RAW COPY

ITU
MAY 15, 2018
3:45 P.M. CET
AI FOR GOOD GLOBAL SUMMIT
ROOM C2
PANEL 4: SAFE AND SECURE AI

                            ***

    >> THOMAS WIEGAND: So welcome, everybody.  Is the microphone
on?  Maybe not.  Oh, no.  Welcome to this session, Safe and Secure
AI.
    We have a list of excellent speakers, and we'll handle the whole
session similar to the one in the Popov Room, where we have a talk
and maybe one or two specific questions, and then at the end we'll
have a hopefully lively discussion.
    If ever we let AI make important decisions in our life, we want
to be sure that it does it, and there are several ways to look at
it.  One is the engineering way.  I'm an engineer.  So the technical
system should work as intended.
    Another one is the ethical way of looking at it, is that it should
be -- it should reflect the consensus of the people that it -- this
thing should even exist, and I also consider that as a safety item.
    And then it should also be protected from malicious attacks,
so we have a self-driving car that is hacked by somebody and it drives
into -- let me put it in a different way.  Imagine a scenario where
1,000 or a million self-driving cars are hacked and be forced to do
things that we don't agree with.  We need protection from that.
    Or imagine we hand over our money, all of our documents to an
electronic file system and that's being hacked, and then the money

is erased and the ownership of our house is erased.

So we have lots of these aspects that we will see in this brilliant new world that we have to work out and have to get right, and if we don't get that right, then we probably shouldn't do it that way and should go back to the old-fashioned way, and this is where we start.

So I'm going to start with the first speaker.

And so our first speaker is Robert Kirkpatrick, and he lives in New York.  I just came there and did bicycling in Manhattan.

>> ROBERT KIRKPATRICK: Dangerous business.

>> THOMAS WIEGAND: It's actually great.  If you ever have a chance, go biking in Manhattan.  It's good.  They rent these bikes and you can go all around Manhattan.  He's a social innovator for more than 20 years, and he works in the public policy of organizational changes.  He's the Director Global pulse, and he's the perfect speaker about misand missed uses.  Please go ahead.

>> ROBERT KIRKPATRICK: Thanks very much.  What I'm going to do is quickly show you five different projects.  This is all about sort of real-word examples.  That word is xenophobic, not xenophonic. Five things related to refugees, since there's so much discussion about the SDGs these days.

What you see here is a prototype that we've developed for UNHCR that's going into implementation.  This is a tool that's connected initially to the Twitter Firehose that is recognizing, classifying, and quantifying content, xenophobic content related to refugees. It's to use that information as a proxy for estimating the types of risks that refugees might face along a particular journey, where you see people reporting problems with border crossings, detention, denial of access to services.  That could be used in a way that you simply draw a line on the map from the origin to the proposed destination, and it will tell you the types of likely risks that you face along the way.  And it uses not only sort of the initial modeling that was done through natural language processing, but you can continuously train it to improve the model where something gets misclassified.  You can drag it over to the right category.

The point is this is now something where we see an opportunity potentially to scale this up to dozens, if not hundreds or even thousands of languages.  It's possible to see the risk of discrimination by any group against any other group in realtime by looking at what people are saying.

We've been working on a project -- again, this is moving now into the implementation phase -- with IOM and UNHCR, looking at refugee rescues in the Mediterranean, and the idea is that, you know, there are a lot of vessels in the Mediterranean that have been involved in rescues, which are dedicated rescue vessels, but increasingly there's been a pattern of commercial ships encountering refugees, intercepting the distress beacons and going to investigate and playing a role in rescuing them, so what we are doing, you can see here we're using data from -- marine traffic great from the AIS data and rescue beacon data to train deep learning algorithms to

signal rescue events when a ship deviates from its normal course and slows down and begins to zigzag, right, and double back.  That's the sign it's involved in a rescue, and by training it to know rescues -- the idea is to develop an early warning system here that can tell you hours and hours ahead of time when people in need are going to be coming ashore.

This is work we've been doing for UNICEF in the camp in Jordan.  There are services provided by third parties that provide sanitation to, for example, empty the septic tanks, all across the tanks -- sorry, all across the camp.  The tanks and the trucks have sensors, the tanks in the ground have sensors, and what we've seen repeatedly according to UNICEF is a pattern of tanks never being emptied because they were difficult to get to, as well as bad conduct by drivers because they didn't want to have to travel the distance to the place they were supposed to dispose of it.  We've used deep learning to develop a model that predicts which tanks are least likely to get emptied.  That could be used to set a Bounty to make sure there's an extra incentive for them to empty those tanks.

Here we've been working with UNISAT on satellite imagery using deep learnings, convolutional neural meds, to as closely as possible approximate the structure recognition capabilities that their current human lists use to annotate structures and understand how many people are likely to be in a camp as well as to track their growth over time.

As you can see, the method on the right is getting pretty good.  UNISAT has told us we have to hit about 97.5% accuracy to be useful, and we're at about 95%, so getting there.  That last mile is hard.

And finally, we've been using deep learning on speech within indigenous languages to reach across the digital divide.  One of the projects we started almost four years ago now was using speech recognition of indigenous languages in Uganda and essentially training the system to capture talk radio shows and news programmes all across the country through a network of little raspberry pies, streaming it into the Cloud, and turning that into machine readable text.  You don't have to get to 100% to be useful, so being able to see people talking about symptoms of a disease or what they're paying for maze or soybeans or cavasa, reporting flooding or landslides is very powerful, you're recognizing segments of audio where somebody mentioned those key words or phrases, tagging them, and dropping them into a queue.  We've been applying this for refugees who are coming into the north from South Sudan, and you can see, for example, there's a perception of rumor spreading that a possible anthrax outbreak was caused by refugees bringing in their own animals without screening.  This touches on the use of hate speech and discrimination.  A lot of these rumors are circulating, and you can see where they're being spread.

So when you do a project like this, I mean, the same properties that make this kind of data very useful to help, of course, can also make it useful to harm.  These data sets are even -- even the public ones, like the radio content, are incredibly sensitive, and there's

tremendous potential for misuse.  This is something that people are increasingly aware of in the world.  People think of big data and they think of Big Brother, right, you think of the privacy issues that come along with the world of ubiquitous algorithms.  When we do these projects, we have instruments we've developed to try to guide our work that look at both the risks -- the risks that come along with potential misuse and the mitigation strategies for dealing with those but also what the benefits could be, and I think this is really important because a lot of the conversation today is really focused on privacy and not so much understanding the costs of not using the data.

We -- we've been working within the UN Development System to develop this set of guidelines on big data for the achievement of the 2030 agenda, data privacy ethics, and data protection.  This has now been adopted by 33 UN agencies, and we have a set -- we have this tool, which is a risks, harms, and benefits assessment for projects using big data and algorithms that's meant for any organization to be able to use to sort of assess and weigh the risks and benefits of any project that you're undertaking from many angles.

So I think in conclusion, I mean -- I think it's important to sort of understand how we think about harm here because every physician understands that the principle first do no harm has two equal facets, right.  One of those is do nothing that could harm the patient.  The other, however, is take every reasonable step to prevent preventable harms from befalling the patient, and I think when we look at what's happened in the space of big data, we think that existing privacy regulations are actually failing on both sides of that equation.

On the one hand, they don't adequately protect people from the unique risks of big data, right.  It's so easy to reidentify people from sets of behavior over time, and it's very hard for people to take the notion of consent for medical practice and translate that into a world of data where a data set that I produce today could be combined with a data set I won't even produce for three years and use to harm me later, but I'm asked to consent now, right.

And then there's an even bigger problem, right.  The even bigger problem is while privacy is a human right, so are food, water, shelter, access to education, health care, and justice, and I think what the majority of the world hasn't realized yet is that the data that we've all been producing without our knowledge that's been used by public and private sector a lot in ways we can't determine could, for many years already, have been used to improve public services, early warning, crisis response, and accountability.  The opportunity costs that people around the world are paying by living in a world where all of this data isn't being used in a way that directly benefits them is very significant, and at Global Pulse, we really believe that we need to move from the sort of prioritization of privacy as the primary source of risk to an approach that's more holistic and looks at balancing the risks of misuse along with what we call the risks of missed use.

