AI for Good Global Summit Session: Breakthrough Groups on Societal Challenges
June 8, 2017
3:30 AM CT

Services Provided By:
Caption First, Inc.
P.O. Box 3066
Monument, CO 80132
1-877-825-5234
+001-719-481-9835
Www.captionfirst.com

***

***

>> We'll give people just a couple of minutes to finish wandering in, and then we'll get started. It's always a tradeoff. I want to reward the people who were prompt. At the same time, I hate to disturb you by having people come in as we're doing this. So we will start promptly five minutes late. Around the topic of humans and computers and then asked to figure out how to do that in a small amount of time. And I will say by way of introduction that half of the planning was trying was trying to figure out, what does not fit under the umbrella of humans and computers in this topic.

And after we failed to succeed in narrowing the topic very much, we focused, instead, on finding an excellent and diverse group of participants who could offer different perspectives. A little bit about how the session is going to work. I'm going to introduce you to our four panelists who will each start by talking about their individual perspectives about what issues about what AI Robotics, the broad set of add advantaged intelligent technologies can do for people for good.

After we cycle through that, I will warn you that each one of them also has some fairly serious concerns about how we might not get there. And we may get derailed, instead, by falling down some path that will not be for good.

And over the course of that, I hope we will come to a point where we have an optimistic future, a precarious path towards that future. And then, it is our time to discuss, are there guidelines, priorities or principles that we can put forward at the end of this

session? That have some hope of allowing the nations of the world, but I don't think you should think of this as limited to governments.

This may be the businesses of the world, this may be the universities of the world to self-impose on themselves a path that leads towards good. In that part of the session, which should be about the second half of our time, I will be putting up -- and you will see me editing, as fiercely as I can in realtime, a draft of these principles, priorities or guidelines.

They will change as people make comments. They will continue to change. Towards the end, we may see to what extent we may see something resembling consensus. And they will then be presented to the plenary after lunch where there will be a chance for people, including you to vote on which of these you are interested in getting behind and which, if any, you might be interested in working on. If ITU convenes a group working on this going forward.

One last note before we go to the fun substance of this -- well, two last notes. Last note number one, we have remote participants. And we will periodically, I hope, hear from them. They're encouraged to send in questions and comments and they will be, you know, read aloud as if they came from within the room because the voice will come from within the room.

Number two, while we did not come up with that perfect scope, we have two themes that the panel has really organized around. We have two people who spent a lot of time thinking deeply about the practical, ethical, and societal challenges, specifically, around robotics and around deploying robots in all sorts of applications. We'll spend about half of our energy around that question.

We have people who thought extensively and practiced extensively in the idea of technology as an enabling and empowering factor in the developing world. And we will spend about half of our time there. And the third half will come from whatever else you bring into this discussion, just because robots can add doesn't mean I can't have three of them. It is my pleasure to introduce first, Marie-Helene Parizeau, the world commission on ethics with scientific knowledge and technology. She's also a faculty member in Quebec. And she's going to start by sharing with us what excites her about the potential of this technology.

>> MARIE-HELENE PARIZEAU: I will present three examples and then some guiding principles, ethical principles. Here there. So my first example is the use of robots. If you analyze the ethical here, robots are used instead of human beings. And it's kind of a self-evident objective beyond efficiency.

So here, we try to avoid putting human beings and human life in danger. That's my first example. So it's kind of a straightforward ethical thing. Second example, industry robots. Car assembly line. Here, the objectives are wider. Efficiency, higher skills, reduction of hardness of work.

Of course, here you have more social impact from the use of robots, especially from the workplace. So questions here are larger. To which extent, industry robots will decrease workers. It is a question that we have a little bit this morning.

As the difficulty of work is decreasing, which was one of the aims of using robots, to which extent strategic know-how are replaced by artificial intelligence. That's my second example. You can see that the ethical and the social questions here are larger.

Third example robots use for the treatment of children with autistic syndromes. Here, we have a peer objective. And then, it's much easier to apply ethical that are standard in practice in biomedical research.

For example, (inaudible) risk benefit analysis. Children are -- the robo groups and the standards of evidence-based medicine, which is the standard medical practiced. Robots are used not as care givers. This, again, this opens the kind of questions which are the interactions the fact that emotional link can be developed between the child and the robot.

Shall I continue on the ethical? Or on to the --

>> JOSEPH KONSTAN: Let's vote next. Mady Delvaux-Stehres you've been introduced to, already. But those who might have come in slightly late to the earlier session, she has a distinguished history in public service Europe-wide. And has led and been delightfully outspoken in an effort to develop guidelines around the use of robots, trying to address ethical issues, economic issues, societal issues. And having heard her this morning talk about the challenges we have to avoid, I think it will be delightful to hear why it is that we're pursuing this anyway. Which is part of what we will start with.

>> MADY DELVAUX-STEHRES: Thank you. I'm less prepared in my comments. And when I speak on the excitement of all of the examples and adding those already, I think mining, for instance, can achieve work that humans are not able to do.

I think the culture avoiding pesticides and so, I see a lot of advantages. I would now take an example of a medical sector. This very intelligent tool that can give you a diagnostic in ways that is far better than diagnosis by human doctors.

And I think this can be very helpful for people all around the world, which have no advance to quality in medicine. But I think it's very exciting is that you can help all of the history of and the travels of the whole world. Which is the first step.

