

FINISHED FILE  
AI FOR GOOD GLOBAL SUMMIT  
JUNE 7, 2017  
4:00 P.M. CET

PLENARY 3: FUTURE ROADMAP - COLLABORATING FOR GOOD

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 is being provided in a rough-draft format. Communication Access Realtime Translation (CART) is provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.

\*\*\*

>> ANJA KASPERSEN: Could I ask everyone to please sit so we can get started with this exciting session? Thank you. So my name is Anja Kaspersen. I work for the International Committee of the Red Cross where I am looking at various ways that AI can be used in humanitarian activities and deliver humanitarian action and how it is being used in warfare.

With me I have my co-Moderator Nicholas Davis and hopefully we will create a flow in bringing in a lot of interesting and powerful voices on the issue of AI and how we move forward on some of these issues and opportunities we have been hearing about today.

>> NICHOLAS DAVIS: Thank you. And thank you everyone for being here. I'm Nicholas Davis. I lead the World Economic Forum Society and Innovation Work focusing on innovation. And they come together in the discussions we are having today around Artificial Intelligence for good. I am delighted to be here with Anja and a fantastic round of speakers on stage, and then a panel discussion to come to discuss the how.

So we met this morning in the opening Plenary and we talked about the current state, the current state of play for Artificial Intelligence. And in the Plenary just come, just been, we spoke about the what to expect from the future, how there are certain aspirations and concerns about what may unfold, but the most important part of the next two days and

this evening is to focus down on what we need to do.

Artificial Intelligence algorithms are being developed and released and embedded in to products. How can we come together as global stakeholders to think not just about people who are lucky enough to come to a global meeting like this in Geneva but the people that Salil Shetty talked about in the last Plenary and people who have less voice and with those people that Anja is often concerned and working with at ICRC.

>> ANJA KASPERSEN: Indeed. When talking about solutions as we are trying to do during this conference it is important to develop a shared narrative where we address what -- not just what is at stake but what are we trying to address. What are we trying to address for whom by whom. The impact of AI alone are in convergence with other technologies which is one of the issues that I work with. If you are on the streets fleeing a war in Iraq or standing in downtown Geneva trying to hail a cab. So addressing the as much as there are opportunities there is also risks and hopefully we will get to some of those risks and dilemmas and controversies that we have to address to ensure that using AI for good. And we preserve those that are the most vulnerable when they need our protection the most.

>> NICHOLAS DAVIS: In this session you are really getting two Plenaries for one. This first section we have five fantastic speakers. We are going to start with Robert Kirkpatrick. And we are going to move to Katsumi Emura, Manuela Veloso, Cindy Smith and Stuart Russell. Importantly to frame the discussions over the in next two days as we move in to breakout discussions and more focused groups for thinking about what it really means for Artificial Intelligence to act for good.

So with that thank you, Anja. You will see Anja who will come and bring our panelists on stage for the second half of this discussion. I want to make one moment and say we are in Switzerland. And I have lived here two years. Anja will bring us to close at 1815 this afternoon. We are going to focus on crisp and concise perspectives from our speakers and thought provoking discussions from the panel. Please don't think of me as Nicholas Davis as the Moderator who is human and understands things like social norms. Think of me as someone who is an Artificial Intelligence agent with an iPhone timer in their hand whose purpose is to make sure we get to drinks and enjoy each other's company and discussions after two hours of fantastic conversation.

And with that I want to invite Robert up to the stage. Robert is not just the person who runs UN Global Pulse, the Big Data focused aspect of the United Nations but also is an entrepreneur, a social entrepreneur in many senses and someone here who can let us know what it is that we need to do

particularly in the use of data with regard to Artificial Intelligence. So Robert, please welcome to the stage.

(Applause.)

>> ROBERT KIRKPATRICK: Many thanks. And good afternoon to all of you. I was asked to speak, of course, today on aspects of a roadmap for Artificial Intelligence for social good. Given that although I'm an optimist with AI, sort of this is an ostrich or a Pekinese I am going to be focusing on narrow AIs and thinking about that roadmap as a start. So we live in a world in which everything increasingly is smart and everything we do emits data as a byproduct of our actions. At Global Pulse Big Data and all this information out there is as a new natural resource. It is abundant and renewable through the power of Artificial Intelligence and high performance computing and can be refined in to new forms of information. Tremendous value for answering business questions and for fighting hunger, poverty and disease.

Our labs are using Big Data and deep learning to map discrimination against refugees in Europe and recognize rescue events from shipping data in the Mediterranean and detect fires in the Indonesian rainforest and predict food prices from Twitter, fight bottlenecks in sanitation services in Satari refugee camp and generate socioeconomic indicators from satellite, imaginary and flows of postal traffic.

Lots of data and algorithms that can be used for good. We have heard a lot about speech technology today. Talk radio programs are the oldest and most original social media platform out there and still the primary one across the world. We are using to generate realtime insights from public remarks on FM talk radio shows. No one had grasped the potential for spotting things like cases of Ebola or small scale natural disasters that would otherwise go undetected. These are small examples compared to the advanced institutions in the wealthiest part of the world are taking. Our new natural resource has fallen in to the hands of an extractive industry, one that is opaque and unregulated. Its benefits are not reaching those who need them the most, and the expertise needed to refine these material are not affordable to all but the wealthiest.

And finally this resource is a bit like nuclear energy. It is a habit of leaking sensitive information in ways that can contaminate and cause harms for years.

So how this information can be used in cyber warfare tools and autonomous weapons. In 1950s when we had seen the horrors of atomic bombs but had scientists tell us that we can use this to safely power science for years safely. How do we contain it safely and harness its power for good and how do we govern it? There is analogy that we can build on.