I think, you know, today, most of the innovation in AI as in other emerging technologies is being driven by private sector, and when there are unintended consequences, historically these things happen and then the sort of oops, we need some regulation here. We've seen recent examples of that. The people in the organization and representatives here are a testament to the fact that that's changing. In certain areas like nuclear engineering or environmental science, you can't just innovate however you wish. There's a mandatory pause at the outset to reflect on what the risks could be to human rights, both in terms of what could happen if it's misused but also what could not happen that should happen if it isn't used, and I think it's -- we really need to look at how we can create, I think, a similar kind of approach in a space of artificial intelligence where there is that mandatory reflection and consideration of how to address risks on both sides of that equation because the opportunities are huge, but we're not going to be able to take advantage of them if we can't deal with the downside risks. Thanks very much.

(Applause)

>> THOMAS WIEGAND: So any questions from the audience?

Well, I do have a question about the tradeoff between privacy and what you could actually do with data. It seems that -- okay. Let me start by me saying I'm a German, and after the census in the 1930s that -- the result of that census was used to kill a lot of people, and that is one of the reasons why in Germany the view on privacy and on these things is very different and Germans are very conservative about it, so what I'm observing is across the globe in other countries, it seems like privacy is given up very easily and without even the proof that the use of the data after the privacy is given up actually provides an actual gain, and for whom; right?

I was at a conference where a founder of a big search company said, well, why don't you give us all your data, we'll make your life better, and about that person, you don't have any data, so how do you see that, and you can probably tell my perspective on that.

>> ROBERT KIRKPATRICK: I mean, there are a couple of things happening, I think, that are interesting in this space. One is that we're beginning a paradigm shift away from data being owned by corporations toward one where -- I don't know if it will be in five years or 20 -- where we, as individuals, will be in control of the destiny of our data. You can already see this changing in the public consciousness, and this is going to eventually cause political change to the point that, you know, companies have -- I mean, the phone company has to have your phone number to route your call, but that doesn't mean they should be able to sell your behavioral information to anyone they choose, right. Eventually we're going to live in a world where individuals can control who gets access to their data and how it's used. We're not there yet.

I think until you have that -- you know, until we get there -- we're entering a society -- the world is moving in the direction where in a few years you're going to have essentially everybody be able

to wear a device that's free on their wrist that captures everywhere they go, everything they think, everything they feel, everything they buy and sell, everyone they interact with, and the functioning of every organ in their body, and Apple Watch does 75% of that already, right, so people call this a transparent society.  That's the stuff of, like, fantasy and dreams and it's the stuff of dystopian nightmares, so what type of methodologies, law, ethics, culture change do we need to make sure we have the means of protecting some kinds of data from ever being seen by a human, some kinds of data ever being seen by AI?  This is a very difficult problem to solve, but I think the first step is to move away from the very binary model we have right now.

        If you look at European law, for example, you have approved purposes, natural disaster, terrorists.  Otherwise, nobody can touch the data, right.  What we're trying to get to is a citizen-led discourse about where to draw the line in particular context.  What's an acceptable reidentification risk in a human GPS mobility data set if that data gets reused without my explicit written consent?  And maybe the question and answer should be, what's the mortality rate of the disease we're trying to stop?  We need people grappling with, you know, the risk-benefit and being in a position to actually determine what they're comfortable with and what they're not.

        >> THOMAS WIEGAND: So I love that idea, self-determination on -- that you can decide what will happen to your data.  Maybe you can click categories, you can say if an -- if a risk is happening, then it's okay, otherwise it's not.  So then maybe make this a little bit more interactive, and every one of you has to show a hand.  Who have of you thinks that by 2050 we will have an overly self-determined data handling in the world and who of you thinks no, we're going to be just giving our data away and people will use them as they wish?  Those are the two options, okay.  So who thinks the first option will be the case in 2050?  And who thinks the second option will be the case in 2050?

        >> (Off microphone)

        >> THOMAS WIEGAND: So you guys are much more skeptical, so more people think that this will go on the way it is and there will be no change to the way data are being handled.

        >> ROBERT KIRKPATRICK: I think events with like what's happened with Facebook recently eventually are going to build up to critical pressure.  That's my feeling.

        >> THOMAS WIEGAND: Facebook has gained more users ever over this.  The effect was that Facebook became more known because of the event, and people -- most people didn't understand what the event was, so Facebook has more users now.

        Anyway, the next speaker -- we have to move on as fascinating as this is -- is Rob McCargow.  Rob is with PriceWaterhouseCoopers, and he works with partners across academia, government, technology vendors, startups, and others, and he wants to drive the innovation within the firm, and I guess he wants to use AI for it, and Rob, please.

        >> ROB McCARGOW: Thank you very much, Thomas.  I'm super proud

to be back here in the ITU as the founding corporate sponsor from last year.  I can't think of a more appropriate mission to fulfill our PwC purpose to build trust in society and solve important problems.

Now, over dinner last night, Steve and I were chatting, and he asked me for my personal highlights over the last 12 months since the inaugural summit, so just a couple of personal highlights in this AI world that I've faced up to.

So first of all, positive highlights.  I had a great conversation with my young kids about the ethics of AI, which I turned into a TEDTalk.  I'd love you all to watch it.  I need to get the view count up.  It's called Robot Rules.  Please check it out.  On the not so positive developments, I had a conference, not dissimilar to this, a bit bigger, where I was introduced as a world leading architect in artificial intelligence insemination.  The wrong thing was grasped, I think.  On a more serious note, I think we have seen this past year a sort of coming together of this global community and really igniting this global conversation across different disciplines from academics and governments, think tanks, regulators, and from NGOs and tech companies.

We've seen an eruption of interest in the huge opportunities offered by AI and the opportunity to start harnessing this to achieve some of our grandest challenges we face, and there's also this enormous significant economic upside to pursue as well, our own analysis from a few months ago suggests that we could see an additional $15.7-trillion GDP growth by 2030, so you can see why people are starting to understand the implications of this.  But it's also focused this harsh glare of sunlight on some of the implications, in particular, discussing on the stage with Wendell, job automation, the ethical (Off microphone), but this is a -- this is a multidimensional challenge, and the scale of which we're not quite faced for generations, and the stakeholder group that Robert just started briefly talking about there that I'd like to speak about today is the world of business.

The greatest impact, I believe, on society of AI will come when the banks and the insurers and the retailers and the health care companies and the automotive companies and beyond start adopting this technology, this dizzying pace of innovation, and a few numbers, from PwC, I'll expect a few.  And we do a CEO survey every January, and the one from last year highlighted an interesting paradox.

First of all, it said 72% of CEOs we surveyed believed AI will be the business advantage of the future, but at the same time, 67% of them believed that AI will have a negative impact on stakeholder trust, and we've seen the precursors of what can go wrong in the past year, and Robert and Thomas were chatting about that then.

It's quite clear we don't have the necessary guardrails in place that AI and enterprise is sufficiently safe and secure nor is it trustworthy and transparent enough, and this has the unintended consequences and potential societal harm, the opportunity to promote and deploy AI for good, can be fatally undermined if AI for bad gains

traction.

We do need a multidisciplinary and multidimensional solution to address this multidimensional challenge. To achieve Safe and Secure AI as this panel's called or as I call it, responsible AI, you need every part of your organization to be engaged in this and be involved as well as every part of your stakeholder and society you serve.

AI will fail if it's categorized solely as a stand-alone ICT project, and, really, what we try to promote as an assurance provider is a methodology that tries to address AI holistically, how is this hard wired into your corporate strategy, how has this taken into account regulation, getting the first batch out to the panel, how is AI going to be affected by that, but also this organizational change concept, as organizations start adopting this, we'll see displacement, we'll see a change in jobs, we'll see workforce implications, do you have aworkforce strategy that takes hearts and minds on that with you?

This goes through the design stage. We'll stress data verification, business readiness, going through all the way that we monitor and operate these systems in business in a way that's consistent, reliable, trustworthy, and giving reliable outcomes that's protected from cybersecurity risks of adversarial attack and the risks of poor data quality.

So, really, this maturity assessment, we believe, is something that's starting to enable companies to have confidence and trust in the technology, to accelerate the innovation of this technology, but to mitigate the risks and avoid unintended consequences.

The next step, which I think Wojciech is going to speak about, is then explainable AI. Many people are asking what happens to my several hundred thousand auditors once the AI and Blockchain is doing the audit? Well, as Wojciech will explain, there's this huge opportunity to become the auditors of AI, so we'll keep thinking down that track.

So to sum up, really, if businesses adopt responsible AI, this will increase the chances that consumers and citizens will benefit from Safe and Secure AI, and with AI safely secured, we can unlock its potential to develop AI for Good at scale. Thank you.

