```
                       FINAL FILE
                AI FOR GOOD GLOBAL SUMMIT
                      9, JUNE 2017
                        ROOM C
           PROMOTING EQUALITY FOR ACCESS TO AI
      Services provided by:
               Caption First, Inc.
               P.O. Box 3066
               Monument, CO 80132
               1-877-825-5234
               +001-719-481-9835
               www.captionfirst.com
                                    ***
```

This text is being provided in a realtime format. Communication Access Realtime Translation (CART) or captioning are provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.

```
                                    ***
```

(Audio pause)

>> Something I think about this idea of the perception of Human Rights being something that is adversarial. Just at the same time, you know, in the conference the last two days everyone is talking about ethics, justice, about equality. So, I guess the general purpose, the general principles of Human Rights seem to be well accepted and well, everyone is striving for that.

Where do you think the issue is? Where do you think when you start talking about Human Rights as a framework, then there is that resistance?

>> I think the resistance has a cost implication, people think it is just too expensive to build in all these fuzzy things.

I think the other really huge issue is around standards, because there is sort of like a parallel sense in one sense you're dealing with a very technical mathematical computational universe vocabulary, on the other side you are dealing with warm and fuzzy abstract ideas, even when reduced to human right standards, they're still not computational standards in the same

way. They're still fairly abstract and involve a lot of
interpretation. So, I think there is this idea of how can we
mathematically engineer for Human Rights standards. A hard
problem. I think it can look different things, transparency and
discrimination. I think standards poses a really, really huge
issue here saying, at a very abstract level we think ethics is
great. Of course, we want to build it in. How do we actually do
it? That is a much more complicated thing and that is where I
think you see some sort of resistance. But it is going to have
so much benefit. Yeah, there are going to be a few small issues
here and there for a few small marginalized people. That
shouldn't stop us doing good for the whole or for the average.
And, I think if you are not achieving Equality for everyone, then
you have sort of failed.

The middle class isn't going to be necessarily effected, you
know, rich and educated people aren't going to be effected. The
whole points of Human Rights standards is to protect the worst
off, and that (Audio pause)

>> Is this working?

>> Yes.

>> Thank you very much. We are going to go through the
funnel now and focus on the first questions around ensuring non-
discrimination kind of reducing or emanating bias and Equality
and Human Rights more broadly. And, I wanted to kind of -- go to
Peter Lee perhaps, Microsoft and computer sigh insist with huge
amount of experience.

How do you see sort of that issue? I mean from your
perspective, the computer scientist, how can it be resolved and
what are the barriers? Especially some of the things around the
difficulty of implementing standards and qualifying those in the
technology.

>> (Audio pause)

>> The one thing to say is today AI practice today is
largely based on machine learning. Machine learning today is
largely based on data, the direction of insights and model and
data. That data today is largely the exhaust from human thought
and activity.

One must understand, that's an exceptionally narrow flight
of the technical possibilities in AI but given that that is what
is what we have today in AI, then fundamental question is how do
we recognize, have sensitivity for it and react to bias says that
may exist in that digital exhaust in the data itself. And, as
Lydia had mentioned in her remark, actually the scientific
community, research community, has studied mathematical
foundations of that problem extensively and continues to study it
and has developed as a result of this and machine learning
technique that are able to deploy properly control for that. For
example, the one example from Microsoft is for the last several
presidential elections, Microsoft has had -- this is not well
known because of the popular 538.com, but Microsoft has had the
most accurate prediction state by state and county by county of
the election result, presidential election, and interestingly,
that data that is used to make this prediction was obtained
almost exclusively from following the game console.

     You might wonder, how can that be? Since something like 87%
of X. box game console users are men between the ages of, I don't
know, 16 and 23, something like that. Cross-section through
multi-dimensional data you are able to develop the model that is
able to essentially control the demographic mix of the data that
you get. And, for the (Audio pause) implies that in practice
such care isn't always done. In fact, it makes sense, because
you need a large number, the more dimensional data that you have
control for this. But I think maybe one thing to say is that to
reinforce that there is a science to this, but perhaps a gap
between the science and the actual practice.

     And then the second thing I want to say is now you could
just allow me to dream a little bit further ahead now. Let's go
ten years plus. If we are successful as a technology and
scientific community to achieve more general artificial
intelligence, we have no reason to believe that AI in
the future is going to be so dependent on data, especially with
data that is produced by people, by human thought.