But that isn't the only challenge. It is not just leaking. We know how hard it is to anonymize this data and also know that bias can be introduced in to it either by the data used to train it or by the purposes of the developer who originally designed an algorithm. There are huge challenges in transparency. And we heard some of that today.

Maybe the best efforts we will have will still have an algorithm telling us why it said a decision. At the end of the -- there is a profound lack of trust and we are facing issues with it where we don't know what the next step is. Rupert Stadler this morning spoke about self-driving cars, making a life and death decision. What if some portion of the society says we think that algorithm should have access to everyone's Facebook pages, how people feel about this and what they buy with their credit card so they can evaluate the value of their life. How do you make a decision, and some portion of people will say we don't want machines making life and death decisions about the value of human life. And yet we are having to program that right now. And others will say if we don't it is not fair.

What about privacy? What makes this data and these technologies so powerful also makes them dangerous. It is not easy to separate. Nevertheless though the opportunity here is unique. Even the relatively narrow AI applications we have heard about today have the opportunity to be used to change the world. It is probably clearer to all of you that the ambitious nature of the Sustainable Development Goals means we are not going to be able to achieve them unless we transform how we make decisions and have massive increases in productivity. Looking to define a roadmap we have to first agree on where we want to go and while we are at it we should make defining explicit nongoads part of that process.

The next step is ground our roadmap and principles contained. So basically access to health care, education, nutritious food, clean water. These are Human Rights. Our approach on the one hand must accelerate adoption. AI must as a tool of empowerment of everyone on earth but on the other hand, it must also adhere to the principle of first do no harm. We aren't doing well at this today. I would argue we need to make sure that AI gets diverted from the path that it is currently already well on, which is increasing inequality all around the world. At the end of the day we are navigating a course of risks of using these new tools and risk of not using them.

In our own work we are encountering cases where a particular opportunity to do good would be valuable but we can't do it because the risk of misuse is too great. That is only going to increase as we look at the applications of AI. So a roadmap, I think divide the roadmap in to two categories. One is how do we

seize the opportunities and the other how do we mitigate risks. The one could start with a fund. Imagine a global fund for AI. Establish potentially as an international institution within the UN. This fund could invest in cutting edge research and promoting AI based tools directly addressing the most pressing challenges of the SDG and scaling proven solutions. Ideally through open source tools where possible. We heard earlier I don't know where to invest. I don't know what the best ideas are. Maybe such an organization with a partnership for AI could help define a common agenda for where the most -- the highest priority investment should go.

Government should invest in building the capacity of their public sector institutions and Civil Society Organizations and promote computer science to increase the pool of talent out there. We are going to need them in an operational capacity all day long as we start to build these systems. I would also say, you know, on the problem of we need to make sure we are investing and not increasing inequality, of course. And, you know, if you look at the opportunity, if you look at the challenge of preventing automation from taking jobs, I'm skeptical that simply giving people universal basic income solves the problem.

Finally we know we are going to need a multi-stakeholder process to get to principles of AI. We could ultimately feedback on how well that's being implemented. We could see international soft law be put in place and one day perhaps a Convention. I look at the existing rights instruments that are out there today and I don't believe that any of them will adequately address all of the risks and opportunities around AI, in particular when we move to a generalized AI.

Last point simply is that I think that what needs to happen next is a series of demonstration projects that can show the value of Artificial Intelligence at the -- and its cutting edge state in ways that the public can understand that also demonstrate the very real and frightening risks and ways this can be mitigated. If we don't address the trust problem we won't be able to move forward. Thank you.

>> NICHOLAS DAVIS: Thank you so much. Stay up here for a second. As I am sure with all the panelists and speakers there are a hundred questions I want to ask as a follow-up. One particular one, do you have an idea of a demonstration project that this group of multi kind of faceted stakeholders from around the world should start to get behind or design over the coming days that can do what you mentioned? Illustrate the potential but also the potential for misuse?

>> ROBERT KIRKPATRICK: I think one project that could be done, for example, would be learning -- I'll give you an example

because we are having some discussions about this sort of thing. And it was raised earlier by Dr. Margaret Chan this morning. What if you could bring together all the data on what people are buying and selling, the movement through cities, particulate information in the air and distances to parks, to gyms, to McDonald's, availability of biking lanes in cities and map down to the neighborhood level the risk of noncommunicable diseases? So say if you live in this neighborhood, these chance of ten kinds of cancers, because of the built environment and diabetes you could begin to understand. If you live ten minutes away, your children have a much lower chance of dying of Leukemia. It could change real estate prices, insurance rates. It will put huge pressure on mayors around the world to alter the built environment. It would empower citizens in a way they never had this before to understand what it means to live where they do. The only way to do that is bringing this data together in a privacy protected way and running a lot of deep learning on it to be able to understand the risks. And then we need to deal with the consequences because the truth sometimes is a problem.

>> NICHOLAS DAVIS: Thank you very much, Robert. Please thank Robert Kirkpatrick from UN Global Pulse. Please.

Robert, you can go and sit back down. Otherwise she will be left alone at the panel and it will be like an episode of Survivor of some sort.

Professor Manuela Veloso, thank you for taking the time to come here and join. Professor Manuela Veloso is the Herbert A. Simon University Head at the school of computer science and leads the machine learning department. If you ever Googled Manuela, you will see the exciting work that got me interested in the augmentation of humans of Artificial Intelligence is the work of the cobots and robots. If you are not going to talk about it I will leave that for the audience to talk about. Thank you very much.

>> MANUELA VELOSO: Thank you. I'm going to tell you a little bit about another form of AI or somehow the form of having robots that move in our environments. If I could have my projection. First time I'm giving a talk in which someone is projecting what I want to say. But that's okay. So what I'm going to tell you is this thing about we kind of understand AI in terms of cell phones and computers and a little bit drones and Internet of Things, but I really want to talk a little bit about what is this thing about having robots that move around. Robots that move combine the ability to perceive the world. They have sensors and computers that enable us to think and they have action in which they move. This is a robot, cobot, collaborative robot that moves at Carnegie Mellon of performing actions and has moved at CMU.