(Applause)

>> THOMAS WIEGAND: Thank you very much. You have written your speech by hand in your book. That is something I have rarely seen.

>> ROB McCARGOW: I always do my speeches ten minutes before panels.

>> THOMAS WIEGAND: I like that very much. Questions For the panel, please.

>> ROB McCARGOW: Not artificial insemination, please.

>> THOMAS WIEGAND: Can you use the microphone?

>> AUDIENCE MEMBER: A comment and a question. I think people had a lot of people coming into this conference had a doubt that AI can be used for good. If they did have the doubt, I think they've seen enough presentations to see AI can be used for good. I believe

in my mind the idea is not about the application of technology, it's about those guardrails you mentioned are missing, and at the regulatory level, the problem is by the time regulation regulations come onboard, it's too late.  As you see, the cookies law that was passed in Europe.  There's so many ways of tracking people now, the cookies law is useful.  You see that every time you get on a website and browsing in Europe.

What can be done by forums like this to actually make sure that for once, the regulatory framework is one step ahead of the technology?  Thank you.

>> ROB McCARGOW: I mean, so clearly this huge has been discussed in the regulatory space, but I can also see Wendell walking in here to talk about the standardization space, the IEEE work, the British standards institute working around the standards of AI, but I didn't address the professional utilization of the workforce.  If you think about the professions in many walks of life, pilots are counted, lawyers, doctors, there are creditors that verify the CPD, the ethical training embedded into the curricula, and the truth is that's not yet mandatory across this discipline in deploying this technology, so I think forums like this are really powerful to professionalize and bring to the fore the necessary level of professionalization to hold people into account and embed that into the curricula so people are aware of the consequences of their actions.  I think that's a very practical step.

Just one other point around governance as well, and I can see Wendell walking in.  I think it's also clear there's not enough appropriate governance in businesses asking the right questions with the right level of qualification for this new technology being applied at scale, is this the right thing to do, have we thought through every consequence, could we incur reputational or commercial damage in applying this?  And I think if there were more effective agile governance and businesses asking those hard questions, I think many of these unintended consequences would be massively mitigated.

>> THOMAS WIEGAND: Sir.

>> Yeah.  I think the way you phrased your questions or what can we do to make sure the regulators can stay one step ahead.  I mean, the regulators are 10,000 miles behind right now.  The cows have already left the barn, and we say how can we prevent it from causing harm?  I mean, the uses of machine learning online today have already proven themselves the most remarkable concentrator of wealth and accelerator of income and equality in the history of the world, so the harm is already very real, and, you know, a few months ago researchers at Stanford developed an algorithm you probably saw, that can take any ten photos of a man from Facebook and predict his sexual orientation with 95% accuracy.  There are countries in subSaharan Africa where it is a death penalty to be gay, so now you've created something that is a weaponized AI.  That's already out there, it's public, so I think the urgency with which we need to figure this out is pretty significant, and we have to get -- we have to figure out how to monitor because with nuclear misuse, you can kind of get those

signals, the inspectors can see something on a geiger counter.  It's very difficult to tell whether some brilliant 16-year-old somewhere is coding something up on -- you know, on their own machine that could be very harmful.

>> I think the lone actor in this scenario is a myth, and regulation has been light years behind technologies and practices because they're dealing with subject matter that they don't understand.  Sorry.

>> ROBERT KIRKPATRICK: I think the consequences are more dire than in the past.

>> I don't believe that either, sir.  I think that's a red herring.  I don't think you should worry about how much the AI can do and (Off microphone) AI is smart enough to pass a (Off microphone) and you'd never know it.

>> One of the things that's been debated is this concept of the ethics board.  Some are transparent or opaque, and there's different varieties of that.  I've had things from DevOps saying I can't be beholden to an ethics committee to look at everything I do, I'll lose the competitive advantage.  I don't know why they sit on that need for boards in places of businesses.

>> THOMAS WIEGAND: Yeah.  To ask that question.  So we need to move on.  Please keep your questions for the end.  If we have time, we'll come back to them, but let's ask a question to the audience.

Let's say the ethics board is an easy question, everybody will say yes, that's easy peasy, so let's ask a more difficult question.  Big data companies develop a lot of things.  For instance, face recognition is one thing or these researchers at Stanford, should there be a self-restriction on technology providers putting out tools that could have malicious use, like face recognition?  You know, you walk through the street, somebody recognizes your face, determines your wealth, decides whether or not to abduct you.  It could be an issue of face recognition.  Should there be some kind of Code of Ethics that creates a self-restriction on those providing it and some of the big technology companies are doing that and some startups and some professors, or should -- or do you think it doesn't make sense to do that, you can't stop progress?  These are the two options I have.

So who thinks the first option is the better one?

>> (Off microphone)

>> THOMAS WIEGAND: Okay.  And who thinks yeah, you can't stop progress, everybody should put out whatever progress they've created, no matter what it does?  Okay.  And the rest didn't understand my two options, I guess.

All right.  So let's move on.  Full disclosure, I'm the executive director of the Fraunhofer HHI, and they're running our Machine Learning Group, so I can say he's the up and rising superstar in machine learning, and I'm glad that he will present his work and the work of his team.  (Wojciech)

>> WOJCIECH SAMEK: Thank you very much.  So my presentation's about the path towards expandable AI and hopefully I can convince

you that if we move forward with artificial intelligence, we need to put a large focus on the transparency issue.  So today's systems are achieved like excellent performances on a variety of complex tasks; however, they are -- they are used in the black box manner, which means that we provide a lot of data to them, they extract patterns from this data, and hopefully provide in the end accurate predictions, but as -- as a symbol -- the black box symbolizes, we don't really know how they arrive at their decisions, why they sometimes fail, and if they're really doing what we want them to do.

And this black box characteristic has certain disadvantages.  For instance, if it comes -- when it comes to trust, you need a certain level of interpretability and understanding.  Imagine like an application in medical domain, in certain domains it's not acceptable to completely rely on the black box decision, so you need to verify it, you need to explain it to the patient, and for that, you need interpretability.  And also, when it comes, you want to guarantee AI safety.  I think you also need to understand how your methods work, why they sometimes fail, and you need to be able to look under the hood.  You see like this example of the stop sign.  So recent work shows that you can very easily fool a neural network, you just need to put these stickers on the stop sign, and the system, which was -- which perfectly detected signs, street signs, will not detect the stop sign anymore, so -- although for us humans, it's quite easy to see there's still a stop sign.

And so when it comes to -- you want to make sure that your algorithm follows the rules or follows some legislation, will you put on it -- if you want to make sure that it doesn't discriminate people, it behaves nicely, you also need to -- you also need to have a certain level of interpretability, explainibility to assure this.

So there are certain implications when it comes to interpretability issues from a legal perspective.  You have implications of social aspect, but like in the other room, when it comes to science, if you apply deep learning, artificial intelligence to answer scientific questions, you need to understand your model, you need to explain, you need some explanation to arrive at specific hypothesis, so if we just have black boxes, this may negatively affect the acceptance of the technology, so my claim is that we need to focus on interpretability issues.

And from the regulation point of view, there has been some progress.  The new data protection regulation clearly states that there should be some rights to explanation when it comes -- when people are affected by a decision of an AI system.  And the good news is that we also have mathematically well-founded techniques today to open those black boxes of deep learning, to open these black boxes of AI and to make them interpretable, so together with my colleagues from Berlin, we have developed a couple of years ago a general method which allows this.  Due to time constraints, I cannot go into detail, but I will -- the method produces heat maps, which is easier, which shows which pixels in the image, which regions in the image were used, the basis for the decision of the system, so you see the AI system

predicts that it's an image of a rooster, and our method provides an explanation for that. It tells you that the system recognizes the rooster due to, like, specific rooster features on the head of the rooster, so you can look at it, you can interpret it, you can visualize it, you can verify that it makes sense or not.

Yeah, so there are methods also, neural networks and no black boxes anymore.

And finally, I would like to present you two examples where we apply our explanation technique and arrive at unexpected results.

So first of all, there is an example from the old days of AI. There has been, like, an image classification challenge, so it's an international image classification challenge, it's universities, top teams around the world participate in this challenge every year. The task was to classify images into 20 categories, like airplane, dog, horse, person, and so on, and you see that, like -- so some categories these teams got quite good results, for airplanes, for horses, for example, and every year these people reported these numbers, but at that time it was not really possible to under -- to explain these models to see if they really behave as expected, so all this research was -- relied on these error rates which were reported.