>> JOSEPH KONSTAN: If you keep wondering why I'm cutting people off before too much negativity. We need to be excited, first. We need to realize this is something we have to do. And then, we can spend a little bit of time figuring out how we have to make sure we do it right. And so with that, I'm going to introduce Apala Lahiri Chavan, who is the President of Human Factors International

practicing worldwide and is, herself, a leader who is known in the research and practice community around human-centered designs.

>> APALA LAHIRI CHAVAN: Thank you, Joseph. Not being really from the field of AI robotics, however, but because of the design of research across Asia and Africa, the impact of technology on large sections of populations, often left behind, it's something that is very close to my heart. And that's what, you know, I have also worked on for a while.

I come from the point of view of that context of understanding that ecosystem and what I do for that ecosystem. I start with that story because it's always a good start. This is a typical gated community where we live. I live in one of these.

And that's not the actual person, but that's Sheetal, a domestic worker. And the story begins about a month ago, before I got this invitation, where one day I was coming to my apartment, going down to the park to get in the car. Now, it was later than usual, so the elevators were sort of not going crowded. So as the elevator stopped at my level, I walked in and Sheetal was sitting in the elevator just completely empty. She was sitting on the stool that's meant for the lift man. The lift operator who sits there.

She was not there, she was sitting on that stool. She jumped up and she told me -- that's why I was sitting there. It took me a second to understand what was going on. She said, no -- please, please don't tell anybody. Just because I was not well, I was sitting here. Turns out, she said, that the source, the men that run the committee that look after the entire apartment complex had said to all of the domestic workers, you dare not sit in the lift. You dare not sit on that stool no matter whether there's anybody there, or not. This is not for you.

So you must always stand. If anybody complains, you're out of a job. And I said to her, that's nonsense. Nobody can do that to you. She said, no, no, no -- this is exactly what she said to me, they're all very educated people. She didn't say they were rich. She didn't say, you know, anything else, they're very educated or not. What they said, it's true, just don't tell anybody.

And that left me thinking these kind of things happen across the continent, but this one left me feeling like I have to do something. So when the invitation came, I went to thinking about her. And also, what can I do for her? You know, I meet her now very often. And I thought this conference for AI for good. When she looked at me thinking, oh, my god, you strange people.

So she depends completely on other people to integrate the world for her. She's street smart, she understood how people exploit her. She can feel something is not right. But most of the time, she's too scared to ever question anything. The education infrastructure, where she comes from is terrible.

As you can see, you know, enrollment may be high according

to government statistics. Dropout rates are terrible, particularly for women. Most women drop out in rural areas because of many reasons. The other reason why she particularly everybody like her who is in these big cities in these positions who are in the strip have similar stories. She came to Mumbai ten years ago, she told me, because her daughter who is 19 years old in the village had been married off and was pregnant and they discovered, as soon as, you know, pregnancy is announced, then, they do the test to see whether female or male fetus. And because with a female child, she was told to abort. So her daughter ran away to Mumbai.

She and her husband came to Mumbai searching for her daughter. That ended happily, but she did not go back. She stayed in Mumbai. And one of these things, you know, why females -- why so many people, so many young children die, why they don't enroll in school? So much of it is because the major illiteracy of women.

They have no formal education or cannot complete a primary education. And then, we talk about artificial intelligence. So how go we connect the dots? So major potential to make some change. How could she afford any access to anything like artificial intelligence? But you know what, there is a full model that has been successfully tried across several continents, but definitely India about how you can eradicate poverty through profit.

So at least, both times it's a win-win. And she's from the bottom of the pyramid. There are some incredibly good examples. The only challenge with those examples is if they're scalable because they depend on large masses of people who volunteer to go and see initiatives and rules in India instead of depending on so many people from over in India, what if technology could step in and scale those kinds of initiatives for the bottom of the pyramid?

Also, there are so many successful cases where community-based ownership has changed the landscape in villages where these have been implemented. NGOs, academic institutes in the west, experimenting.

So community sharing, collaboration, collective usage. This is part of the culture's DNA. Why not use that as a business model to provide access to something like artificial intelligence? But what would it do? What if there was community-based AI systems that would feed language and basic life skills to people like Sheetal? What if that could be done? There are other technologies that have been adopted through the community-based ownership. Why can't we look at AI? In education, it would change in there. 50% of women in India could just be more confident, self-sufficient, aware of their rights, know how to deal with life every day. No matter what information.

If only that would happen, so could there be design of systems for that? Or what if those of us who want to volunteer, you know, a lot of us feel terrible that we do nothing for people like Sheetal. What if there was a way we could have access to AI systems

that we could, then, use to teach?

And I tell you, this part of the story, Sheetal said to me a week later -- I told her, why don't you get some literacy? A literacy program run by the government? Why don't you go and get that? And she said, I will talk to you later because she was flustered about having been caught on the stool. She had no thought to process anything. She came back a week later and said, I want to.

I gave her the address of a neighborhood one. She came back a week later and she said, the teacher there, I'm so scared of her. She tells me you, it's too late, just don't come here because you're a waste of time. She asked me, will you teach me?

All of the bleeding heart, everything, but the practicality of -- okay, I really want to, but how? I don't know how to teach. I don't know how to teach language. And it's not just me, so many people in India want to be able to say something. Could I have some AI systems that I can use to teach her? And so, that's where I leave you with the thought. Of course, there are many concerns.

>> JOSEPH KONSTAN: Thank you.