In this talk I am going to very briefly in the eight minutes that I have to tell you about three things, three things. One is, in fact, that these machines that move around need to be able to localize, need to know where they are and they need to actually plan for routes and execute them.

And then I'll go in to two second topics. So in this first topic, if the slides move, let me just bring you to the core of what is robots or these AI machines are all about. So they have sensors and these sensors are not our eyes, not our touch, are not our ears. I mean it is a different world. In the robotics world they have sensors and in this particular case I am giving you an example of a Lidar, which is a scan of rays that tell you how far the obstacles are. And, for example, they have a 3D camera that gives also what is like the scene in 3D with very limited kind of distance, six meters for the regular connect or maybe can get ten meters. And you get these obstacles which are the walls.

With this information they could have other sensors, but these are common sensors in robotics. With this information they basically have to be able to transform the image which is on the left side which is a depth image in to eventually mapping whatever they are seeing, the distance at which they are seeing walls to a map that they are given or that they learned. So somehow robotics and a mobile robot has to map whatever they see with this very limited sensors to whatever they know which is the map of the environment. So the way they function it is actually through these -- oops. The video that I was going to show, so they are supposed to map and they actually move at Carnegie Mellon as you saw before, moving four kilometers. They move for more than a thousand kilometers. If you come to Carnegie Mellon tomorrow, you will be escorted by a cobot. A cobot is escorting some visitors which I scheduled yesterday and it is escorting visitors today.

You don't come to my offices by knowing my office but a cobot and you follow the robot that takes you to the office and which it does.

So that's a very interesting kind of phenomena. These robots don't have arms, and it is kind of like understanding that is like some limitations that are embedded. They don't go upstairs. It might be hard for them to open all the doors of the world. So it becomes this kind of like realization that our robots have limitations.

So back about five years ago these type of like realization led me and my students to introduce this concept of symbiotic autonomy. These robots ask for help. So these robots are left in to the environment. But when they find something they cannot do, they don't understand, they are lost and certainty is too

high. They just stop and ask for a human to generously help them in pressing an elevator button and telling them where they are and put things in their basket. They navigate autonomously in this kind of symbiotic with the other robots because there are four of them and with the Web.

So if the next video works, I'll show you an example of these in which if you put the sound -- well, you can tell the robot please bring a coffee to the left and the robot says I don't know this object. What is coffee. And I'm going to check what's the probability of coffee being in these locations.

>> Also for object coffee which location do you think is best.

>> MANUELA VELOSO: The human says go to the kitchen to get coffee. This robot goes to the kitchen and he is going to the kitchen to get this coffee. Now I can ask you how is it going to get coffee from the kitchen given that it doesn't have any arms. It can't do anything like getting coffee. The way it does it gets to the kitchen and talks to the air and has no clue if there is people there and it just says hello. Could you please put the object coffee in my basket. And then there is someone in the kitchen that very generously actually does that and puts coffee in the basket. And then says okay, now I can go.

So there is this kind of like now we can have one day the robot have arms and it might not ask any more to put things in the basket. But up to now all these things it can't do yet and many of the things it has been improving on it asks for help. And it gets to the lab and says I brought coffee to the kitchen. Tell me what I do and there you go and you are done. There is this new paradigm in which the robot asks for help.

So the important thing here to understand in some sense which is novel with respect to other things in which robots had like being supervised is that this is a proactive ask for help. They find out when to ask for help which is specific generalized situations. And then eventually they substantiate them as needed, but it is not that anybody is supervising this robot. And then finally -- and they actually learn from this interaction. And I want to tell you something that's important. So those robots started knowing nothing about which objects were where or where Manuela's office was and everything was learned. Hundreds of facts were learned through this interaction with humans, human machine, the kitchen and then it saves mapping. And this robot has acquired all this knowledge about the environment. We can instruct it in policies.

One final thought when this robot appears from the front of someone's office and this is the last difficult problem about robotics and AI, you wonder you did not see anything. It just stopped in front of your office. And there is this nightmare



about finding out what happened. So when a cobot stops in front of my office I want to know what happened while you were on its own. Did you escort successfully Stuart back to the elevator? Did you -- what are you going to do next? How many times did you go to the kitchen today? So the autonomy and this AI has to -- we have to realize that they are going to do things on their own that we actually don't know about.

So we have introduced this concept of verbalization in which the robot can generate automatically at the description of what it did, I went straight for 15 minutes and I turned left and then there is verbalization to generate explanations of different levels of abstraction and the robot can give different levels of detail in its explanations.

So in conclusion the three points I wanted to make, if this goes, so in conclusion, so there is this great challenge or this great opportunity for researching, developing and really producing things on human and AI interaction. And we have been doing this on humans and web and robots this coordination and this task distribution but, of course, hopefully for the joint benefit, co-existence and transparency which is something that we are all researching on to have these machines that learn, and they have to do things automatically, explain themselves when humans express explanations. Thank you so much.

>> NICHOLAS DAVIS: Thank you for summarizing what has been more than seven years of research in eight minutes. But here you have an audience of stakeholders from Civil Society, from the tech world, from the UN, of course, from other industries. And I guess the question that's burning in my mind is what new areas or extended areas of research should we be thinking about or investing to extend the lessons that you learn about collaborative robots in to helping a broader set of stakeholders and not me or us when we come to visit you at Carnegie Mellon.