And so what we thought quite recently is to apply our explanation technique to the best models, the models which won these challenges, so we did that, and for the class horse, we saw a very unexpected result, so we saw that -- I mean, this image, which you see there, was classified as horse. It was classified correctly. But the heat map which we got tells you that the model doesn't really focus on the horse, it doesn't identify the horse in the image, it focuses on the left lower part, and looking more closely to this example, we realized that there was a copyright tag, and then we looked at the data sets which were used for training these models, and actually, a lot of horse images had the copyright tag. They were taken from some horse websites and were -- had this copyright tag, and the method -- all these methods they learned to associate the copyright tag with the category horse, and people didn't realize it for years, they just reported the error rates, and nobody realized that there was like a strong bias in the data. And if it would be like a sensitive application, this would be a high risk that you train the model, you think it performs very well, but actually it does something else.

And another example with were the state-of-the-art models, so we trained the deep neural network, got state-of-the-art of the models for age prediction, age and gender prediction, but I'll show you the example for age prediction, and you see that, like, both ladies were classified correctly as 25-32 in the upper one and the lower one 60-plus, and both ladies also are laughing, and we -- we also -- we visualized these heat maps, and if you look closely at the heat map, then for the young lady, the laughing -- the fact that she's laughing speaks for the prediction, so it supports the classification that she falls into the category 25-32; whereas, for the older lady, laughing speaks against the prediction, 60-plus,

which means -- so you see because there is like a blue color, which means negative evidence, which means that the model learned on this data that laughing -- it associates laughing with young age, so -- also like this example shows you there may be artifacts in the date. Maybe these data sets, old people didn't laugh enough, and the model thinks only young people laugh and associates this feature with age, and this example shows you that there maybe hike these artifacts in the data, and we need to be very careful when we apply AI in sensitive applications, and we need to look under the hood and to understand what's going on and to be able to verify the prediction, and so -- and, yeah, we have the techniques today, so neural networks and no black boxes anymore. At least we can assign -- we can -- with these heat maps, we can see where the important information comes from, and, yeah, I would be happy to discuss with you about this topic. Thank you very much.

    (Applause)

    >> THOMAS WIEGAND: Any questions from the audience? Please.

    >> AUDIENCE MEMBER: Yes, sir. The examples you gave were visual where it was easy to see a heat map. Do you have something equivalent for things that are more textual or language-based?

    >> WOJCIECH SAMEK: We also made it for class clarification. If you analyze a text, you can see which words are particularly important for those classifications decision. We also applied it for EEG data, and you can get meaning interpretations if you know how to read EEG classifications. We're applying it for other data. In principle, you can apply it to all domains, but for us it's very easy to interpret these heat maps visually, if you have, like, some signals, like sensor measurements, it may be harder to make -- to understand what this relevance means, but we -- for EEG, for FMRI, it also works.

    >> And our colleagues recently also found a way to solve this equation, ten to the times power of faster with these methods, and it was also applied there, so we tried to use machine learning as a tool. It's really more for us, we are engineers, computer scientists, physicists, chemists. We see it as a tool to have some progress, and what I'm sometimes seeing is that there is this overwhelming -- the questions are sometimes much too big when you look at what machine learning can do. It's like is it going -- are we going to have a terminator scenario, are we going to model the entire world? No, there's a limit to computationability, we can't even compute a quantum state, so we cannot even model, then, ten quantums or a billion quantums, at least with what we've got, so we have to basically, A, figure out what we should learn on and what data we should use, what are the important ones that actually help the planet and that help the sustainability goals, that help humans and others.

    And then second, once we are there, we should apply similar methods to actually verify that what we learned actually is generalized to the actual problem.

    Imagine you have these artifacts in your credit line approval process and you don't know why you won't be granted credit, for

example, et cetera, et cetera.  So I hope that people who are actively researching AI and machine learning are careful with their predictions of the future because there are these very clear limits. We have a mathematical explanation for these limits when it comes to communication, and we have to have it also here, and otherwise if in a few years the real progress through machine learning in really hard areas isn't showing up, like an equation or chemistry or biology, medicine, medical diagnosis -- I'm still looking for this big cancer breakthrough, medical diagnosis for machine learning.  If those problems don't get solved, then it will suffer from it.

>> I feel that sometimes the problem is in the judging criteria. It's very subjective.  When we looked at those examples, I didn't see any problem.  I think it's -- what you're feeding it and if it's using -- what you're feeding it is probably not sufficient, and if you were to feed it more, then it probably would behave better or maybe you're engine is looking into certain things.  What I'm trying to say here is that from our perspective, we -- we're looking at it, we're like, okay, well, it's able to tell it's a horse because of the thing at the bottom.  That's a good thing.  That's -- you know, if it noticed that in many of the data provided -- we provide it with the data, right, so it's -- so I'm not sure if I actually see the problem from the same perspective that was indicated because I think the judging criteria might be very subjective.

>> THOMAS WIEGAND: So you don't think it's necessary to have a human inspection on what the machine learning is doing?  Is

>> Well, I think the human inspection is subjective.

>> THOMAS WIEGAND: Say again.

>> The human inspection is very subjective in whether it is doing its job or not.  It doesn't have to do the job the way we do our job. It's subjective the way it's doing the job, so my point is if we're subjective in judging it and it's doing a certain job, perhaps we should focus on -- rather than judging it, we should focus on what are we feeding it, and if we're feeding it wrong data, if we're giving it wrong engine to do certain things, we need to address those rather than focus on the judging criteria because judging criteria will always be subjective.

>> Yeah.  The rest (Off microphone).  It used to be the focus in AI development in the '80s and '90s, there's three colors to AI development, right, there's the math model the that's translating to algorithms, the data sense and computing power.  If all three of those are not in line, your stuff doesn't work.  The focus decades ago was on the math models and the model (Off microphone).  Diverse it is, the better your machine learning, so I would (Off microphone).

>> (Off microphone) needs to be investigated, and if that logic --

>> Yeah, but you're not saying anything differently.  I'm just saying those three colors, so where do you look, do you look in the data or the algorithms, and the rest of the world is saying look at the data.

>> I agree partly with you that looking at the data is very

important and may present -- prevent some of these -- some of these effects which we saw, but it's also not an easy problem, right?  If you have, like, millions of images and -- how do you make sure that you are not -- that you get all these biases and that you -- that you make sure that this data is representative, that it doesn't have any flaws?  I think it's also not easy.  It can be one part, but still, I think, in looking at the models, it's also important, especially if you are -- science, physics, certain quantities, for example, cannot be negative, and you -- it may be reasonable to make sure that your model also assigns relevances which, like, reflect the physical laws.

>> It's use specific, but it's not always a model.  More of the data scientists and AI practitioners in the world would argue with you, always go back to your data sets, reclassify, clean it, and put it through again.  That's -- I mean, I'm just saying for consumer-based AI or human-centric AI.

>> THOMAS WIEGAND: Well, we'll leave it at that, maybe have a discussion at the end, but nevertheless, our next speaker is Toufi Saliba, and Toufi coauthored a fully decentralized Blockchain protocol, and he co-founded TODA Network and is the chair of the ACM Practitioners Board conference part, and we just met in New York on Friday because I'm advising there as well, and he is also the CEO of PrivacyShell, a 23-portfolio, and he has (Off microphone) and is having more in the future.

>> TOUFI SALIBA: I'm actually Blockchain now.  Thanks for having me here.  The question I put in, is Blockchain a branch of AI or is AI a branch of Blockchain, and it's -- I'm going to stick to the first part.

Many -- and it's -- it goes to what I stated earlier, our subjective way of viewing intelligence and how we can define what intelligence is vs., you know -- I think there's a certain consensus that we've all agreed on, and it was stated by Alex Weissner, and it's a force that acts so as to maximize the future freedom of action, okay, so if you agree to that -- and we started looking at certain elements and certain machines, what they're doing, are they intelligent or not and how the bar is shifting from a certain thing that's called AI to now it's not called AI because now we have an explanation of how it works, but before it was called AI.

I don't remember the days because I wasn't born then, but supposedly the calculator was considered AI.  It was supposed to get rid of all accountants' jobs.  I wish it did.  Apologies to all accountants.  I meant you guys could get a better job.  And then when it's done, it's not AI, and then we looked at the spreadsheets.  Oh, wow, this is going to be AI.  We looked at, you know, Google, the search.  Are you going to be able to search and find information that quickly?  That's going to be AI because it's impossible, but we can explain it, no, it's not AI.