     If you think about what happened today, the fact that AI is
being trained on data for human thoughts is actually a
limitation. It actually places on limitation on the power of
today's AI to the power of human citizen thought. But in the
future, we can imagine more general forms of our (Audio pause).
I think that is an interesting question how we would think about

question of fairness in the more fundamentally non-human form of control.

>> Thank you. So, if I can think, so essentially, I mean, my, so the mid-, perhaps the science for correcting for bias is there, and that the main difficulty is essentially it takes time and money to make sure that that has happened and that is not always the case. To what extent, I mean, I can imagine it is very sort of very difficult to answer that, but how often is it that AI kind of AI development does that, takes the time the do that? Is it something that happens regularly or is it the exception rather than the rule?

>> I don't know. So, I think one of the hardest -- right now if you look at the technology community, whether you're in the University or in Microsoft, data is extremely valuable and getting access to data and (Inaudible) is a privilege and extremely valuable and can also be extremely difficult.

If you are trying to build a speech recognition system, there is tremendous amount of effort (Audio pause) people end up being fairly (Inaudible) and they work with what they have.

One other thing I would say also is, not to embarrass myself, but I think I'm not alone, because I think every technology company is in the same boat right now. We increasingly do have learned systems around the world that are learning dynamically around the world, and that is generally a good thing, but it also creates lot of embarrassing situations. We have situations just a few months ago where our translation system learned to translate the word Saudi Arabia into Diash.

Some of you might remember about one year ago my team had really but social chat pod on Twitter and learned some very unpleasant things from people. And, so, there is also there is something fundamentally risky about learning from people and also fundamentally limiting about learning from (Inaudible). On the other hand, the fact that our systems learn from people make them more human (Audio pause). Alpago is not plain gold. It is really alien form of play. And, so, I wonder how these issues will look when AI feels and looks.

>> Thank you. Go to Vivian. You're a scene your research officer in the Human Rights Technology Project at the University of Essex, and from your perspective, what are the key things that you are seeing, either up front from a perspective happening or

not, where are gaps are that you see at the moment, as well as any other action.

>> (Inaudible) so the question should be good gaps and some operation to lay out the context here which my colleagues and I have been observing from (Inaudible). It has already been projected in previous sessions and beyond that there is a need to think about how the design of technology, be it artificial intelligence, machine learning, and so on, and then integrate some principles and values as a safeguard. And, we see this in privacy and constant technology, et cetera. And, we share the optimism of some (Inaudible) artificial intelligence operates opportunity to achieve other goals as well, and perhaps automate a decision in processes that can help take away, neutralize or help make some decisions that you make more accessible.

But, as another speaker has highlighted in this morning's plenary, has just said as well about deliberation of human thoughts based on that, that perhaps no matter how well designed was from this standard, technology is only a magnifier of capacity. And, to recognize. As much as I share that optimism, as that's said, that artificial intelligence has great potential and (Inaudible) in helping us achieve the expertise and improve our efforts in global problems and it can help them protect and naturalize our rights. I'm also very confident of the pitfalls that need to be very cautious and confident so that we can fund this one potential on the one hand but be able to prevent and address the risk that we have been observing, therefore I would consider a few things that I think we need more thoughts on perhaps to the gap we've been talking about.

Even the application of AI for good or not I think need to be thinking about even when the application that we're talking about are for likes and kinds and purposes, as (Inaudible) had said earlier, just because we have certain mathematical robusts of what we've seen in proportions, trying to paraphrase what she had said, that there is a real risk that automated decision making can also introduce or empathize by it and make a problem worse instead of (Inaudible) bias data or if we have historical prejudices, we enforce and perpetuate that that data is the application. For example, if we see a social interest, historical data that targets directly or indirectly and we face new direction on that and it reinforces the problem, but then what we need to do is think about, well, how can we address that.

So, I guess the question really is how can we deny the valid and apply and monitor AI in such a way that it allows us to present, address, and mitigate potential discrimination. And, this can be done through internal and external (Inaudible) that are based on rights, nor imprintables that we've seen on Human Rights. And, thus bring back to what (Inaudible) was saying earlier about having some sort of IRD for companies, and argue if you want to think about how we can teach AI or teach AI to teach itself (Inaudible) as we said earlier to come up with new solutions, then how can we do that? I think it's critical that we think about conducting Human Rights and tech assessments at each state designed for application for monitoring of the use of such technologies. And, it should be guided by a consciousness of the fact that it is not just discrimination, but our rights are indivisible, enter de-pen dabble and interrelated and if the risk of freedom of expression direct to (Inaudible) and I think we really need to think about the full wind of life that could potentially be implicated, because as (Inaudible) impact or not, otherwise can be effected and I think the challenge here is people.