(Applause)

Continuing to -- as we progress through this, increasing in levels of inspiration for the potential this year, I bring us next to Professor Nick Bidwell who has previously also held the title of Professor at the University of Pretoria is a particulate of some of the challenges that are being faced by people in sub Saharan Africa. Technology has some interesting ways of amplifying things.

>> NIC BIDWELL: Thank you very much. I and my students in Namibia in particular are excited. My work and many of their interests tend to designing interactions that suit the knowledge and communication practices of rural and indigenous people.

And we're excited by something similar by the potential for AI to increase access to resources for marginized people. And to ensure that they really have a voice in the future oh, sorry, next slide. There we go. So this is pertinent from the big goal. I'd like us all to pause for a moment and say, all of the discussions so far have been in English, but we don't live in an English world. We live in a far more diverse. We'll talk a little bit more about diverse later on.

So AI promises to assist us with translating between the world's at least 7,000 languages. I'm going to take the two countries that most of my research recently has been involved with. So that's South Africa. I still have Ph.D. students and so on where I live now. So we have 11 official languages in South Africa, but there are at least five others widely spoken. And we have one official language.

And for those people don't know what it is. It's two hours up towards Angola. The most recently liberated African country that gained the independence in 1992 from the occupation of South African forces and other forces.

For political reasons, decided to adopt English as an official language. But, again, over 16 language groups. And despite schooling in English, most people don't speak or write English very well. And this is reflected in resources in people's home languages. Almost none in Namibia. South Africa has -- it's represented (inaudible), but there's many, many languages that aren't represented online.

And this acutely impacts on technology and literacy. And on human capacity building. My students do computing at university unless they have done well in English. And for many of my students who have come to university, I would say are better at English than my Australian students. But they are in the minority in the country.

So they have to work and think in English. And this is different from their failed experiences. Their experiences of their identity, their experiences of their relationships with others are in the natural language. I would like to see, potentially, that AI can assist people like that. I would like to see how AI can be involved in translating between languages to enable people to take their identity and experiences and use that to build their own future.

We have a huge commitment in Namibia to solving our own problems for ourselves, including using technology. But we are somewhat limited because of the dominance of the computational world.

I believe, and my students would like to believe, that AI can help us to move on towards that future.

>> JOSEPH KONSTAN: Thank you. And there's a point that we want to remember, as we've heard -- yes, that if we keep hearing about, you know, the wonderful power that we've had in machine intelligence, machine translation, that so long as that power is based on finding sufficient text, it will be harder and require more conscious effort to deploy it to lower-used, smaller population languages. Because they simply don't have the volumes of text. And that creates an interesting research challenge that I hope some of you may be the ones who are up to.

We're going to take pass 2 now through our panelists, hearing the broader set of issues that we should be aware of as we make sure that we do, indeed, try to deploy AI for these lofty purposes.

>> MARIE-HELENE PARIZEAU: Okay. Okay, there's not much transition, which I'm trying now to propose. It's more the result of the reflection of the working group on robot ethics. And this is a very short presentation. The major principle that we have identified that can structure some reflection upon how to use robot, how to justify their utilization and how to help to regulate social practices.

So we designed a few principle and values that I will just briefly comment. In order to show that there are some existing

principle that can help us to regulate the use of robots in different society. The first principle that we have a consensus upon is we do not harm principle. This refers to laws that I think are still accurate.

The first one is the robot may not injure a human being, allow a human being to come to harm. The second law is the robot must obey the orders given it by human being, except where such orders conflict with the first line. The third one, a robot must protect existence as long as protection does not interfere with the first or second one.

Of course, in our working group, if we morally take seriously this principle -- if we are morally serious about this principle, then the question of armed drones and autonomous weapons should be banished. It's a strong commitment we have made in our working group against autonomous weapons.

It's not -- of course, I won't discuss this here, but we have really hard discussions on the topic. But they do not harm principle makes balance between robots and humans. And that is the idea of this principle.

The second principle that we -- oh, sorry. Identify is, of course, human dignity. That is a core value related to human rights. Here we explain human dignity as related to a principle of autonomy that's also expressed through recognition of relationships between humans and between humans and animals and between humans and environment.

So what is important here is how do we understand the question of the use of social robots, for example? Does it enrich our relationships? Or does it reduce them? And of course, the question is, how can we assess those consequences on the relationship between robots and humans?

Another aspect we have discussed as we entered around the interdependency is that it's implied that robots are part of our creation. Our artifacts surrounding. And this is also environments. Most of the time, there's no discussion about the impact of using robots, especially. The ways that are generated by robots, the energy consumption and the CO2 commissions. The technology footprints of robots.

Robots are not virtual, they're material. And those aspects must be sensitive. Another set of values that we have, also, identify is, of course, a principle of responsibility and liability. And this is to include -- to have a more inclusive perspective with researcher and industry and government.

To create a dynamic where everyone is partnered in the perspective of innovation. The robotic development should not be reduced to economy productivity and efficiency. There are others that are relatable to responsible innovation.

One of the -- consensus were that humans should be always

in the loop and researchers and industry should find ways to control robots by different means. For example, responsibility. In order to maintain human responsibility at the core ethical principle. Because this related human responsiveness is moral responsibility, but also legal responsibility.

So those are the, I would say the -- very important principle. Another one is a principle of professionality in technology. And it's always asking why are we using a robot? But take into account the social context assessment and implementation.