>> MANUELA VELOSO: I do make the case let AI be mobile. I entered this building. There was not a single robot moving around itself. Go to the hotel and there is no robots moving around anywhere and we are surrounded by AI technology that does not move. So my first point is like if you can in your companies believe that, in fact, this mobility is possible. And then the second, I mean the second case is, in fact, let's acknowledge that AI systems are not perfect. And it was going to take a long time for them to understand all types of scenes and deep learning. There will always be something that they don't get it. There will always be something they cannot do and what about thinking about these machines as being able to ask for help and say I need more data. I cannot classify these anymore. I need to ask another -- I need to ask a human. So these -- so the first thing is mobility and the second thing is acknowledgement

that we can have AI for good. AI deployed for great results even if we have not solved the whole thing yet. If we enable our systems to ask for help and declare they need more knowledge. And the third is not make these black boxes and always try to have the systems be as much -- explainable as possible. And this is a research question. And so that's basically the three things that I think we can all work together on is these aspects. Okay?

>> NICHOLAS DAVIS: Thank you so much. Put your hands together for Professor Manuela Veloso.

(Applause.)

>> NICHOLAS DAVIS: And now it is my pleasure to introduce Mr. Katsumi Emura who is the executive vice-president and chief technology officer for NEC Corporation and the Strategic Council for AI Technology. We are shifting here to thinking about the roadmap on a national level and fascinated to hear your views of what we can learn from Japan's experience here and where we can take it in to other countries. Please.

>> KATSUMI EMURA: Thank you. Good afternoon, Ladies and Gentlemen. As Nick introduced me I'm from industry. And actually the NEC Corporation but today I'm presenting for the Strategic Council for AI Technology in Japan. That's today. I introduce Artificial Intelligence industrialization roadmap developed in Japan for realizing Society 5.0. I will explain Society 5.0 breakdown. Can I get my slides? This is not mine.

>> NICHOLAS DAVIS: That's Stuart Russell's slides.

>> KATSUMI EMURA: So when we are discussing or developing a roadmap there are two ways. One is on kind of forward casting. So in a technological people always use the forward casting type of roadmapping. But in an industrialization roadmap in Japan we are using a broadcasting approach.

So in the broadcasting approach we have to advance in the society in the future. This is one example. This shows a global state in 2050. The population in 2050 is expected to be 9 billion. Maybe 30% increase from today. But because of the urbanization the people living in the state will be increased 1.8 times, resulting in a huge increase in resources like energy, water and food. How to cope with demand is our challenge.

And the SDGs, the 17 goals, this -- these are also showing our directions. And NEC is committed to contribute to SDGs. And -- but we have seven themes for value creation. And the -- these themes are pretty much aligned to the SDGs. The question is how to -- how do we create social values and roadmaps that align with these targets. And to create social values we have to use AI technology in the sense of visualization, analysis, and prescription.

And the -- by considering these we are developing the

industrialization roadmap. And the -- this explains Society 5.0, the society which we want to realize in the future. The Society 5.0 is a Japanese vision for upcoming society. Following on from hunting and gathering society that is Society 1.0, Agrarian society, industrial society and information oriented society. Information 5.0 refers to a social mode that realizes the advanced fusion of cyberspace and physical space. The second provides goods and services regardless of location, age, sex, language or any other indication. And thirdly bring about the human centric society in which people can lead high quality lives full of comfort and vitality. So realize Society 5.0, evolution of Internet of Things and Artificial Intelligence are the keys.

The Japanese Prime Minister announced in April last year saying that the Strategic Council for AI Technology is established to develop an AI roadmap by gathering the top experts from industry, academia and government.

The aim to create an AI industrialization roadmap is to promote the social issue resolution and enhance the Japanese industry competitiveness. First we define social issues to be solved. Then we envision the future society. For example, as I mentioned to overcome the resource requirement we must realize, for example, zero waste society. Under this vision we develop the roadmaps.

There are many potential areas to develop a roadmap. But considering the next step means that to realize a research project, those kinds of things. We try to pick some of them to realize the roadmap. The criteria were really challenging social issues. Contribute to achieve Japanese GDP target and AI is really contributing. And then we selected three areas. Those are productivity, health, medical care, welfare, and mobility.

And the -- this shows the vision for the society under these three areas, productivity, establish, an ultimate ecosystem that will enhance productivity in society as a whole, meaning from production to consumption. The health medical care and welfare create a society where people at the age of 80 are able to work actively. We may redesign social and industrial systems. Mobility, get rid of accidents caused by human error.

And this morning we heard that 90% of accidents are caused by human error. So minimize the social cost of transportation and the -- when the autonomous driving has become a reality we could use moving spaces for workplace or entertainment.

When we develop the roadmap, we have a technology trend in mind like semi-conductor architecture for Artificial Intelligence, the kind of data, and the multi-core architecture and their functionality. This shows the whole aspects of roadmapping activities. As I mentioned we see social challenge

and technical trend. And the -- with that we like industrial -- the industrialization roadmap in a phased manner.

So this shows the phase. The phase 1 is approximately to year 2020. We already use Artificial Intelligence but in individual domains. In phase 2 to be around 2025 to 2030 where we apply -- we can use AI's cross domains. So the different domains are connected. And in phase 3 Artificial Intelligence as a service becomes common and with that any industry is connected easily and wholistic solutions are becoming available.

This shows the whole view of the industrialization roadmap. Each squares indicate not the technical goal but the indicator, what we want to realize in the society. So our roadmap is indicating as such.

So this is one example. Maybe you cannot read. And the Japanese materials are already issued end of March. And good news the English document will be available soon. So -- and you can read it and check the detail.

So in conclusion we are seeing the Society 5.0 and create the industrialization roadmap. And then the next step maybe we do an innovation program that is a research program and to realize social barriers. I believe sharing the vision and roadmaps are an effective way to align activities and good use of Artificial Intelligence technologies.

So thank you for your kind attention.

(Applause.)