What is the full implication and how do we assess that? How do we look at the connections?

On the other hand, what are these direct and indirect elements of discrimination, because if we can detect something that is explicitly bias or discriminatory, that is all well and good, but what about the unintentional effects?

So, I think availability of the data that drives AI and technology also drives this paradigm shift that in a way that our rights have been effected, and I think it's really important that we consider how -- how can and how should we address such direct and indirect discrimination.

And aside from identifying technical problems, how can we ensure a (Inaudible) in this context. To secure remedy whose individual rights are effected.

I think what Malavika said about the solutions to mathematical achievable ones is really important, because that is something that my colleagues and I spend very much to work on this approach because it's not enough to just look at it from the view of law alone or (Inaudible) alone. But I think an intersection of successes to identify real solutions moving forward. And, I think in the challenge of trying to look at

unintentional indirect effects and trying to secure remedies for discriminatory effects, we have this problem of what if the individual is not even aware if it is unintentional. It is part of the data set but the positions that were made about them was based on something that was not based on the data sets. And, what if the part is to approve and distributed and aggregated and what about the complex human relationship or state person relationship, the non-active relationship.

So, I think in some way we may need to think about how a Human Rights based approach can be implemented so that we can harness the potential while managing that others have been talking about as well, and how can we base that on international Human Rights principles and integrate that throughout the process of the start, designing, developing, and the point where we use technologies. And, I think we've acquired a multistakeholder approach that this forum is trying to view as well, have this variety of perspectives. So, I think we really have to apply both these general principles, as well as find active solutions to specific challenges within different sectors, different stakeholders, and be able to translate these principles to express and actionalize an effect.

>> Thank you. You give me I think something that you said that I'm going to, was the question to the audience in a moment around what happens if based on the data that we have you cannot actually correct or advise with what you do is application. So, I'm e going ask this question in a moment. Thank you very much.

So, our last part is Frederike Kaltheuner, Human Rights with Privacy International. If you could, perhaps, you know any response of course you have on anything that has been said, but also maybe from particularly looking at those gaps again. So, the gaps that exist now and what you would like to see happen instead.

>> FREDERIKE KALTHEUNER: Thank you very much. Also for the invitation to be here today.

I very much like what Lydia said at the beginning. The example of (Inaudible) shows we're faced with a very difficult moral dilemma. It shows that it is very appealing to use technology in cases that we were very much unable to tackle as a society today. However, you said something very interesting, as well, and that was we should use these technologies in a way to involve people and in the full awareness how the data collected

and reads. And, Privacy International we are looking at AI in the context of a data environment where this is fundamentally not the case. So, I'm a bit more pessimistic about this being able -- about this being possible to be realized.

Let me give you an example. So, one project I've been working on is the use of Internet of Things, data as evidence in criminal proceedings. This is quite fascinating. There was a murder case when Amazon echo, the police wanted access to data that may or may not be on the disease on the Amazon server based on the assumption the device may or may not have recorded voice doing an ongoing murder investigation. And, what is so fascinating about the use of IOT data as evidence, it shows how most people are fundamentally unaware there is a fundamental information (Inaudible) with what kind of data is being generated, collected, processed, shared, and what people know, and furthermore what people have access to.

While I'm speaking here, my colleagues on cords in the UK challenging bulk data collection and government surveillance, but not just in the private sector that people are unaware what is happening to their data. The same is happening in the public sector, and I very much like what you mentioned the word digital exhaust that this is what machine learning is currently relying on, so given that, most of machine learning is relying on digital exhaust and the word almost implies that there is a fundamental lack of awareness about this being record, what kind of inferences, predictions and judgments can be made from that data. Even that this is the current case, I find it a distraction to talk about the future that is not yet there. About intelligence, about singularity, when even the very limited machine learning that we use currently is so unregulated, so out of control.

