for what they should do next so as to stay ahead of Cylance

.

PAUL JACKSON

That's always been one of the issues to a degree with neural networks is it is a black box.

KATHRYN HUME

That's right.

PAUL JACKSON

It's saying yes or no. You're not quite sure why particularly once it gets that sophisticated. You've been very patient down the front. The final question.

ANTHONY CARUANA, CSO MAGAZINE

Thank you. I'm Anthony Caruana from CSO Magazine in Australia. I was interested in the point about human bias in AI development that Kathryn made. The Go computer was defeated by the human when the human made a completely unanticipated allegedly wrong move and it confused the AI. I guess the question is really for Stuart. It's - how do you guys overcome that human bias in the program when the bad guy does something really completely outside the model and unanticipated? You mentioned then you kind of like that because it teaches you something. But that's not helpful for your customer that's just got broken.

STUART MCCLURE

No it's not. No it's not. I mean luckily we're in the 99.9 per centile today. So if they were dealing with a - 200 attacks last year, they dealt with one today for this whole year. So it's dramatic improvement. However you're right. I mean ultimately everything can be bypassed. So for us the beautiful part is we can retrain constantly. The quicker that we can identify the bypass and get it into retrain mode the better. For example we have a new module that's coming out

next month called Optex. What Optex does is it actually records every single activity on the system itself sort of like a black box on a plane. It records everything that happens on it so that if there is an attack we can take all that activity and replay it into our simulation and figure out what are the features that were present that bypassed our technology in that moment?

We can get that identified very quickly and pushed out. So in the worst case scenario it would impact customers by minutes, not by hours or days. So we're getting pretty good at that. It's something called centroids that we've built with every bypass. Luckily there are very very few so we don't really have that problem. But inevitably it will occur. The goal is just to limit the damage when it does is the bottom line. I've always said this. Security is not 100 per cent. It's not a finish line. Some people just think well we should be able to get to 100 per cent security. It's just impossible.

The bad guys could plant an engineer in my company and create a back door code that would bypass the whole thing. Now in theory our software development life cycle process should find it. But if it was collusion with the auditor as well - I mean there's always some way to do something malicious and bad. The trick is to elevate the expense and the time and the effort that it would take for people to bypass it - to only those that quite honestly can spend hundreds of millions of dollars to try and do it. Because that will at least eliminate 99 per cent - I always said our goal is not 100 per cent security. It's to motivate the hacker in Minsk or Shanghai or in Herndon Virginia to go get a job at KFC and Taco Bell. Just find another job. This is too hard. That's all we wanted to do is to lift the bar to those that honestly can't. Because most can't. Most are just ankle-biter hackers that really don't have much skill.

Oh and one last thing actually, kind of an important point. Now Kathryn you might - I don't know if you believe this as well but - I really believe today the other reason why AI is really working in this space and others is because there is nothing new. There really is nothing new. I wrote a book called *Hacking Exposed*. It's been published now for I don't know 16, 17 years. It's over a million copies, translated into 30 languages. It is an encyclopaedia of how the hacker gets into computer networks and systems. Every year or two I update the thing. I shuffle deck chairs. There's no new content. I'm telling you. Trust me. This is the big reveal I'm giving you. There's nothing new.

[Unclear] buying the new issue?

[Laughter]

Don't buy the new issue. I'll send you a free copy. No but literally there is nothing new. I mean sure there might be new twists to an old technique but it's sort of like, how many ways can you break into your house? I don't know. I mean the garage, the front door, the back door, the windows, the chimney. There's no teleportation device yet. So that's not a technique. Now if teleportation comes in 20, 30, 100 years you're absolutely right. We're not going to be able to detect burglar that comes into the house.

[Unclear]

Drones - yeah drones - maybe micro drones is another way. We haven't been hacked by micro drones before - or burglarised before. Now we have. So we find that as quickly as possible. We learn all the features of the micro drone or teleportation elements or whatever and then we can do it. But you know, there's only an epic change like that in the threat surface area every 20, 30, 50 years basically.

KATHRYN HUME

Yeah in terms of [got new] stuff in outside of security so most of these systems are good at - they're not as smart as we think they are. They're good at automating things that are relatively repetitive, things that aren't new. It's precisely because they're not new that it lends them - they lend themselves to a lot of these systems. I think where things get exciting is especially in the realm of text working with language, computer systems that do things like summarise or try to write their own novel tend to generate stuff that's very weird. In the commercial world in order for them to be viable for business processes they have to behave in ways that seem rational to humans. So we're going to exclude out - the market will exclude out weird creative stuff. But it doesn't mean it's not there. There are groups of fringe artists that are doing incredible things, generating films, generating novels by rearranging past human text in ways that's weird. I think there is incredible creative opportunity there but it's going to sneak in through back doors that aren't necessarily where we think they'll come from.

STUART MCCLURE

Probably no more time but…

PAUL JACKSON

Yeah. Kathryn, Stuart. Thanks. That's been absolutely fascinating. So thanks very much for your time this morning.

STUART MCCLURE

Appreciate it, absolutely.

MANEK DUBASH

Fantastic. Thank you very much Paul, Stuart, Kathryn. Really interesting. Don't you get the sense that we are somehow right at the edge of a brand new abyss? Minority report, Google cars, you name it. There's going to be a whole new world of stuff out there that will be able to do stuff automatically,

autonomously. We won't know why it's doing it. Isn't that interesting? Maybe that's kind of me up on stage. I don’t know. Anyway I'd like to move on and introduce Andrew Braunberg from NSS Labs, a well-known security testing lab who is going to talk to us about defending against the unstoppable, ransomware. Andrew Braunberg. Come on down.

ANDRES BRAUNBERG, NSS LABS

Good morning. Oh yes please, panel come on down. I'm going to get started.

[End]