I think we have two good examples where it's important to see how robots or AI can be used in different social context. And how we can be useful and determined to -- does the technology develop imposed to people? Or is it designed for the people? And eventually, weakness for the people? So how does the creation of robots and use of AI are related to the needs of people.

So this implies another value, which is cultural diversity. Robots can be used in certain settings and not others. That makes a general setting of principle that (inaudible) identify as an element of governments and also industry in the development of robots. Thank you.

>> MADY DELVAUX-STEHRES: May I commend the presentation of my neighbor? I have to say, when I think of all of the discussions we had, I feel ashamed when I say we have problems of rich and developed countries and we think these problems are enormous. And when I listen to you, I think we should really concentrate on where are the biggest need is. Because I believe that although there is a lot of -- we have to prioritize. And to where it's more money invested and where it's the biggest need. So I have to say, that I learned a lot from you and I'm quite -- you convinced me that this -- because I was a bit skeptical, the developing countries, but you convinced me that, really, there is a lot of -- a big group of progress. But I want to come back to my example of the medical diagnosis. And I have to -- if I understand our moderator, I have to say, what are the risks. And I believe practical as we are working on the pathway. If you have diagnosis, such as to medical, it can be in certain circumstances dangerous. But we have also need to train doctors and medical and the nurses on how they can manage this interaction with this diagnosis.

Who will be liable at the end? Who will contest the diagnosis by artificial intelligence? And our artificial intelligence and decline all of the responsibility. And say, well, told me that you have this or the other. And I decline -- how will we train our doctors in the future? This is a very practical presentation. And I have to say, that I do not have the answers. Because what I know from the studies to be a medical doctor is that they have to learn an enormous amount of statements. But they will never be as good in data as artificial intelligence. What will be

the interaction between human and machine? And how do we acquire this expertise?

And on the other hand, this enormous memory of every data that are available. And on the other hand, what is the quality of the data on which artificial intelligence works? Is it really all of what is identified? Or did they make connections?

For me, it's initially what we have to look at.

>> APALA LAHIRI CHAVAN: Okay. So I feel that, you know, a lot is possible. Like, I was talking about. But just as Marie was talking about ethics, there should be some guidelines, some principles. I feel a system of set of values are very important to embed in those, you know, sort of DNA of artificial intelligence systems.

And this is just an example of the new universal set of values that we often use in our cross-cultural research and design to understand different cultures, particularly when designing new systems. I'm not suggesting it has to be this exactly. But just values with which exist across all cultures.

They all exist. So why not take values such as, you know the values of universal, values like that and build that as the base DNA and the systems artificial intelligence systems will never breach those values no matter what positions they are taking as they learn and decide to act. They never reach. It's very important to have that set of factors. That's one.

Also, core creations. In design, core creations often get the best results. No designers sitting and creating by themselves. And not the users of the design systems creating by themselves, but working together. And it's a philosophy of a system mainly from Scandinavia. And I think it's got to be a co-created new reality where the co-creation is happening between these artificial intelligence systems and talking to the developers and designers of the systems. And the users of whom for all of us whom these systems are meant for.

It can't be done in silence, you know, where it's just nothing at the how the intelligence, get even better and get even smarter. But it's got to be in the context of working in that human ecosystem. After all, what's the objective of artificial intelligence?

So I think that co-creation doesn't quite exist. And it must be part of looking at the future of AI. It also, then, will enable us to keep in mind context and culture. And to enable local development, you know, of AI to happen, we've got a huge movement in India now waiting to rule in India. I'm so surprised. Why not use those kinds of movements to also see if artificial intelligence, developers can exist. That everything doesn't have to be important.

So that, I think, is important from my perspective. And this is, you know, a little out there, but I thought I'd leave it, anyway. This place is one of the places in India where we have a design

studio, it's on the southeastern edge on the ocean. What was very special about it, he's considered a seal.

He was the revolutionary freedom fighter who fought against the British rule. He escaped from prison because he was going to be executed by the British, he escaped to a French colony. Once he reached there, within a few years, he underwent a complete transformation and started a movement. And it's the base of that. He's passed on to his co-worker, the movement remains.

In Ponde Cherry, there's a new township they created. This is a meditation space. That they created. I often joke that isn't that a spaceship? Come on, it's not a meditation space. You get up and you're going for meditation and you'll not be there. And certainly, when I was thinking about AI, I remembered often talked about knowledge -- this is the full integral truth of anything. I've never understood what that means.

But it's plastered everywhere in Ponde Cherry. And I thought, is artificial intelligence, perhaps, what he said was this super mind is coming into existence to help humanity evolve to another dimension. To enable that journey to another dimension for humanity. And I'm looking at history of artificial intelligence. He said this in 1956. I found -- I don't know if this is correct, or not, that at a workshop, the beginning of AI research really came about. That's the turning point.

And guess which year, 1956. So, you know, it's one of those trivia that then makes you think that if nothing else, one should look at philosophies from the east. Various kinds of philosophies in other areas. To see the guiding principles to fashion AI so it doesn't become an end in itself, it's a means to help and enable that evolution. And to really actualize the potential of human race, which is not happening now because so many people are behind.

But all of that is great. And can think about it. Say, okay, whatever. But some real concerns. Since I think having a set of values is very important. But whom is going to do this? It's got to be shared across everybody who is creating artificial intelligence systems, how can this ever happen? So I wonder if this, you know, it stays in the pipe dreams.