Now, when I look -- and when we all look at an organism -- and if I were to get a show of hands here by describing a certain organism that can -- that we created, okay, so it's artificial, that we would

agree on, it's capable of defending itself, it's capable to live by itself, it's capable to provide incentives to people here in the room and other people outside of this room and possibly in the thousands, possibly in the millions, in order to enhance it, evolve it, and get to work on it, and it's able to be replicated and evolve without its creator, so if were you to say, okay, is that AI, without giving me a show of hands or if you would like to give me a show of hands, how many would think that that is AI?

Okay.  How many would think that this is not AI?

Okay.

>> (Off microphone) artificial life, no.  Artificial intelligence.

>> TOUFI SALIBA: I think I would agree.  If we were to have the artificial intelligence itself that evolves, would that make it artificial intelligence, it evolves by itself?  Okay.

>> AUDIENCE MEMBER: Can it learn (Off microphone)

>> TOUFI SALIBA: So who doesn't think that this is AI?  Show of hands.  Okay.  We have like 2.38% of the room.  I'm quick at doing that.  I have an AI thing.

So we do have an example of those machines that's been running for nine years and three months and several days, and it's called Bitcoin, and when you say it is Bitcoin, most people say this is not AI, it does not self-evolve, it doesn't do all of those things, but when you compare it to, let's say, a homo sapien, how do you evolve or how do you gain certain intelligence, how are you able to incentivize people and so on and so forth.  How are you able to get people around you to give you money so you can go and buy food, so you can go and do certain things versus the machine that it's able to incentivize people to give it money, to compete for resources that we have, compete for resources other machines have, and so on and so forth?

So perhaps we are subjectively viewing something that is self-evolving.  It is self-evolving.  It's advising over 3,000 people, they work days and nights and it's self-evolution, it's providing over three million jobs, it's got a market cap almost equivalent to the largest company on the planet today, and it will surpass every single company, and I bet my life on it, and it's -- nobody owns it, not a single human owns it, it owns itself.

So that's -- if I were to say this is AI and let's say I were able to convince you, I can still tell you this is a very rudimentary shape of AI, but we need to start looking outside of our subjective way into identifying what is AI, and that's my argument here, what I'm actually bringing.  Perhaps Blockchain is a branch of AI that we would need to incorporate into a lot of things that we're doing in AI for several reasons, and the most important one that I believe in is the autonomous decentralized governance.

So many of the problems that are indicated here in the room, and folks would say, okay, we need to have AI that is going to be, you know, abiding by certain regulator or whatnot or it's abiding by what we would like it to do or what we think ethically is so

important to us.  I disagree.  This is not how it's going to work.

If you were to be the president of a company or managing a certain process and you have an AI engine that might screw 1,000 people's lives or 100,000 people's lives and it's not very ethical, but you're going to lose your job if you don't let it do this certain thing and you're never going to get a job again, what will you do?  Nobody's going to find out it's going to screw 100,000 people's lives.  You don't have to answer me, but deep down inside you, you know that what kind of choice we would make when we are faced with certain decisions, when you're going to look back at your kids and say, sorry, kids, I can't feed you anymore because, you know, I lost my job because I had to make a conscious decision.  I didn't want to screw 100,000 people's lives because it is not ethical.  We will not make those kind of decisions, okay.

So an ethical AI is something that I'm very skeptical of.  We don't think of AI as a skeptical, we need to start thinking from cryptoeconomics what kind of thing is going to win more, and are we incorporating that in the governance of AI by design?  And if you are not doing that, then we should be questioning a lot of things of what we've built in this humanity, are we going to save humanity for the next 200,000 years or we're going to liberate it?  That is the question that needs to be answered and how we're going to be providing governance of AI by design is the solution and not by telling it?  It's not going to respond in future, I can guarantee you.

A lot of folks that I've talked to, some of the top AI scientists, they say, if you're afraid AI, then turn the switch button off or just, like, unplug it.  The entire Chinese government trying to turn off Bitcoin and they cannot, and, yes, it's called artificial life that it's rudimentary, and I'm calling it AI and it's also rudimentary, but an entire nation cannot turn it off.  You're not going to have the chance to turn off AI if it's actually, you know, answering a smaller group of people to do a certain thing that is not necessarily ethical.  I think if it were to be concerned of AI, it's not what it's going to do on its own, it doesn't have the motive to enslave us.  I think we need to be fearing AI if it's governed by certain group of people that they might think that they are in control today but they're not going to be the ones in control tomorrow.  It's going to be a different group of people that the devil within them will act into making things that are not ethical, not the machine by itself.

If you look at the entire history of humanity, the worst evil that ever existed is humans themselves to each other.  If we were to put so much power like we did with the electronic money system that was supposed to be servicing each and every one of you, yet it's used as a method of control, you don't sense it because you live in a world that's fantastic, but if you look at 70% of the world when that money is used as a matter of control, people that they work the entire day to feed their family, they still have to respond to certain people that are controlling them if they can't take that money they've

worked for to feed their family.  It was not supposed to be method of control, it was supposed to be something that simplified people's lives.

If you do the same mistake with AI, we may not have the chance to fix that, and that is my message today.  Thank you.

(Applause)

>> THOMAS WIEGAND: So there's a question in the back.

>> AUDIENCE MEMBER: Hi.  Okay.  This is a question for both Rob and Toufi, then, building off of your points.  So how do we incentivize industry from using a system like cryptoeconomics to -- for responsible AI?  So maybe Rob has insights from working with industry.  So the argument is that Blockchain has incentives and this is what determines the activity in the system.  How do we make sure that AI is responsible and ethical using such incentives?  Because right now I don't see why industry has large incentives to self-regulate and self-govern beyond not wanting to have a big accident and losing market share.

>> TOUFI SALIBA: I'll let Rob answer first because I think it's alphabetical and R is before Toufi.

>> ROB McCARGOW: I'll go in alphabetical order.  I think as we start to move AI use cases from ones of less consequence to ones that have got far more consequence, organizations simply have it in their interest to harness this in a more responsible and ethical way because they will be exposed to risk.  I know we can debate all day long about advanced governance to protect, you know, sort of the integrity of AI, but I think businesses already are very regulated and businesses will hold back of adopting it at scale because they know they're going to unlock risk.  We know GDPR is two weeks away.  There's clear risk of breach there, which is 4% of your global annual turnover.  That would put companies into insolvency.  We've already seen less consequential use cases lead to fairly reputational harm to companies, maybe not all of them, some better than others.

I actually think there's already quite a lot of regulation in place that can be applied better than it is at the moment without overregulating this.

As for the future of how crypto regulates AI, I'll defer to my expert friend alongside me.

>> TOUFI SALIBA: Thanks, Rob.  A lot of responsibility, and I don't know the antihistamine to that.  I can tell you with certainty I've been in the industry since 2001.  I've seen a lot of failures in crypto, I've seen a lot of successes most recently, and I personally came out of the -- what we call the crypto closet in 2016, precisely in June.  About seven years ago, what I did was considered illegal in most of the nations across the globe, so was slavery about 100 years ago, and I probably was a slave.

In any case, what -- to answer that question, what I have done -- because I don't know the answer to -- because I know that we must find the answer, I've created something called AI Decentralized and presented it to ATM.  It's a not-for-profit, and with AI Decentralized, it's geared to collide AI practitioners with crypto

folks.  We have a lot of practitioners and crypto folks, and we're hoping that those questions are the right questions, and if they can get the right answers, hopefully over the next few years, then perhaps we can continue doing what we have started, which is fabulous, which is the machine that can liberate the remaining of humanity's life or whatever left of it.

>> THOMAS WIEGAND: Okay.  Well, we have to move on.  Our next speaker is Andy Chen.  He's a renowned innovator, in the IEEE Computer Society Board chairman, is that correct?  Wow.  That's pretty amazing.  And he's also the VP of Professional and Educational Activities.  He also runs his presidency of a global consulting firm, and I'm glad that we have him, and we started 15 minutes later, so we will not stop 15 minutes earlier, so we'll have the session until 5:30.  It was planned until 5:15, but we started 15 minutes later.

Okay.  So we have 125 minutes left, so the speakers please say what you want to say.  Don't feel hurried, don't feel rushed.

>> ANDY CHEN: Thanks.  Can you hear me okay?  So what I'm going to do is I'm going to make it a little bit lighter.  This is the end of the day, and I'm going to look forward a little bit more what's in the future.