So, the two -- AI is also very broad term. A dangerously broad term. I think I said this yesterday on the panel. What are we actually talking about? Are we talking about robotic? Singularity that may or may not ever come, or talking about machine learning to make decisions? And to think that we as a privacy organization or as a civil liberties organization are interested in is the way in which particularly machine learning from the (Inaudible) is used to make inferences, predictions and judgment about people. So, the fact that we can learn from data such as phone based network you can make inferences about people's personality. And, I think people are very well aware

with their share online to understand what kind of inferences can be made from seemingly mundane and boring data.

The second point we're interested in is how these inferences and other sources of data are used to make decisions. Get a job in the case of employment, in policing, these decisions may or may not be automated, but this is -- those are the two scenarios that we are interested in.

And, the point that I would like to highlight is I think the harm is unequally distributed, and for example if you look at the case of hiring software. Unfortunate currently that when I apply for a job I get an interview, I can explain myself, I can put data into context.

If you apply, at least in the western world, if you apply for very cheap neighbor in, let's say, in the food sector and (Inaudible) sector, this process is being automated and data driven. Which means there are certain people that are in a position to question judgments and explain themselves versus the bias of discrimination and automation and data use effects those who are already marginalized disproportionately more. And, sort of like this unequal, the unequal effect of discrimination and harm.

You also see this in the case if you compare how secure operating systems are, if you compare the more expensive iPhones to Android operating systems that I use around the world that are cheaper, they are much less secure.

I think I forgot what the actual number was, but there was a study for a couple of years back that showed that Android operating systems are subject to up to 11 vulnerabilities at any point in time. This is concerning. This means that their data can be accessed by government, by police, by unauthorized parties, by jealous, stalking ex-partners, et cetera. And, these harms are not distributed equally.

There was a quote, I forgot who, I think Peter said, the software to eliminate bias is there. I think the hiring software case also explains that or illustrates that a lot of time when we look at discrimination we look at traditional markup, then the race, ethnicity, sexual orientation, but one thing that we do not measure, because it's very difficult to measure, is economic background. And, I'm a bit worried, and very much a fan of the accountability experience community that it actually works and it shows that technologies are very concerned and willing to make

this better and fairer, but this are things that are very difficult to measure and quantify.

We can have -- we also need a discussion about what do we mean by fairness. What kind of society do we want to live?

And finally, since we focused on these micro things, we focus on inferences, predictions, decision making and to fairness and discriminatory nature of these, sometimes easy to forget the bigger picture. And, I think there are three big developments that we have to discuss access in NAI. And, one which is the provocative term is that we're living in a form of surveillance capitalism. The only thing what this means is to say there are companies whose primary business model is personal data. To make predictions of these data, to use these data to sell goods. And, in this ecosystem, AI is something that is added on top to make it more easier to make predictions from data that will make it more easier to nudge people to buy things and do things that they don't necessarily want or that they are not necessarily beneficial for the greater good.

And the second one, I think, is we are seeing a trend towards predictive knowledge and sort of like we are very eager to make risk scoring. We want to know things before they happen in advance. I think in 2005 it was revealed that the Obama no fly list is based on predictive knowledge, not actual evidence, and the same is happening on predictive policing. So, the police are no longer just predicting areas with a high propensity of crime, but assigning individual risk towards the people in some areas. I think this deserves a more general discussion, do we want a society that relies on predictive knowledge? I think there are cases where this is wonderful and very beneficial. There are also cases, and I think terrorist watch lists are some of them where we might not want to rely on predictive knowledge.

And the second, or the last larger trend is the trend towards personalization. AI or machine learning is a technology that has contributed to the fact that some, not all, online spaces are highly personalized towards our assumed interests. It's difficult to discuss the implications. This is a complex issue, but as we're moving towards a software becomes ever more embedded into the real world. I can imagine that we also will see personalized spaces off line. The world is going to be full

of fences (Inaudible) that make I can decision in realtime and this will raise some very important questions about what kind of society will that be? Will this be a fair, just society? What kind of intentional, unintentional accidental tools of discrimination of harm and resolve?

>> SHERIF ELSAYED-ALI: Thank you, Frederike.

There is so much to talk about. We can spend a few days on this. Okay. So, this is what you're going to do. We're going to switch a bit format slightly. Going to go into very lightning format.

So, I'm going to ask a question, a couple of questions to the audience. The first one is in cases where currently or in the future if bias cannot be corrected because of the type of data that there is, because of where the eventually to invest how to implementation can happen, in things that touches people's lives directly, so whether it is bone decisions, policing in the sentencing decisions in the criminal justice system, would you say that we should go ahead, then, use AI and sort of improve on it over time, or not use it until we can get this right?