In countries, just as I told you, so much means so much. There's so many by people who manage the building apartment complex to everybody else in their daily life. So if AI cannot be accessed by everybody, nothing will change, nothing will change.

Yes, our most fancy hospitals that use AI, which they do already, you know, the system, all of that robotics, all of that the change for the rest of the country, which is the majority, really, of the country. I think access of AI to everyone is very important.

And it's about being open to different business models. Different ways of disseminating and making open. And at no conference ever, I think that's really, really ridiculous. And how can we change

that? It's so difficult. But if that doesn't happen. If we don't -- not just hear their voice. I am now an intermediary, I'm giving you a story. But they should be giving you their perspectives directly.

So I think that's very important to make happen if you want to have artificial intelligence for good. And then, finally, looking at my own -- I don't find any mention of design or the social science, sociology. I don't find mention of that in all of the talk about artificial intelligence. I think it's very important to be disciplinary, because bringing to the table the understanding of technology.

So these, then, are my concerns.

>> NIC BIDWELL: So without wanting to undermine the speakers, we welcome funding and partnership from you. But it has to be a fair partnership. And we have a long tradition of lots of aid coming into Africa. You've probably seen the stats in the news recently. The amount of aid that comes into Africa.

It doesn't match the amount of money extracted from Africa. I'd also like to say about the participating design. I can now see and we have to do it differently. I wanted to put those perspectives out there and perhaps we can put them up.

I can't see my slides up there. To the issue with translating. Perhaps a bit something more specific. And I hope it's not too, too specific. Let me wait until the next slide. So while most mainstream products, like, Google Translate focus on the most resourced languages, like Joe mentioned, it seems as if -- can we have the next slide? Oh, it's me doing the next slide. That's why.

Sorry. Okay. There we go. So it appears from my fairly superficial reading that systems with human machine partnership are the future of the language translation. So the quite easy and accessible model for me to understand was one that I saw, put forward and I think represents a mixed role of many different models, a generic model by Manning's group at Stanford. So in this model, the idea is that there's an English input. And then the machine suggests a translation in Arabic. In this translation suggested the inflexion is the wrong gender. The user has the opportunity to turn it into the language it wants.

There's a relationship between the humans and the machines. However, there is a routine. This genre of correcting by the user may not be acceptable or entirely unsuitable for the user experience of many language groups. We know that the forms put into print change the form considerably. And that changes the experience of the person reading that form. And that experience connects us to ourselves, to our identity, to our relationships, to other people.

So the interactivity used in the symbiotic model is a particular model of interactivity, a particular conversational style. And if you haven't had a chance to learn that style, if you

haven't had the chance to learn that style because you haven't got access to technology or because you didn't learn English in school or because you went to school in another kind of style where you learned and interacted with your teacher and fellow students, then you are going to feel alien from that training. It's not going to feel like you training the machine at all.

My students were trained through so-called international standards of design. But they frequently tell me about the significant gaps and breakdowns between the language and logic of the AI I teach and the textbooks teach. And they tell me what they do in their heads when they are using the model they learned at school, university with me. And then, what they need to do to try and adapt it to rich home languages. We know from philosophers that language, self-experience, identity, are intertwined with each other.

And we don't get these conversational rules there, we are going to -- oh, sound. Can we turn the sound down? I wasn't expecting the sound. Sorry. I was just expecting the video. Sorry. Okay. All right. A concrete example. And it's work I do for a long time with traditional living, speaking people.

And what you see here is the young group -- of young ladies who are showing each other and interacting to help them share oral resources. Just to manage their community practices, share information, could be quite personal information, could be funny information, could be traditional information.

But one of the things that -- it's a shame I can't get it without the sound -- you see people's hands interacting with the tablet. The tablet was designed for one person's hand. The intimacy between each other. And they're not sisters. They're people just going to church once a week with each other. The intimacy between them has adapted.

They have a wonderful turn taking between them. It's really a lovely video because hands are coming in from all over the place. And they've learned that because they have embodied the conversation between them.

Now, that's one age group. There's a little girl there. But most of the women there are similar ages. If we now take reasons in conversation between different age groups or between genders or between people with different titles, we see a whole different set of reasons. My students have to work incredibly hard to critique and come back to me because it's not comfortable to them to question the people that they see as elders and authoritative figures. I have to get them to work with ACI the way they expect them to work with a critical voice.

I don't feel very comfortable at that to be honest. I'd like to this that we put out there like a conversational turn taking model that is used in that AI learning model is also conflated to something else we have to do in Africa.

So much like Apollo mentioned, when her -- when -- the mixture of the values, also, this internalized sense of inferiority with people who are more educated, people who are clever enough to design tablets and interfaces.

I want to end on an up note. I'd like to -- why industrialized educated. And in our weirdness, we are rather constrained compared to Africans. It's very rare when I go to a rural area in Africa where I'll meet somebody who doesn't have some fluency in another language other than their home language. But there may not be English or another colonist language. It may be the language of the next town or the next region.

African people, and I suspect this is true in other colonized places are monsters at language. But then masters at spoken language. In fact, there's masters inventing new languages. And I would like us to think of that as a resource for our own creativity and partnership when we design AI.

I'd like to leave that to the issue of the village next to mine where there's been lovely things. I'm a bit careful when I describe this. It tends to become a romanticize, oh, look at those poor Africans building things from things -- because they haven't gotten anything else. What I'd like to point out is the creativity involved in that. There's a group of people who have made a welding machine from microwave parts.