>> NICHOLAS DAVIS: Thank you very much. So this is I'm sure fantastically detailed but also represents a lot of discussion and debate amongst Government, industry and social partners as well. Can you maybe just briefly let us know was there a particular technique or lesson about how to bring together those stakeholders across industry and Government and Civil Society to come to a common vision for Japan's future with Artificial Intelligence?

>> KATSUMI EMURA: The key is to start discussing the future issues, not the technology. And the -- for example, like on mobility, to reduce the human error based accident. Then we start to discuss what technology can contribute. Then we start discussing autonomous driving or mobility system. So with that different stakeholders can discuss on the same issues. So that's a way we do it.

>> NICHOLAS DAVIS: Thank you very much, Mr. Katsumi Emura. Thank you, Katsumi Emura from NEC and the Strategic Council on Artificial Intelligence. You can see there is a trend starting to emerge in the discussions about setting social goals and understanding some of the technical aspects about how our society and human aspects relate to the technological capabilities and data and other aspects below.

I want to hand over to Cindy Smith who is the director of UNICRI, the United Nations Interregional Crime and Justice Research Institute. And Cindy mentioned and will mention I'm sure the fact that the center -- that UNICRI has announced a new center focusing on Artificial Intelligence.

So it is a delight to have you here. And we look forward to your remarks on what we need to take care of as we think about the impact of AI on multiple parts of the future as well as, of course, crime and security. Thank you.

>> CINDY SMITH: Thank you very much. Just in case I don't get all the way through my message in my eight minutes I want to tell you the important part first and then I will tell you how I got there. We were asked to talk about how, what do we have to do and what we believe at UNICRI has to be done is the Cs. My name is Cindy. So you will remember this. The Cs, conversation, collaboration, cooperation, coordination of and as you just heard from the last speaker private industry, academia and Government. Putting those three partners together at the table and making sure that they have a conversation, that they collaborate and that they cooperate and that they coordinate.

With that in mind, Distinguished Guests and Colleagues and friends, thank you for inviting me here today. It is an honor to be here. In my short time with you I want to touch upon three key issues. I will talk about first the benefits of AI. Second the risks associated with it. And third the necessity of international cooperation to maximize the benefits and minimize the risks. The first of these issues, potential of AI to bring about benefits to global development and societal change has been the topic of discussion so far. We are living in truly an extraordinary moment of history. The world nurtured by science and technology has become increasingly interdependent and interconnected. Rapid advancements in the field of robotics, coupled with the rise of computing power have exponentially increased the breadth of tasks that can be assigned to robots and systems based on Artificial Intelligence as well as autonomy with such technologies operating.

We are only beginning to understand how these advancements can really be applied. We at UNICRI believe the benefits of AI will be of paramount importance to society in the years to come. Of the United Nations Sustainable Development Goals, for instance, AI may be used to identify and correct for gender bias in the labor market. A wide range of critical fields of relevance to the SDGs such as health, education, energy, economic inclusion, social welfare, the environment as well as crime prevention, a topic close to my heart, security, stability and justice all stand likely to benefit from the progress made.

The second issue is that advancements in AI will lead to

serious challenges and risks. The wide range of economic, demographic, environmental, legal and political implications of these have yet to be truly understood. Looking at some of the challenges and risks we should start to think about what new threats to international peace and security will emerge from the advancements in AI technologies. If you build it my folks will use it. And my folks are people who usually end up in corrections. In fact, some of these technologies have already been used for terrorist purposes. We will -- what the -- what will be the availability of new instruments of terror mean more the international community as it strides against the tides of terrorism.

Are we ready for this kind of terrorism? The recent increase on cyber attacks takes a whole new dimension. We must capitalize on AI in this fight against crime and terror and find ways for AI to contribute to enhancing global security.

Moving beyond, in a strict traditional security context we see even more challenges and risks. We see, for instance, that AI is beginning to disrupt the global workplace, change the types of jobs available, and reshape the skills that workers need. This leads us to the question how these changes will be felt in the developing world and what it will mean for labor migration. Whether AI leads to unemployment, increases in inequality or new waves of crime and terror will depend on a proper understanding of the global implications of AI. Having looked briefly at both the benefits and risks I turn to the third and final point that I would like to emphasize and that's the importance of the Cs that I spoke of. The cooperation, as we strive to strike a balance. We in UNICRI believe that given the technology complexity and the rapid rate of development in this field we can only succeed to minimize the risks and maximize the benefits by bringing together the various communities unifying national and international public and private, academic and practitioner communities.

In 2015 UNICRI launched its Artificial Intelligence and robotics program. The first at the United Nations General Assembly. Firmly positioning cooperation as being central to our approach. A High-Level Event with preparation of world renowned experts in the field of AI on the 7th of October in 2015 at the 70th session the United Nations General Assembly marked the beginning of this program and some of you were there. Since then UNICRI has sought to advance understanding of the various aspects associated with developments in AI and robotics and contribute to building bridges between stakeholders. We have organized a number of high level gatherings, period of times, meetings and training sessions. The most recent of which took place at the University of Cambridge in February of this year,

to bring together a group of stakeholders to discuss the risks associated with AI and robotics and how we can leverage these technologies to the benefit of the people, countries and societies of the world.

We believe this work is essential. Societies have to be ready for the technological revolution. In this regard we must strive to avoid stifling innovation and work to contribute to the increase of the human well-being.

During a High-Level Event at the 71st session of the United Nations General Assembly in 2016 I announced that UNICRI was in the process of opening a center on Artificial Intelligence and robotics in Hague, The Netherlands to serve as an international resource on matters related to AI and robotics. The process of opening the center is nearly completed and for this we are very grateful to all of our supporters and stakeholders. We are particularly grateful to the Government of The Netherlands and the city of Hague for making this historic initiative a reality and enabling the creation of the global center to discuss advances in this cutting edge field.