Yeah, interventions.

>> Yeah. Whoa. Okay. Okay.

Hello. Firstly, that was a fantastic panel. I think it was the best one that I've seen at this conference, so thank you very much for all of that.

My response to that, I think, would be coming from the medical world is that we're often faced with decisions about, you know, technologies or medications or therapies or things that were we think might be good, maybe in this clinical trial that we've tried with this very specific population, and then we have to make decisions on a cost benefit bases. Forget economics, cost benefit for the patient. Is this going to help more people than it hurts? How many people do we have to treat for this to be effective? And I think that one way to approach this is it is not an all or nothing or shouldn't be an all or nothing model. It is not do we use it or don't we use it. It is if we have a model and we think the benefits might be very, very high, can we come up with a regulatory structure to test it in the real world in a safe way where we can mitigate the consequences in a timely

fashion like we do a clinical trial, see if it works or not, and then decide whether or not to scale it up. We don't have to abandon the technology, but we do have to be responsible about how we implement it. It doesn't make sense to have a global implementation on day one. Just like it doesn't have a success for Lipitor is going to be day one for everybody. You try it over a significant period of time until you're sure that it is safe and effective.
        >> SHERIF ELSAYED-ALI: Thank you very much.

        Any other thoughts on that question? Please. I think if you keep your hand up so they can activate your mic.
        >> Nit from University of Victoria.

        One of the things I feel it keeps to me, and I agree this is a fantastic panel, thank you all, I learned a lot, is there is a sort of tendency that we sort of think about this machine, we think about regulations or institutions that would help us manage this Human Rights, and we don't think about ourselves as users and exploiters and, what do you say, people put out or use that. And, one of the -- sorry about that -- I can hear I'm going in and out here.
        One of the things that I know from my own experience in working for quite a long time is my own personal -- my own personal journey about superiority and inferiority. And, I think one thing that AI could allow us is a little bit more time to go on that journey. And I'm sorry if that sounds very metaphysical and very sort of personal and psychological and all of those sorts of things, but the thing that comes out to me, living in the most recently liberated African country, and working with African people all the time in a black university is that sort of internalization of inferiority and superiority that plays very hard into the Equality context. So, when I think about sort of institutions, regulations, working with Human Rights and things, I would think it's beneficial for us all to think about why we as individuals think essentially do have a superior, everybody here has some kind of superiority complex. I've been working with mine with black people for a long time. On the flip side, the marginalized people have this kind of internalized inferiority. And, I think I would just like to sort of put that out there that maybe that all seems a bit touchy feeling, a bit scary and

confrontational and all those sorts of things, but I think there is room here to take on that sort of personal identity engagement in things to start thinking about what that means when it comes to deploying these tools.

>> SHERIF ELSAYED-ALI: Thank you very much. That would be a good thing to take up.

Okay. Go ahead.

>> One thing that comes to mind is as we think about policies moving forward, it seems important not to overly privilege AI. We're all dazzled. I'm dazzled by the possibilities. It is just an element and journey that people go through. I'm reminded of a specific type of case study example. You have two doctors, Dr. Alice and Dr. Bob. They're both surgeons. Dr. Alice being the woman is paid 70% of what Dr. Bob is being paid, even though they're both the same type of surgeon in the same hospital. So, we do become data driven and the data shows that Dr. Alice has a much lower mortality rate in surgeries than Dr. Bob and it is clearly unfair that she's being paid less, but now further data and further machine learning shows that Dr. Bob is being assigned much more difficult and much more challenging cases. And, related to the type of (Inaudible) you could argue Dr. Bob is the superior surgeon and maybe justifies the higher pay.

Further investigation of the data, as we continue this journey, show that there may be a gender bias in the assignment of cases. Privileging Dr. Bob with the harder cases and on and on. It ends up being kind of as more data is collected, as more thought and more analysis, it is not in itself an answer to the fairness question that is injected into the process into the journey that we go through in order to succeed fairness in all of our human endeavors.

>> SHERIF ELSAYED-ALI: Maybe I was just going to ask if anyone on the panel, do you have -- would suggest a food for thought on the question. So, what this proposal has been named a food for thought on what we -- how to protect, how to protect Human Rights, how to protect Equality in AI.

Lydia, please.