This morning, on the opening keynote, Roger Penrose, who was just here, on the session they asked three questions, the top three questions were -- the first question were -- I have it here.  So it's how do you incent the young professional to not just go after the money but go after the ethics itself, the good of the event

And the second one was talking about so how do you know when he's going to actually reasoning and think, and the third one is saying, so when is he going to be able to get a conscious, when is he going to be conscious?  So those are three really good questions.

The trick is I think I have all the answers.  I want to share with you our use case.  Just go through my viewpoint, and I think Robert -- Thomas did a really good job, so I don't really have to say any more.

If you have ever heard of a U.S. advanced robotic project -- agencies, it talks about three waves.  The very first wave is back in the '60s when people started to figure out there's a problem and they go turning to the code and they actually execute it through the task, like the calculators and others, and it's called the handcrafted knowledge.

The second wave, then, after that AI, we had a really long AI winter because it seems quite a longest time, up until late '90s and 2000s.  A new wave comes out about the neural network, machine learning, and deep learning.  They are now able to do something really significant and playing something really advanced, so they could play game, they could drive the cars, they could recognize your faces, and they could convert text messages and others, so it's very much data driven.  The more data it is, the better it is.  The better -- like we said before, not enough data, they're going to make some mistakes.

Now, the third wave is what they call the AI system where

construct model that would explain how the world works, so this is what the third wave is.

Now, today I want to share with you about the third wave.  We, as a venture capital company, we look around the world for investment in AI and Blockchain.  Me and my partner, we look around and we look kind of -- I look over at IEEE, we have 400,000 members across the globe, we have the top researchers, so I think all are good innovator approaches, but what we thought the use -- the use cases we're going to share with you is one of them that we haven't seen too many around.

So it's called Mined AI, and it's not -- what it is, it's a new symbol -- symbolic paradigm.  It switched completely the other way.  And it uses natural languages, so it could be any languages, English, Japanese, French, any language.  It can do reasoning, and it used augmented topologies network.  I'm going to go through the details of the concept of how that arrived, and it does linear qualitative reasoning process, so it can actually trace back to how it makes that decisions.

How the decision was made when you made a mistake, it's almost like a white box.  You can actually clearly down to the basic unit of where the error is, and it's human understandable logic, something that we can understand.  And one thing I like a lot is it doesn't need a lot of data.  Now, I say -- so how do you make decision with no data?  Well, I think that's the difference that I think this engine is able to drive.  I'm not saying it's the best in the world, but certainly it's different.

As I say, it's a --

>> I'm sorry to jump in, but I have not seen anything better.

(Laughter)

Full disclosure, I'm not involved with that company whatsoever.

(Laughter)

>> ANDY CHEN: Okay.  At least I got Toufi to agree.  Toufi never agrees with me, but today it must be -- what did you drink?

>> TOUFI SALIBA: (Speaking non-English language).

>> ANDY CHEN: So it can do any language, okay.  So let me play a two-minute video about my AI.

(Video played)

>> Artificial intelligence is the new arms race of our time.  With the world's wealthiest corporations and nations hoarding the most innovative scientists, engineers, and researchers, it's deterring our collective brain power from creating problems that may solve global humanitarian crises to merely systems for profit, surveillance, and war.  The AI arms race rapidly progressed us through the second wave of AI in which deep learning from brute force statistical computations led to super human achievements in domains, second wave produce extremely accurate results, but its mistakes can be fatal, as we've seen with recent autonomous vehicles.

To make matters worse, it's extremely difficult to pinpoint what went wrong within these black boxes.  Mined AI was created to tackle the limitations of these second-wave systems and bring forth the third wave of AI, Mined is an artificial intelligence powered by

revolutionary internationally patented data structures to perform
deductive, inductive, and abductive reasoning.  By parsing natural
language and breaking down information to its most fundamental level,
Mined is able to contextualize every detail.  This entire process
is transparent, allowing us to pinpoint and surgically correct exact
points of error.  We plan to open source the Mined engine and
distribute the ontological database to promote a democratization of
power in the AI space today.  Building a community of intelligent
minds around the Mined ecosystem can allow us to turn the tide of
the AI arms race, promote benevolent progress rather than narrow
self-interests, and ultimately advance humanity as one.
        (End of Video)
        >> ANDY CHEN: So you think, so how does it do that?  Well, it
has a pretty simple logical -- the lower unit is what they call a
canonical unit.  It does -- what it does is that it -- that model
there, it allows you to abduct about a subject, an idea, or it can
deduct what it is, and you have an inductive part where it actually
is trying to figure out what's that different -- what is the
difference.  Say a mouse is an animal, but the mouse is with whisker
but it's animal, but it can also do it with a bird, bird has the wing.
        So Mined AI uses the basic canonical models to construct their
database, and when there are different canonical units linked
together, they call them the canonical which is augmented network.
        So this is the reasoning part.  It -- the engine itself, with
the ontologies database, which I'll talk in a minute, it could
actually start doing reasoning, asking questions, devolving ideas.
        The ontology versioning is something that you take the existing
law, like knowledges, say, general -- say -- now we know there is
a general relativity which is a better idea than the gravities.  Now
that engine will go to -- go for the -- if you wanted newer version,
it will retrieve that, but you can also go back to the old versions.
        So my ontology database, it's comparable.  It says that -- you
know, today that engine, it's almost as smart as seven years old.
How do you do that?  Maybe that I'll explain in a minute or I'll have
somebody else explain in a minute.  It can learn.  A year from now
it will be as smart as college professors.
        So what -- one of the things I introduced is called meta
theotics, and this is -- once you've got together all the critical
masses you need on the database, ontology database, you can -- the
machine itself can start devolving theories.  It will come up with
its own ideas, it will think itself, so that is what -- I think what
separates the disruptive part.
        One thing I like about it is that this is completely open.  It's
an open source.  Everybody can contribute to it, and what it's trying
to do, it's trying to democratize so that it's not democratize, it's
not going to go to the corporations, it's going to give back to
everybody.
        So I want to talk about -- a little bit about the ethics that
everybody's talking about.  We talk about all different things.  I'm
in the scientific organizations.  There are quite a few around the

world that are doing the ethical design, and I just want to know that it's almost a global effort.  We have European Parliament talking about setting a fund and having all the smart robots to register, and Stanford had its 100-year study and the Stanford Institute also has these ethical designs, and ACM has these things call the ART, and they were talking about it.

The last part I want to talk about is cybersecurity, AI and cybersecurity risks.  Right now we're using AI to -- we detect fraud on a credit card with speech recognition in the airport and then we do image, you know, processings, and we have driving -- self-driving cars.  All that, I think, at the end of the day, we need to provide some proper training and education for not just the students but for everybody to be able to catch up to it.

So that was -- that was my message to you, but before you ask the questions, I just want to introduce Paul Lee.  He's the -- he's the founder -- one of the co-founders for Mined, so if there's any tough questions, I'll refer back to Paul.  That's it.

>> THOMAS WIEGAND: Thank you very much.

(Applause)

Could we maybe first hear from our next speaker because we are a little late, and I don't want you to not having to -- not being able to say what you want to say, and then when she's done, we'll go back to questions, and we can run over as long as you guys want, as long as the ITU starts complaining, so Susan is our speaker.  She's the co-founder and CEO of MKR.AI, and to give you the picture on her web page for the company, it starts with the following sentence: We live in an era of bullshit.

>> SUSAN OH: That's right.

>> THOMAS WIEGAND: And I think that's enough of an introduction.

(Laughter)

Because that's an excellent sentence.  I like that sentence. So please go ahead and tell us what you want to say.

>> SUSAN OH: Hi.  Can you hear me?  This is working?  Okay. Great.  I'm the CEO and founder of MKR.AI.  I use NLU, NLP, and sentiment analysis, machine learning algorithms to essentially build up a bullshit detector.  Everyone is trying to crack the fake news and false reports, it's always been my premise that interpretation is too wide and mathematically impossible to model, right, because every space you have then has to have a numerical value and you're snowbawling value and storage.

Instead of treating it as a fact-checking problem because as a journalist of 25 years, I knew that nobody really cared about the facts, ive decided instead to do a dip into neural linguistics and was only tracking patterns of deception of which there was a rich 200-year history a study on in terms of propaganda and misinformation.  I take that data and run it through sentiment analysis for the intentional intent to evade, right, and then I grade

it.