The aim of the center is to enhance understanding of the risk benefit duality of AI and robotics through improved coordination, knowledge collection and dissemination and awareness. We want to leave no one behind. Our desired outcome is that stakeholders such as policymakers and Government officials will possess the knowledge and understanding of both risks and the benefits of such technology so that these leaders can contribute to the discussion on these risks and potential solutions in an appropriate and balanced manner. Such solutions may, for instance, take the form of improved policies. The knowledge and understanding conveyed to policymakers should be founded upon real and accurate information derived from the technical experts in the field. A series of activities have already been identified to facilitate the production of the outputs and to contribute to the objective of the center. For example, risk assessments. Training and mentoring programs and so on.

In conclusion with more than 50 years of experience UNICRI has structured its work programs so as to ensure it remains open and proactive to the needs of the international community. Now we are ready to take up the gauntlet with respect to AI and I truly hope that you will feel encouraged to share your expertise with us and exchange ideas with us because only together can we make this rise to meet this challenge.

Abraham Lincoln said you cannot escape the responsibility of tomorrow by evading it today. This conference I believe is the best way to start shaping our future and the present. Thank you.

>> NICHOLAS DAVIS: Thank you so much, Cindy. I have one I

think very important follow-up question which goes to all of your Cs. In the 50 years that UNICRI has been bringing together Intergovernmental agencies and partners from very different parts of the world to tackle incredibly important issues, what have -- what has UNICRI learned and what have you learned as director that can unlock this power of cooperation. What kind of one lesson or tip can you give us about bringing people together to make real progress?

>> CINDY SMITH: This is my favorite question. We think about this at UNICRI in a triangle. And if you think about academics over here and I'm a retired University professor, so pardon me if I say something mean about me. But academics speak to themselves usually and they talk to each other. And they don't relate very well with the other parts of the triangle. And over here we have practitioners. And many, many, many, many years ago I worked in corrections. And so I was a practitioner. And I know that we spoke a different language. So think about academics speaking Spanish and practitioners speaking French. They are not going to make the change in the Government. It is the people at the top and the policymakers that make the change and they speak Italian, for instance. They have words in common but they don't always mean the same thing. UNICRI sits in the center of the triangle. And we are the interpreters between the language. So when the corrections people say I don't understand why you want me to do this. And the academics say but I have research that demonstrates you will do better if you do it. The policymakers have to be able to say we are going to make a policy that works but they need both of the information from both of those legs in order to make good well informed policies. We will interpret and you guys sit at the table. Thank you.

>> NICHOLAS DAVIS: Thank you so much, Cindy.  
(Applause.)

>> NICHOLAS DAVIS: Now speaking of academics who speak multiple languages we have here Professor Stuart Russell. Stuart is someone who you may have seen speaking on these issues before. And I am glad this is not my first time of being up on stage with Stuart, but you are not just a leading light in AI research on the goals of AI and Artificial Intelligence and you obviously teach a variety of courses on this. But you also work on the applications of AI particularly in medical fields in a sense of computational biology and other areas well. You can bring us home with the mix of technical language and the idea of where we need to take AI from a technical and design sense and what impact that has on us as people who speak multiple different languages around the world.

>> STUART RUSSELL: Thank you very much. It is a great honor to be here. And now we can have my slides with -- so Lisa



Doll is one of the great Go players of the world. And here he is having what my Silicon Valley friends call a holy cow moment. He is realizing that AI is progressing faster than we expected. Coming down the pipe we are going to see, for example, technologies that can read, that can read text and understand it, extract information in a useful way. Once that happens, very soon after that machines will have read everything that the human race has ever written.

So now you have systems that can look further ahead in to the future, like Alpha Go looks further ahead on the Go board and they know more about the world than we did. And if they can combine those two capabilities then it seems inevitable that AI systems are going to be able to make better decisions than humans across a wide range of activities.

And there is a big upside and people have mentioned the upside. There are many things we can do. First we solve AI and then we use AI to solve everything else. So just to give one example from my group we made some development in AI. And we were able to use that to solve the problem of monitoring verification of the nuclear test ban treaty which is run out of Vienna.

And basically everything that the human race has that we value is the result of our intelligence. So if we can have more of it, then this will be a step change in our civilization. But at the same time there are these risks as other people have mentioned, Cindy, for example. And so we see headlines like this. So all of a sudden we find that AI has been manipulating the information environment that everyone lives in. So it is not -- this isn't our physical security. This is our mental security. And I actually believe that we have a right to truth. There is a human right to live in a roughly true information environment. Because humans are defenseless against information environments that are grossly corrupted. And I think this is a very important thing and it is messing up a lot of our political and social processes right now.

So what do we do about it? Well, Cindy talked about malware, the use of computers by criminals to do crime. That's a very serious problem. When we add AI it is going to get much worse. And the use of malware has been a catastrophe. The world has essentially completely failed to defend its populations against malware. So direct theft from bank accounts is now over 100 billion dollars a year.

And almost nothing is done about it. There is almost no traceability. There is no cross-border enforcement. So we have a really serious situation. We are about to lose the Internet. Large corporations are no longer using digital payments, for example, because it is too insecure.

So I think we need to work on malware and I think by solving malware we will set up a paradigm for how we can start to think about controlling the misuse of AI. But we have to do this soon.

Another area of negative use of AI is in killer robots or lethal autonomous numbers as the United Nations calls them and the issue these will become a weapon of mass destruction. Because the thing about autonomy is it decouples the number of weapons that you need to launch from the number of people that you need to launch. The CCW is moving forward with the Treaty. That's very good. I think we need more public awareness. People need to make movies about what this would be like for human populations and our professional societies need to agree that we are not going to do this. Building machines that can decide to kill humans is not okay.