>> LYDIA NICHOLAS: I will quickly repeat what I said before, that the places where I see people able to control the information, what happens with their information and the kind of questions that get asked and the kind of impact it has on their life has been in places where people have genuine ownership of that information, and although at the moment the lot of the information is machine learning is the typical exhaust of thought, I'm seeing really interesting businesses and stuff working on getting information that is much better cure rated, much better polished direct from people in ways that they can fold. I've worked at stuff called citizens me that keeps all of the data on your phone encrypted there so they never see any of it, it they run surveys, and it can be health issues or rights and political issues and you decide what to share. And, people are willing to pay for that because that is better date that you that they know is not kind of exhausted as fuzzy and noisy and messy. So, there are alternative models in which people have a much more direct link and get dividends whether in the case of some (Inaudible) genuine monetary dividends or in the case of health data things like guaranteeing or at least pointing for access to treatments this they contribute to the production of donating their data, their time, their information. So, I see situations like that as one approach within a much larger ecosystem in which regulations and industries standards and ISOs and things like that are all part of it.

>> SHERIF ELSAYED-ALI: Vivian.

>> Thank you. Second food for thought.

So, I think we need to be able to brake a pathway about (Inaudible). I think who we are referring to make those decisions is important to think about.

>> And, Rapporteur for organizing this discussion.

>> AI is (Inaudible) source they should be (Inaudible) and how data is collected, analyzed and used to access the impacts of AI and my marginalized community. And, I would like to say leave no one behind. The key principle, key consideration. I would summarize as this.

>> Very quickly. Any ten-second reaction to this or addition? >> Maybe this leave no one behind thing because it

reflects employees. It is this fascism that everyone must be connected or without asking people why they are disconnected. The reliance has done great research for the people that haven't come online because they have access but the ones that have access are not using the Internet. That is really instructive why people might choose to stay away. When we brand everything if everyone must be rah-rah technology and rah-rah tools, it might be some people want to stay off line or not eager to participate for all kinds of reasons, political, stubbornness. I think when we have this broad-brush approach that everyone must sign up to the shiny new universe, we reject choice.

>> Yeah. Got that. Great. Thank you.

>> Not to be difficult.

>> Any last thoughts on the question of protecting Equality, protecting Human Rights within AI? Any last thoughts, interventions? Got three. We will do it 20 seconds, please.

>> There we go. It back.

Thanks, Sherif, for moderating a great panel. Really concrete examples from you guys that was missing throughout the whole conference, so thank you for that.

A term that I wanted to introduce here for promoting Equality and access to AI was this principle of democratization and the Montreal delegation sitting here, we've given it a lot of thought and we came up with certain principles and I guess we didn't get a chance to really get the panel starts on that, but I'm really going to throw them out there quickly and if I can get five-, ten-second reactions from you guys.

Some of the ideas we looked at were accessibility, low cost representativeness, standardization of data and interoperability, giving people a voice in this, and most importantly, some education measures that not only empower people to understand what the implications are, but also to help them utilize the tools in their specific localized context.

Thanks.

>> So, thank you very much for that. That is fantastic and I should be very good to talk about the access. I will just pause that for 30 seconds just to collect that any, one last --

any last intervention on the previous point on promoting technology. I think there is someone there.

20 seconds, please.

>> Just because for the translation and the recording, I think use the mic.

>> Level. So, when each of you were thinking about questions of fairness, questions of Equality, you are thinking about it in a very much outcome way. So, a process is fair and equitable if the outcome is not discriminatory or treats different groups in equivalent ways.

But, there is another way of thinking about how we want to demand treatment from AI systems, which is to say that when we treat it in a particular way we should have the right to know the reasons why we treat it in that way. So, if I'm, you know, privileged extra scrutiny by the child protection services I can say why have I been treated in that way? When you look at technologies like machine learning, which may be essentially inscrutable, how can we provide those reasons that is consistent with our obligations in terms of Human Rights, equity and fairness.

>> On that, last thing. Peter, if you could direct very quickly to that.

>> I think for high stake decision that effects individuals, that individuals in many cases will demand these explanations. I know when we deployed something for salesmen and AI says offer this customer 20% discount. No salesman anywhere on the planet is going to succeed to that. They want to know why. When the Audi spoke on the first day about cars making choices, what is the one person or another person they're building, I think the fundamental need is we all want to know what is the reason, what is the reason they built that. We want to know why. I think this is actually some of the fundamental needs that we all have. People enjoy the technology.