And, of course, in that process of training your AI, you have to, of course, take it out, reclassify it, clean it and put it through, and in that process you can actually weight the relationship between the words, do a line by line analysis, word by word analysis, and even weight -- put a differential weight between the relationship between words.

And then from that, I take -- I create original content around it, getting, you know, very famous skeptics like Penn and Teller or actors to look into the camera and say I think this story is 70% bullshit, these are my biases and this is my methodology, and you open it up for anotation for other people to then give -- volunteer and crowdsource and validate their own citations, and what that's doing is basically shifting the conversation from I'm right to you're wrong to this is my methodology. My audience is a very thin margin, maybe 5% to 7% of us that will be moved by reason and not rhetoric.

I started building this in a year and a half ago, and in the last year I came across Blockchain and cryptocurrency. I know that sound a little bit scary. There are people who will tell you they love Blockchain and hate cryptos, and that's pure ignorance. Your every missing the best part of what could be a great opportunity to fractionalize anything that we term value, right, into an asset that's measurable, definable, and shareable.

How that works on my platform is you give me a better citation then my celebrity guest, I reward you with tokens. The machine learnings validates on command. Those tokens you can trade for anything else or access an open source platform that my partner has built to either build or train your own machine learning. It's it's a democratization of AI. The way I look at it, if we're looked at transparency of transactions, machine learning and AI really needs Blockchain and cryptonomics behind it because there's going to be automated machine-to-machine transactions that you're not going to be aware of unless you want to go back over the data because, of course, you know, data's the only thing that doesn't lie.

This is essential, so I'm also going to posit that, you know, Blockchain, which is essentially just a data structure, and unwieldy one at that, needs AI to validate the anomalies as well as the desirable results because there's a pattern to all of human activity.

The only way we can bring this full circle and realize the full benefits of both these nays ent technologies is to give sovereignty of data to each of the people in society today. Everyone in modern cities and modern times, no matter what you do, you're kicking up data that is worth billions of dollars, except there's Facebook and Google and all these behemoths who don't give you back anything for it. They take that data and get to you weaponize it to buy more things, right.

So if you give people sovereignty over their data, they can choose and decide and stratify which layers they're going to volunteer in to your system and you can reward them with tokens for it. It moves faster and is more effective than regulation because

regulation is a two-step fold.  It is a law if, A, you can't enforce it, and, B, you don't know what you're tracking, right, so instead of having it be a twofold action, why don't you have a sentence, give people sovereignty over their data, give them their choice of being able to share that data, tokenize that model, and be able to reward them for things they would use anyway such as airtime on their phone or be able to, you know, pay off parts of their electricity bill, something that they use all the time?  Those examples that I've given you are actually used in different parts of the world today, right.

You can even tokenize -- I'm working with a group called My Far Global out of Zurich here in Europe.  They're tokenizing the future IP of the top scientists and neurologists in the world today towards the development of general AI.  The reason why this is -- the reason why this makes for safer AI is because it's more human, it understands what human interaction is, what human experience is, and all -- it creates a working framework at a market whereby researchers and people who've devoted decades of their lives can monetize parts of that IP in a safe and trustless transparent system.  That doesn't occur today in either academia or large parts of the market.

So Blockchain and cryptos allows you to -- allows anyone to really open up a trustless system where it makes it an open inhave I tatings to collaborate -- invitation to collaborate and incentivize them, where it isn't in AI today.  It is the new arms race and data is the new gold.  We've gone from commodity-based economy to an information economy, and then Web 2.0 gave us the knowledge economy, right, the value is in who you know and what you know, and now you're looking at a data-driven economy.

The more we recognized this and create systems for collaboration instead of forced regulation I think is the better way to go.  Thank you.

(Applause)

>> THOMAS WIEGAND: So are there questions for Susan?  Please.

>> AUDIENCE MEMBER: So just on the onset, 100% agree with all what you said, but the place where I'm skeptic about is we talk about democratizing data and giving people power.  Do people really care about it, as when mark Zuckerberg was giving testimony in the senate, his value went up $9 million in those six hours.

>> SUSAN OH: Those markets, by the way, are corrupt.  I don't believe in stocks.  The only thing I believe less than stocks is probably statistics, right.  I want to go back to the data in the space, and all of that is driven by sentiment and splash.

But to your point, I think people do care.  I just think they're walled into a sense of apathy where they don't realize what they can do about it, and the way that -- you know, it's like when you show someone what could be, right, and you show them how much they can make and you show them by giving them a token and showing them how a system works, then they can't go back to it because you can't unknow something that you know can work better and can work better towards you, right, for your own benefit, so I do believe that people care.

>> Just further more to that point, beyond the caring of people,

what I like about that model, it introduces economics into the equation, and if at the end of the year you're generating, let's say, $400 of data, that you choose to sell it to Facebook at the end of the day, then Facebook may not necessarily have that money to buy it from you because you're the one who is making money, so if you're equating all of that, then you can bring that sovereignty to the user or take it away from them by paying the price for it. It's much fairer than the data we live in today.

The other thing I like of what Susan said is the weaponizing it, and this is something that not a lot of people are paying attention to when it comes to the invasion of the privacy and confidentiality of people. It's more like the frog in the boiling water. We don't notice how much it's taking away from us until it's all gone, and privacy is part of liberty, and if that is taken away, we may not necessarily have a lot to live for.

>> THOMAS WIEGAND: So to complete my job and then we'll have the free part, is there a question for Andy? Please.

>> Who is Andy?

>> ANDY CHEN: Yours truly.

(Laughter)

>> My question is not for Andy, it's a comment for the lady. I do share your sentiments in the way you have actually approached the way in which you could actually solve the issue of data protection, but I would like you to -- I would like to add that it cannot be one method at the exclusion of law enforcement. I'll tell you where I come from. I'm a medical doctor. To this day, the best place to conduct trials, it's in Africa because you can't go cut a branch and nobody's going to stop you. You can even do the worse thing that you can never imagine because Africa will have no protection.

So if you use that level of self-policing and protection, Africa has got a lot of people that are illiterate, and it will be for everybody, including your system. Everybody who sees that your system has actually blocked access to data, they'll go to Africa, and I speak on behalf of Africa.

>> I actually -- I want to jump in here. It's a very good point that you bring, and actually, I use that quite often when talking about a lot of the solutions that are being presented to Africa, including electronic money system. Every time it's presented to, let's say, a country like Madagascar, hey, we're going to ask you -- include you into our banking or whatnot, but in -- the people on the ground, they say $2, when you park it to your neighbor, your neighbor has $2, you get a cup of coffee. With every electronic money as many that's been used, it's going to extract through every single transaction. Nine other people get paid outside of that economy, and if you quantify that, 96% of the money gets extracted out in three years.

So that's only thinking one thing, and that's to your point, and it's called electronic money system, so if you take that and propagate it across, what do we need? We should start asking what

do you have?  You have the people, you have the network, and you have smartphones.  If any solution's going to need an additional component to those things, you should say no, and that's my simple answer.  Sorry about being super passionate about it.  I am, and that's what I kind of live my life for.

>> THOMAS WIEGAND: So, yeah.  Let me pick on that.  If there are no more questions for Andy, then I would like to pick on that.  I'm -- I guess I'm old fashioned in my way of thinking what progress is.  I think progress should be making better food, using less water, using less energy to bring you from A to B, loving your children, loving your spouse.  These things, if you can improve on that, I would call them progress.

First question, do you agree that's progress?  Second question, how do some of the things that we discussed today actually contribute to that old-fashioned progress?  Because I want to be sure that we are leaving this session with a AI for Good and good being defined as something we can agree on thing.

>> SUSAN OH: Well, automation is going to be -- see, machines are meant for productivity and people are meant to be human and creative; correct?  So it's -- with this mass amount of automation, we're going to have a great deal of productivity.  With that productivity, why can't we make sure, hopefully through Blockchains and cryptonomics, and smart contracts to Airdrop people these tokens so they have a minimal basic income and they can decide are you going to then turn more human and humane and do things that machines cannot possibly do because food cooked by humans tastes better, nobody will be able to rehabilitate and raise children better than human beings for at least the next ten years or so.  I really believe that.

(Laughter)

Or are you going to become a steward of the bot and create better robots and systems and people, so all this productivity has to go some where, and that's not a technological problem.  Tech is agnostic.  That's a human problem.

>> THOMAS WIEGAND: Over there.