And the same people who have also made a sandblasting machine from old washing machine. So they can do this. They can do it. The problem is that for us to make AI, so, for example, an African person could show us the genre of turn taking that we need to make the interactivity work, we need them to be able to pull it apart like they can the washing machine or the microwave.

Otherwise, they are going to remain excluded. So I suppose my end note is we need for the creativity that consoles some of these ethical dilemmas for us. We needy verse people involved in AI. Diverse on all levels.

>> JOSEPH KONSTAN: Let's take a moment just to thank all of these panelists for their contribution to the conversation. And now, it's time to incorporate your contribution. We're going to bring up a document where I have attempted as best I could from the combination of what I knew before that people were going to talk about and what they've been saying live in realtime to digest this into a set of possible guidelines.

We're going to very briefly mention those. And then, I want your input, your edits. Don't worry about language, worry about principles, your suggestions for what's missing and your questions. These are at the moment in six guidelines.

One, this principle of the ability to examine what is inside, take apart. Second, accessibility, usability, and cultural suitability. How do we make things designed so they can be used by

wide ranges of people across some of the issues that people have brought up here?

Third, perhaps this sounds more governmental than anything else, but impact analysis. I took several of these different points of saying that if you're going to invest public infrastructure resources into some project of -- in this AI space, there should be a study of what are the safety implications, autonomy implications, de-skilling, redistribution of resources, you know, before you make the investment, not discover what happened afterwards.

Fourth, this notion that was brought up of human accountability and control. And specifically, that there should be some mechanism for not only giving control, but for points of accountability and liability.

Including ethical criteria in the programming, fail safe systems and evaluation. And this now broadened to make sure it extends to ethical criteria that are community-centered and not just, certainly, universal principles or valuable or insufficient.

And the last one was the catch-all of the idea that governments, NGOs, industries and others should articulate and fund high-priority areas. I've captured the ones talked about here. Avoiding danger to health and human life. Working in areas where algorithms can outperform human experts, like certain areas of diagnosis, scaling up applications that require automation or automated coordination to deliver their promise, like literacy, tutoring and devoting some significant fraction of research development and deployment to serve challenges of the under resourced and underserved populations of the world.

I am sure I did not get everything. And I'm sure we did not get everything. Now, it's your turn. Please, go ahead. And introduce yourself.

>> AUDIENCE MEMBER: I'm from West Virginia in the United States. And although I'm a lawyer, I've been involved for more than 30 years in working with the software development and the advanced technology communities, more generally. And Dartmouth College with AI, I remember in the early days working with Bob Kohn. The organization Bob founded. And it was the early large people like Robert Minski and Bob McCarthy. They asked, what do you think about AI? You've funded this, organized all of the meetings. He said, well, you know, in the early days, people thought it was kind of magic, these things happened, you know. He was more technical.

But he said, once it's understood, it's just computer programs. Right? And people like computer programs in those days. What's advanced is programs can write programs, you know, and it's more interesting.

We had a program the other day and one of the panelists mentioned. I won't ramble, too much, occupy too much time here. He mentioned about the fact that you have the ability to manage the

information for medical purposes and that's a particularly interesting thing. When do you trust the -- people say, well, this robot told me, so obviously, they must know. And it's sort of a trust. You trust somebody, a doctor.

Well, a doctor told me. This is true in the developing world. And so, you have this trust. But then, my point gets around to, I chaired -- I was the moderator of a panel yesterday in a meeting across the way on the internet of things. Well, they're looking about when you have computer programs, they're embedded in things, like a robot, which is maybe some common project piece of plastic or whatever. But the brains of it is the program. We're talking about computer programs here.

The man had this whole system he developed. And the panel I was on was IT sponsored. And it was sustainable development goal 16. Peace, justice, and, you know, really trustworthy institutions. And so, he was talking about his example for his system was mental health.

Well, a couple of weeks before, I had a judge in my old law school in Washington for an iron tech law where he had these young people developing apps. By the way, that could be really helpful, the applications that addressed particular needs.

And one of them, they had a social application for people that had health and mental health challenges. And they tried to link them up with social services. They were like the person in the elevator you were talking about, they could be educated. Here's the script, they were afraid of approaching it. Here's the things you're entitled to, and here's where you might be able to get it. You might print it out and bring it with you.

So the whole idea was that the man that made the presentation was saying, well, the technology can handle all of that. That's fine. And my lawyer hat. The sustainable development goal is rule the floor. How do you balance that?

Say, for example, this app that helps makes -- pours out all of the information about the medical condition. There's no privacy, no nondisclosure agreement. Nothing like that. But it gets out to their employer and they get fired. So I would add through your point here. And this is probably not in your purview, but I'm suggesting it should. And it came up yesterday, too, when there was a little -- so respectable and legal. Legal and regulatory environments.

So ethical criteria, whose ethics? The ethics of the program author? Talking about copyright and patents? So, basically, you have to look at the law. And how that would integrate in this new environment.

>> JOSEPH KONSTAN: I think we get the point of adding legal criteria. I'm going to collect a couple of comments before we ask the panelists respond. We'll go one, two and more to follow.

>> AUDIENCE MEMBER: I've been using robots in critical care and working with students in nursing homes. And one of the interesting things is, I think, the ability to use robots as Avatars. And patients seem to like them a lot better than talking to television screens.