>> Thank you. If I can go back to very quickly, if we can just say in three words, just say three principles that you mentioned again, that would be great.

>> Yeah. So, the principles were accessibility, lowering the cost, representative ness, standardization of data and interoperability, and giving people a voice, and education

measures that help people not understand and ought so utilize AI based technology.

>> Thank you. Going to the panel. In terms of the second question which we don't have that much time to talk about.

So, kind of reacting to this in terms of ensuring equality of access to AI, you know kind of directing to these three ideas and building on them, any thoughts? Again, 15 seconds, please.

Anyone.

>> I think what the speakers referred to the way I would put it as humans being things and Internet of Things, and so a fundamental aspect of this is will that create another type of digital device. Some people less instrumented than others and therefore (Inaudible).

>> We definitely already see this, and (Inaudible) I could go on for hours.

While I talk about public access, I absolutely understand the problems we have in general understanding of math particulars and statistics. I've done focus groups of people in quite developed countries where people are reasonably well off and they denied the existence of influence and statistics because how can you know that? Are you God? How does knowing anything about anyone else, anything about me? If you don't use these things in a daily life and don't understand modeling, we need a lot, I think, stories, as well in terms of data strategies, ways of getting people interested that upset them is why I think we need to know the easy route and work in health and local government, because people can generally see if you stand back and provide information then I know that my children will know more about their genetic predisposition, that is something that I can understand. So, I think implication and accessibility a lot of times working with local issues, local problems, local communities and their interest, necessarily getting everyone to (Inaudible) which I can.

>> One is I think we really need to learn from alternatives and explore options. And, the lessons from different communities and different histories of discrimination. And, one resource I would show to you is (Inaudible) from the MIT media lab. We're so inspired by the address by Megan Smith at MIT but 30 years of women at MIT that she wrote a globe post collecting 70 features on Equality discrimination and bias. I would urge you a median

to see this, and a way to decide (Inaudible) by the media lab across different technologies, different sectors within communities, within regions. That is a really helpful resource.

And the other thing I would say is with a lot of technology we sort of assume we don't need to know how a car works in order to drive it. Why is it for AI we assume everybody needs to understand how AI works in order to benefit. On one hand, I do agree that we need to understand, on the other hand, you have product liability standards, you have to take them at the platform level. You inject a free (Inaudible) into a design state so that the end user doesn't bear the burden of exercise choice and routes that are save. You make sure that there is a minimum baseline with other tools and technology.

Why is it that AI we're trying to push everything down the chain, particularly down to the end user? So I think it is very reductive analogy and the whole (Inaudible) now everything is a problem, but I think it is really instructive in saying if we expect it of other tools, we don't under need to understand how something works in order to use it or benefit and we can use it, why is it with AI is effective, is it so different, is it the case the opportunity for bias are so much more that we expected case, and if so, what kind of education do we think is useful to help people navigate those choices and how much of it resolves to the user, how much of it needs to be the problem of industry and government and policy makers? Because I think one of the things we see now is that the rules of the road have been framed by people creating the technology. Government and policy makers are slow to catch up with all the big companies creating AI and using it with best practices and guidelines.

As a lot of lobbying, I see this especially in Asia, where it is seen as low rights jurisdiction where a lobby government to come up with a set of guidelines or robot efforts, in developing countries (Inaudible) in a vacuum of legal safeguards where they don't have (Inaudible) rights, and safeguards (Inaudible). So, this kind of jurisdictional sort of optimization and the form swapping I think can be really harmful. I think this is something that we all need to look at as a community of end users.

Just to end, as colleague Alex made a great comment saying

there are only two industries that talk about users, the drug

industry and technology. And, I think that says a lot about

where we are.
        >> That is a very, very good action.

        >> Access to technology means that regulators need to be
able to scrutinize systems. That is currently difficult. We
need sort of, like, there is lessons to be learned. I'm also
like Malavika and become critical for vague demands for
transparency. What do we need by transparency. Transparency for
whom, and also we sometimes say transparency when we actually
mean different things. Transparency is a means towards an end.
It is not an end in itself and there are lessons to be learned
from open government initiative, information requests where we
see that sometime these transparency tools can lead to the
opposite. That they can lead to less accountability because
Governments can say yeah, but it is on the Internet, so you could
have checked. So, I think if we want, again, not for everything,
but in cases where critical decisions are made, we need the
ability to oversight and scrutiny. That also requires people
with skills that are able to actually conduct such oversight.
And, I think for example data protection authorities, I'm not
sure they have the capacity. They don't have enough in Europe
countries, don't have enough money to go in and investigate, even
though this is in the law.
        >> Thank you.