Another risk that people talk about is employment. And this is a very serious question. Economists, many Nobel Prize winning economists think this is the biggest threat facing the world economy in the next 20 years.

So what can we do about it? One of the things that you hear is well, it is not going to happen. Robots are -- they are not smart enough and they are not going to replace human jobs. I don't think this is a wise strategy. You could do some careful economic analysis and this has not been happening. The thing we really need to do is have a destination. And we all talk about having transition plans to help our population through this big transition. You can't have a transition plan unless you have a destination. What is an economy that you would like to live in where most of what we call work now is done by machines.

And a hint, that economy does not have billion data scientists working in it. Okay? That's a complete nonstarter. That's a tiny lifeboat for a giant ship and that's not the future.

But this is what I actually want to talk about. So if we have AI systems that are more capable than human beings. One of the risks that people talked about it sets up a terminal conflict between the human race and machines. So people say it is not going to happen. People say don't worry about machines becoming more intelligent than humans. We can always switch them off. That's ridiculous. That's saying we can beat Alpha Go. You will find the AI community in denial about this question.

The AI community is now saying after 60 years of saying we are going to succeed, they are now suddenly saying we are not going to succeed. We have wasted our lives because AI is not going to happen. This is an unwise strategy. We have to anticipate and avoid the failure modes. We better be quite sure the purpose put in to the machine is the purpose that we desire. This is the primary failure mode. Intelligent machines will be incredibly good at carrying them out what we ask them to carry out. If we

ask them to carry out the wrong thing we will be sorry. And this was a point known to King Midas, the legendary king in Greece that asked for everything to be turned in to gold on everything he touched. We have a long history of failing to specify the purpose correctly.

So this is a technical problem and this is the main source of risk from AI. So there is a version of AI which I'm calling human compatible AI which is different from the way we have conceived of AI up to now. The first principle is that the only objective of the machine is to maximize the realization of human values but the second principle is important. The robot does not know what those are. It is precisely the single-minded pursuit of an objective that causes a problem. When the machine is uncertain of the objective but obliged to help the human achieve it you can switch it off and so on. It will ask questions. It will hesitate before doing something and ask if this is the right thing to do. And then the source of information about what humans want is what humans do. This is a long story about why that's actually right. So I'm not going to go in to that.

So my recommendation for what we have to do is actually change the very definition of AI and the way we pursue the research in AI. To make it provably beneficial. So this is an oxymoron. Beneficial is a vague notion. That we have provable, rock solid, cast iron mathematical guarantee that the systems we build will be beneficial. And the recipe is the recipe that I just gave you. We have made some progress on the basic technical principles. And we have actually shown systems that are provably beneficial in the sense I described. As this technical work progresses we need to roll it out through education, through standards, eventually not now but we will need regulation for the design templates of AI systems that are allowed to be deployed and connected to the Internet put in to self-driving cars and so on. It is premature to pursue it now.

The biggest problems that we face is the humans. Humans are incredibly complicated and evil in many cases and lazy, greedy, violent. And so actually reverse engineering all of that to figure out what it is we really want on our good days is a really difficult problem. And I think in the process of doing this we will actually learn a little bit more about how we should be, right? We'll not just be that way on our best days but maybe be that way more often. Thank you.

>> NICHOLAS DAVIS: Thank you very much, Stuart. As we bring this speaking section to a close and Anja comes up to bring the panel on, one question for you, Stuart, because we have seen an array of alignment here in terms of the importance of bringing the human beings in to the discussion, but we have also seen some tension emerge then between Manuela, desire to

bring mobility in to robots. And the concern you are expressing here before we do this we think to think provably and socially about the negative consequences of experimenting in this area. Is that a false tense we are seeing here? A way to resolve that we can have our cake and eat it, too?

>> STUART RUSSELL: It is a good question. Up to now we haven't had to worry because AI is way too stupid to do anything. In some of our other talks I talk about a domestic robot. There is nothing in the fridge and it cooks the cat. None of our robots could catch a cat, skin a cat and none of them know how to cook. We have been pursuing a definition and approach to AI that only works when AI is stupid. And when AI starts to become sufficiently capable to have a real impact then it is that definition stops working. And we have seen this already in computerized trading systems which have crashed the markets on several occasions and almost crashed the world economy in the process. And so once you get systems that can have an impact on a global scale, you have to think differently. I just want to say one thing, why is it human centric? So Nick has a story of AI gone because it makes paperclips. It turns the entire world in to paperclips. If you are ferroxidans which is a bacteria that eats iron you are thrilled with a world covered in paperclips. We are the ones who are supposed to benefit and not the bacteria eating paperclips. It is inevitable that the human will become the mathematical framework of AI.

>> NICHOLAS DAVIS: Thank you. In the ideal world we would ask the speakers to come and take this forward. We have a panel that Anja is going to moderate on the same topic. With that please can you put your hands together and thank the fantastic speakers. Thank you very much.

(Applause.)

>> ANJA KASPERSEN: While everyone is getting seated I would just ask for indulgence to go over time. We started 15 minutes late. I apologize for that. We are trending on Twitter. We are No. 1 on Twitter. Thanks to all the excellent Twitter feeds and updates from the participants here. Quickly I will come back to each one of these as we proceed. So we have here Chris Fabian with UNICEF. We have Lan Xue from the school of public policy and management at Tsinghua University in China. We have Peggy Hicks from the Office of Human Rights Commissioner. And we have Chaesub Lee who is the director of the ITU Standardization Bureau and part of the organizing committee that has brought all of us together. A big thank to Mr. Lee and his staff.