You have autonomous movement. And we find that, also, with robots and kids who can't go to school. Where you put a robot in the classroom and allow a student to be able to stay in school while they're going through chemotherapy or through the hospital. And I think one of the things that's been interesting to me is this business about privacy. Parents worry when students are having a robot in school.

Who else is watching the classroom besides the student? Are parents watching classrooms? I think there's a lot about the interactivity that we need to think about as we put, from my perspective, robot Avatars into multiple places where you get an opportunity to not only have the robot interact, but the person interact together.

So I think thinking through those things. One of the things I found fascinating, we put a robot in the classroom and the person I had the most trouble with was the teacher who basically said, I don't like having somebody from outside watching my classroom. Even though, she said it's only a student. She said, how do I know it's not the parent? How do I know it's not this? There's a lot about interactivity with robots in health care and education that need to be understood a little bit better.

>> JOSEPH KONSTAN: Okay.

>> AUDIENCE MEMBER: Hi, everyone. I work on a bunch of different projects, but I co-founded a nonprofit. And what we do is community building education and promotion of neurotechnology, specifically accessible neurotechnology. My question is not related to my -- not my question, but my point.

A lot of the concepts of the conversations that I've been seeing for the past couple of days is really, I guess, aligned with how -- what are we expected to happen in all of these different areas? And how AI's going to affect them. But the one thing I want to quickly bring up is maybe you just also making sure that, you know, as we maybe have a list of priorities of what we're going to be trying to implement as policies and whatever, I think there also has to be a focus on ensuring that --

You can only use so much planning, right? And once the actual implementation happens, there might be some side effects we never considered to happen in the first place. I think there has to be really rapid visions on policies that are made simply because impacted the AI system may be so quick to happen that, you know, we can't wait months and months for any revisions to happen. It has to be revisions a little bit sooner.

>> JOSEPH KONSTAN: We'll take one more here and then we're going to see if we have any remotes and get a couple more and give the panelists a chance to respond.

>> AUDIENCE MEMBER: Great panelists, thank you. Great contributions. I just wanted to go back to your principle number one. I am from the Gates foundation. My background is speech recognition. I appreciate the issues that was raised in terms of the research challenges there.

But I just think that guidance principle number one is too specific. Based on the comments from the panelists, the two panelists on the left-hand side, my left, we need to broaden it a little bit to something around human center design. The way you put it right now, just two facts come up.

Number one, conversational turn taking is universal. Right? So the idea of looking inside cross linguistically is true, whether it's African language or English. And then, we're able to look at issues of context and the variations that come out of context. That's the biggest struggle that AI and artificial intelligence is struggling with right now.

Our understanding of context brought into how we model meanings. That's missing right now. That's my comment. I think you call it scrutability. What's the second point, sir? Then, my bad. The second principle is what I referred to, yes.

>> JOSEPH KONSTAN: (Speaker off mic).

>> AUDIENCE MEMBER: Especially context.

>> JOSEPH KONSTAN: (Speaker off mic). And other environments (inaudible).

>> AUDIENCE MEMBER: Right.

>> JOSEPH KONSTAN: (Speaker off mic).

>> AUDIENCE MEMBER: One more point. If you look at it programatically, put yourself in the position of the designers. If you do that, it becomes syntax.

So issues of all of the disparity you have in this cross linguistically because we will never be able to have machines that can really do what we're seeking to do.

>> MADY DELVAUX-STEHRES: A Ph.D. student working on Namibia identity at the moment and is running focus groups about different ideas of privacy when there's interoperability. And one of the things -- one small element that can indicate how concepts of privacy can impact on these sorts of things is a situation he has remaining at birth. So -- in many cultures and it's certainly not African -- just African cultures, many cultures will not register a birth, but they will not publicly express the name of the child for some time.

And that makes -- so when you're making registering the birth and making identity documents at that point were you to then put the name on those identity documents, then. You'd actually be

culturally quite undermining.

This is a small example about how the privacy concepts differ across cultures.

>> MADY DELVAUX-STEHRES: I would like to adhere to your words because I think for normal people, I consider that I'm not an expert. So I take the reactions of people who are artificial intelligence in their work, in their private homes, this is a big concern. How can we protect privacy? If robots are in your home and there is also for the workers in industry.

For instance, the robots will control every action they make. And how can we have (inaudible). I understand a robot cannot function without collecting images and data. So I think this is the most difficult issue that we -- I don't know the answer how we can organize except say it should be proportionate and this is a very general principle. And I don't know how to apply it.

And concerning the legal issue, I cannot agree that we have to look at the legal issues. But, we have -- we must not forget, either, that laws can be changed.

And there needs to be consensus or common understanding of what should be legally allowed or forbidden. And this is also a discussion we have now because new technology and new questions. And we have to find solutions. But not have the education too early because we don't know exactly where we are going to.

I think the dilemma of (inaudible).

>> JOSEPH KONSTAN: I want to make sure we get the people who have not had a chance. I see three I'm going to make sure we get to in the time we have. I do want to make sure that we give you a chance collectively to agree or disagree with these before we present them.

So we're going to go one, two, and three, and we'll see what we have time for.

>> AUDIENCE MEMBER: Hello, thank you very much for your presentations. I'm a representative of Geneva innovation. It's an association. And first of all, let me start with -- I'm a sociologist and there's a lot of things I don't understand. But nevertheless, I want to pose the questions I have.