        >> (Inaudible) used to achieving responsibility and that is
really enforcing. Just thinking about what access means. We
need to think about the disparity (Inaudible) one example.
(Inaudible) and I think it's important to detect the problem in
different forms. I think there is a tension that Malavika that
has mentioned about legal safeguards lacking for world that is
coupled with lots, a very high level, very advanced adoption of
technology, and I think it is tension that makes it dangerous.
It's not that people aren't using technology, they are using
technology. Safeguards that is the risk that we've been talking
about with regard to facility, and I think we should be able to
speak about the specific challenges that are inherent in their
society and not just particular to the application of AI. For

example, if you have political ability, if you have transitional government (Inaudible) that we've been talking about, if you have infrastructure of (Inaudible) where even being connected is an issue and these are problems that not only pertain to be success AI and I think that excess on another plane is kind of a remark. It is not just thinking about access of platform of technology, but also being able to achieve some sort of fairness. And, I know that that has been a big problem that we cannot define cleanly, but be able to access content that is not discriminated in a way because you have, you know, as a result of having personalization and user content, we have very different types of information that are filtered by users and in bubbles that are difficult to begin with and give us an exercise and information to begin with.

>> Thank you. I'm going to go to Ahmed to see if we have any concrete food for thought.

>> In the previous (Inaudible) limited to AI. The audit should be clearly effected and impact of AI on Human Rights that can be (Inaudible).

>> Malavika, do you have something very quickly on that? I'm going to say we have two minutes left, so I'm going to ask if I can take an extra
two minutes. Is that okay?

>> MALAVIKA JAYARAM: Around labeling, I think a lot of people don't realize that AI is being used to do things on their phone and the Internet is being open in various ways. I think the areas of product liability and data symbols, think there is a lot of room to play with symbols and the way IA is collected. The idea of labeling is really important because labeling is so huge with machine learning in terms of categorization. I think a different form of labeling would be really helpful to people.

>> Thank you very much.

>> Coming up and take away or kind of profound thought on access, the issue of Equality of access.

>> I'll go back -- I'll ask this in one second. Just going back to the RIB question. There is just should this be within, like, whether this should be independent or not.

Peter.

>> I think that most just speaking for Microsoft, because I think this is for all of our peers in the tech industry, we have wildly extensive internal reviews. Much more extensive than what I spent working in both the government and the University, so I think maybe surfacing the term Human Rights at the top level is probably a good idea.

It wouldn't surprise me, I don't actually have the verbiage in my head, but there is actually a rather simple policy. We may already be there.

I also just wanted to say I also endorse the Montreal

schools.

>> SHERIF ELSAYED-ALI: Thank you.

Frederike, were you about to say something?

>> FREDERIKE KALTHEUNER: I was just wondering when we want to constrain ourselves by saying that these reviews only happen internally and they're great companies that are responsible companies do review but as a liberties organization I'm always interested in the bad guys where this is not happening, and I think those are the ones where internal review will be window dressing or not happening at all.

>> Equal access to AI or do we mean Equality vendor to AI. And, I think those are two very different things. We need to be really cheer when we talk about access what we're talking about.

>> Just to add to that, it is worth saying for a publicly traded corporation like Microsoft, access is not fair. There are groups like the Klu Klux Klan or ISIS and those are groups that do not get clearance. So, there are always gray areas here that you can worry about, but at the moment, it's not strictly open.

>> SHERIF ELSAYED-ALI: Vivian, you're the last person.

>> Ten second suggestion. Perhaps we should focus on the review and not limited to the internal, in my opinion.

Thank you very much.

>> Thank you. Just the Montreal principles, generally just a show of hands of how people feel about them, if you feel a good starting point.

Okay. Take that as a yes.

Okay. So, just maybe one last thing to include, perhaps, there was a point around the expand able AI that I think would be good to reflect even if we don't get to a conclusion about that.
Lydia, Malavika, Peter, Vivian, Frederike, thank you all very much, and enjoy the lunch.
(Applause)
<center>***</center>

This text is being provided in a realtime format. Communication Access Realtime Translation (CART) or captioning are provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.
<center>***</center>