>> My question is for Andy and maybe for all of you.  Today I'm in a mind that AI brought something else.  I mean, we call it distributed AI, and when it comes to distributed, I mean, Blockchain has been there for many years, and what Blockchain brought and especially Bitcoin, we failed, actually, because it was after 2008 -- I mean, the latest financial crisis, and today we are failing the world.  Why?  Because all the value that -- I mean, virtual values, I mean, are true cryptocurrencies that we have created, I mean, billions.  They are concentrated in a few hand hands, he and we can manipulate the market very quickly, and you can see that, so when it comes to distributed AI, how can you learn from Blockchain and the token, et cetera, not to avoid but to mitigate that risk, because combining Blockchain and AI and on top of that new technology that is coming is -- it will accelerate the decision-making process.  We can do more harm than good.

>> Can I jump in here and -- this time it's not alphabetical,

it's Andy, Toufi.  When you look at the existing Blockchain and we look at Bitcoin, we need to take in effect exploitation that happened on Bitcoin five years and five months ago that not a lot of people paid attention to, and that created something called minors and a company called Bitmain.  These companies do one thing, they burn the planet to create a certain thing that is called cryptocurrency.  It is not intended to do what it's doing today.

The initial intent was to provide each and every human the ability with what you have in your home, your computer to do that mining, but instead, it started evolving certain parts of that machine that was not expected to make it one single machine will have hashy power a more than entire Europe combined, and you can go, just, like many of those machines are competing and start getting something that's like, hey, I've mined these coins and I've burned half the planet.  Currently Bitcoin consumes more electricity than the country of Iceland multiplied by two.  Bitcoin is on the chart of nations, number 153, more than 152 nations in the con -- consuming electricity, and it does only five transactions per second.  It was not intended to be like that.  It was intended to be in the hands of everybody, and it's that exploitation that got it where it is today.  There are solutions to that, there are solutions in effect, and I think many of the guys are looking into benefit from the Blockchain, especially, you know, if you're looking into all of the research and all of the things that are coming, they would be using things that are back to the intent and not of that exploitation.

>> ANDY CHEN: So as I said on my slide, I think the most important part is two things.  It's -- especially in the kids on Mined AI, it's open and it's decriminalized.  It actually gives the power back to the people, and the whole thing about the ICOs and the Blockchains or what it is, it will fail if you do not have a good ecosystems.  If you don't have everybody uses it, it doesn't matter what you do.  At the end, it's the people that -- the ecosystems is actually going to fail what the success is, and I believe that to build an ecosystem, it's got to be something that benefits to everybody, it's open to everybody, it's accessible to everybody, so that's got to be the first step that we do.

If we start using a big company and trying to come up with any big standards, I don't think that's going to work.

>> I mean, the history of technology, you go all the way back to the invention of the club, right.  The history of technology is that when something is invented, it looks like a great leveler of the playing field, and then you fast forward a certain number of years and it always turns out whichever individual or group was in the best position to move fastest ends up with a wild asymmetry and everyone else is left behind.  We see this -- we go to one place now for books and one place for Search and one place for auctions when the Internet -- when the worldwide web was created, everybody said I can have a web for my kids with Bill Gates.  Everything we do is being scrutinized and used by various parties.  You know, you look at Blockchain and you say everything that we do going on the Blockchain

is just another way to have companies and government have even more robust evidence of what we did that they can use in the same way because -- from one perspective, Blockchain is just another database, but it is also the first technology I've ever seen that I think has the potential to change that equation, but maybe that means that everything for now the AIs do has to be on a transparent and public Blockchain, and humans should have the ability to work on a private chain.

What we're seeing now is completely anonymous networks that can be used, eventually not only for transmission of value but also execution of smart contracts and distributed computing. The cifher punk movement is coming back (Signer) and we're going to look at -- as we have as a society, looking at these tradeoffs, privacy is the foundation and guarantee of our liberty. It is also the price we pay for our liberty, and we have to be willing to live in a society where people can be completely anonymous, to hold power accountable, and also in the process empower terrorists. There's no way you can have both, but I think that Blockchain is the first thing out there that really could play a role in making sure that AI is monitored in a way that can help mitigate some of --

>> Sorry, I want to jump in here because you've mentioned Blockchain and you've mentioned database, and that goes back -- we should give credit to Deloitte because they coined that term "distributed ledger technology." But in reality, theres not a single Blockchain that is distributed out there, and also, Blockchain is not necessarily a ledger, but the good part Deloitte, they got the third part correctly, the technology, so we should give them credit for that, but that ended up creating a lot of confusion with folks and started talking about on the Blockchain as if it's a database. It is not a database. The Blockchain is a block-off time of events and they are chained together. They can be on the ledger, but they don't have to be.

The intent of the Blockchain is also to be open and not in the private. If you were to have any single Blockchain that is in the private and you bring in a smart database developer, they can build you something that is better an more efficient, so there's a lot of things that started when Deloitte came up with the DLT. I hope nobody's here from Deloitte. Please don't take offense. Whoever came up with that, I kind of know them personally and I said it to their face. They made a mistake. They admit to it right now, but it's already propagated across a lot of folks around. The

>> THOMAS WIEGAND: So we have to come to the end, so let me go from over there, start with Rob. You have three sentences, and you can say them about anything you want, three sentences. Don't try to make them too long each.

>> ROB McCARGOW: The Blockchain for Good Global Summit next year, possibly. We've spoken about Blockchain more than AI in the last half hour, haven't we? I think there's a great opportunity in the next two days ahead to start moving this forward. I think there's been a lot of long-range discussions today. Some of the discussion

today is too long-range for harnessing this this year, and I think Blockchain is one of these crypto discussions that might take some time to mature, so I'm concerned about talking about that.  What can we do the next few days to move this debate along?

>> THOMAS WIEGAND: Okay.

>> WOJCIECH SAMEK: So I'm engineer, and -- but I believe that AI can help us to -- to progress, to find solutions to some of the problems we are facing, and I'm pretty sure that if we meet next year, we will -- we will move forward on -- along some of these lines.  Thank you.

>> SUSAN OH: We heard a lot of gloom-and-doom scenarios regarding artificial intelligence.  That's not what I believe.  I think there's more of us who mean well and who are actively working every day to create better things, better experiences for human life. So artificial intelligence can automate 80% of the work that we do and all that productivity can be pulled by machines, as long as we develop it transparently and responsibly and building business models around it that necessitate collaboration.  In this, every city in the world can now become a living lab between artificial intelligence and tokenomics.

>> TOUFI SALIBA: I just needed to mention something I've started with and the title of this event, AI for Good, which I've had huge issues with.  Not anymore.  Thank youfully to fabulous organizers that they were able to realize that good is subjective and trying to do things around it.  If we look at wars, each and every warrior from both sides, they say we're doing it for good, but that good is not equal to each other, so how is it different from 400 scientists going to be sitting together in the conference and determining what is going to be good for the future of humanity, and the organizers kind of realized that and kind of invited me here, so thank you.

(Laughter)

So anyway, that was the joke part, but it's -- in reality, they do realize that it's more around governance of the machine and having the machine -- managing the machine rather than just having us being subjectively judging that it's good and -- yeah, and thank you.

>> ANDY CHEN: Yeah.  So I was quite happy to hear this morning, one of the keynotes talked about the ethics, the ethical approach, and she said, really, it started with doing a good thing, something that's good.

Now, I think in this summit, I think it's -- the onus is on us to find the good for all of us together.  Something good for me, something good for you, something we can build something together, so I'm looking forward to this summit.  I think it's pretty exciting.

>> ROBERT KIRKPATRICK: So just a sort of pragmatic note.  We -- in the environmental law, there's something called the precautionary principle, which essentially says that before you create something that could have a cascading catastrophic effect on an ecosystem, you have an obligation to stop and model the risks provably and develop mitigation strategies that are sound.  You have to be able to show that you've considered these risks, and my sense

is that we should probably look at that as a model for how we think about AI.  You know, if we -- it doesn't mean regulate it to death and stifle innovation, but it means having criteria for knowing when we need to pause and think things through, when we have a reason to believe that whatever we're considering creating to meet a market need could have a negative effect from a human rights perspective.

>> THOMAS WIEGAND: So I want to thank the speakers.  I think that the world needs lots of improvement.  We need a better financial system, we need better ledgers, we need better production, we need better handling of immigrants, we need democritization of AI and we need many other things.  And we have a lot of work to do.  Thank you very much.  Give it up for the speakers.  Thank you.

(Applause)


(Session concluded at 5:50 p.m. CET)