So the first question is, referring to how important it is to allow for this model of interactivity. So I'm kind of -- is this way of interacting and working on this the best way to go forward no to tap into the creativity of the people within the room?

The second thing I want to refer to is the new reality. My question would be, in co-creating the reality, is artificial intelligence going to be part of the co-creation? I'm talking about artificial intelligence because we're already mentioning that if we talk about knowledge. Then you need to consider.

And the last thing I want to say is the guidelines and impact analysis. And because the perception was that if we put the

public stance into the analysis, it means it's too late for me. There's a guideline of foresights and core creations of principles that would lead the way of artificial intelligence being developed in order to kind of have principle to come forward.

And I think it would be wise to consider.

>> AUDIENCE MEMBER: I'm from the University of Cambridge. Very much these issues, and these are a great set of guidelines and you've done a really good job of distilling the important principles we need to bear in mind.

I was wondering if we could just have another look at number 6. I think it was number 6 or the one that mentioned ethics and we're going to move quickly, number five. Thanks.

I was worried a little bit that referring to including ethics in our guideline is passing the buck a little bit. If we're not talking about what we mean by ethics. Industry standard principles are our ethic with regard to how we think these systems ought to be developed. And I wondered if that guideline was sort of passing the buck to the next committee.

>> JOSEPH KONSTAN: (Speaker off mic).

(No audio).

>> AUDIENCE MEMBER: People could disagree on how far we've got. So I wonder if we might think about a moderate version that focuses on ensuring that a given system reflects the values within the community in which it's going to be implemented, for example.

Refers more to understanding the particular values of the context in which it'll be used.

>> JOSEPH KONSTAN: We'll put universal in quotes to make sure. Because universal is probably not (inaudible).

(No audio).

>> AUDIENCE MEMBER: Sorry. I really am intrigued by this idea of the universal set of values or maybe even a step further in thinking about the universal that writes around our interactions with machines.

I think the biggest concern is that really our AI reflects our values, our society, our differences as well as sort of the speed of the development. And I think those are both to make sure our system evolves appropriately. Advocating for some work to be done to hold that high standard of what industry should be accountable to and what we should all be accountable to.

>> JOSEPH KONSTAN: She's raising the point that this is the concept of universal set of values is a challenging one. And she's intrigued by that and the notion of, perhaps, a universal set of human rights with relation to her interaction with an AI technology. The stuff is moving rapidly. Getting on top of before it's already there being important.

And I think something -- I hope I can say we would all agree on is that there are things that are aspirational that we don't

know how to do, both in technology and in ethics.

And, you know, there was a huge benefit he could write rules of robotics and novels in which they would be enforced without having to be able to code them. And even then he could find that they would be coded around such rules. And even a simple thing as saying, a robot can't harm a human.

Ethics to face that dilemma when you have your -- you're on a speeding train going down the track, you know, do you let it go and hit five people or pull the switch so it only hits two? And the robot would be crippled. Well, I can't let it hit five people, but I can't hit the two. So the robot destroys itself and the train hits the five people.

We know these are hard problems. And this is part of the reason that the philosophers and the ethicists as well as the technologies need to work together.

I understand there are at least three of you who are now signaling you have more things to say, and I'm going to invite you up at the end before this, but I'm going to do the very fastest show of hands we can do before you escape to lunch. If we can just quickly start at the bottom there. Without getting into all of the tremendous details, does this group feel it is a wise thing to have a guideline that suggests that priorities for areas of great positive impact should be set to a large extent by governments, by industry, by NGOs, foundations, nonprofits? By the people who are going to likely invest in making these a reality? We can quickly go over what the priorities are.

If you agree, I'd love to see a hand. If you disagree, I'd love to see a hand. That's good to hear. And if you don't have a hand -- we're not going to there. But we all know in any room, there's a bunch of people that don't like raising hands and that's okay.

Do we agree about the notion of encoding the ethical and legal criteria into the building protecting and evaluation of robotic and AI systems? Recognizing all of the challenges of developing the legal criteria? Agree, disagree?

>> It would have to change if you get into ideas of freedom of expression.

>> JOSEPH KONSTAN: Yes, you do. This is why it's good to have somebody with a legal background here. Do we agree with human accountability control liability? That we need -- that the humans stay in charge? Or do we think it's important that systems become autonomous without human control? That's the alternative.

A lot of people believe in human control here. How many people believe that's sending us the wrong way? We should be going for autonomy?

>> AUDIENCE MEMBER: We're already doing that. You have computer programs as information systems. Say you have a Tesla riding

around in the car. And --

>> JOSEPH KONSTAN: We understand the concept. We're going to go through this quickly.

>> AUDIENCE MEMBER: I know. Each one --

>> JOSEPH KONSTAN: We're not taking comments on each individual item. Do people believe in the concept that we should be focusing or asking for impact analysis before the governments of the world get involved in investing and supporting these technologies? Yes? No? I'm not seeing a very strong positive. I'm seeing a mild positive.

Human centric design accessibility, usability, cultural suitability. Is this an important guideline? How many would say no? That one's closer to universal here.

Examability, transparentability. People would say that's a yes. A no? I want to thank you. I'm sure our panelists would be happy to stick around for a few minutes if you have further questions if you want to ask them about their areas of expertise and experience. I will stick around here, as well. And please do continue to participate as we pull this forward over the next three rounds of these sessions and try to evolve towards something that we can support. Thank you all.