We have Sam Molyneux who is with the Zuckerberg Foundation. And at the end here we have Paul Bunje. With no further ado I will throw out a couple of questions. We have had a very rich discussion so far. And so this panel is not really meant to

bring up brand-new issues, though I hope there will be some holy cow moments, but to reflect what has been shared with us by the distinguished speakers, to give some framing discussions. I hope you can think of some interesting questions to pose from the audience when we come back to you.

So let me start with the application side of this. So Christopher, you work with the UNICEF innovation unit and the ventures team which invests in a lot of these technologies which you call frontier technologies. UNICEF being a UN organization, working similar to my organization in a very contested environment. What are some of the experiences that you have made applying these kind of frontier technologies in your situations? How do you deal with data integrity, the forensics of data being used in your technology? And what are some of the hurdles of how you overcome them? So Christopher.

>> CHRIS FABIAN: Hi guys. We're almost there. It is almost the end. Do me a favor before you answer your questions, raise both your hands for a second and just do this with your fingers. Doesn't that feel good? All the blood is going down. Thank you. All right.

Getting us through this last panel. Thank you to the partners for this event. You can find more on the UNICEF innovate Twitter account and my Twitter account. Four things I want to raise and 30 seconds each. We have set up the first venture capital fund in the UN. We realize that with emerging technologies we have to be able to make small, early stage investments in open source companies and pool their technology together and start to understand where the world is going. We have invested in drones and block chain. And I would like to talk more about that on Thursday morning. Don't tell the ITU guys.

The second thing -- I got like one laugh from that. There we go. Second thing, who knows what this is? Raise your hand. One person, two people, three people, four people. This is an MUAC. Mid upper armband. Monitors the nutrition status of a kid. If you have a healthy fat little baby your band is green. If you are moderately malnourished it is yellow. If you are a kid that's suffering from malnutrition your arm looks like this. This is a valuable tool for famine response. We are looking at investing in companies that are using the same thing using face recognition. This is a simple Artificial Intelligence but it is something that is important and affect the lives of hundreds of millions of children. How this works within the UN I ask you to Google a few things or use whatever search engines are. One is principles of innovations. These principles guide the basic ethical approach to new technology.

Second the UN innovation network is the body that ITU WFP, UNHCR. Guiding that network have been important to us in

bringing the community together. And third products like the magic box. To give their data in to a data collective and use them for outputs allows us to work with groups like the Global Pulse and others to understand the world better. And finally partnerships, like the partnership for AI for Good allows us to have that overall view of what's happening. So that's just a very quick set of things that we have seen over the last few years work as we understand and explore new technologies or don't understand them and happy to talk more about that in the question and answer.

>> ANJA KASPERSEN: Thank you for that. And we will come back to some of the dilemmas that you didn't touch on. Sam, I know you come from this data, more data scientist background. We talked and was mentioned by several of the speakers this open source as a platform for innovation. But although as Cindy was mentioning also as a tool to democratize destruction. If it is acquired or used by people with malicious intent. From a data consumer experience side do you think around open source and ways of moving forward to ensure safety and how we developed AI and related technologies that AI infuse?

>> SAM MOLYNEUX: Absolutely. So to answer your question I will tell you a little bit about what we are doing and then I'll get to that. So the first thing that I want to say is that at Meta we think that AI is part of the solution in partnership with humans. In areas or especially looking at areas and activities that are really at the root of how we make innovations that underpin human prosperity and human potential. And these are smaller markets from a revenue perspective but often have an outsized impact on humanity. And science and education are two of these. And in fact, in many ways underpins solutions to the 23 SDGs that we are talking about.

So within the sciences the speed of science in many ways is limited by two main problems which we think that AI is well positioned to help with. So the first is basically predicated on this notion that science is globally distributed and a decentralized activity amongst 10 million people in several hundred thousand labs. And it is hard to understand what everyone is working on. Scientific awareness is a key challenge.

And the second one it is extremely difficult to comprehend everything that is now. Even within subfields and to reason across that and understand it is more than one human can do or a human brain can do. And so in Meta with Chan Zuckerberg we have been thinking about how AI can be used to solve or the use of AI to power scientific knowledge, production itself. And we are in the process of making available a global scientific awareness system that can accelerate advances in sciences and make it available for free, open to scientists, researchers, funders,

policymakers, et cetera.

And one of the reasons why we think is this is an important thing to do scientific advances and discoveries themselves are recursive. And science compounds itself and this is a nonlinear thing which is important. Building on themes that we heard earlier this morning. A great example of this is the discovery of crispr which is a method that's unlocking a variety of new stem cells and the cycle continues and accelerates that we have seen for a century. From the perspective of free and open source we have been working with the -- operate on all scientific content, including closed access content. There is a tremendous way to do that based on this idea that it is important for everyone to know what's happening in the sciences. And our tool as all tools that operate in this way is available for free and should be made available for free to individuals, to institutions, to teams and to enable everyone. So that's basically how we think.

>> ANJA KASPERSEN: Thank you for that. So the last person I will ask more on the application is Paul Bunje, one of the co-organizers of this event. You come from a very strong background of global conservation and environment. We are at a pretty critical juncture in terms of future commitment to environment. How do you see AI being used in this context?

>> PAUL BUNJE: Thank you very much. It is a great question. I also greatly appreciate the pronunciation of my name. As an American we go coarsely with Bunge. There is a number of great examples even on the wheel behind of ways that AI can apply to a number of Sustainable Development Goals and those in the environmental space in particular. It was mentioned earlier that right now the UN is also hosting a conference on oceans. Simple machine learning comes from a program called Global Fishing Watch.

One of the things they have been able to do is utilize the data in fishing vessels and apply machine learning algorithms to identify where those vessels are but what kind of activities they are engaged in. Those data have become so successful that there is even one case where an illegal fisherman was fishing in a protected area of Palu and was successfully prosecuted as a result of the data that came out of this. Even though they then offloaded those fish to another vessel out, another ship outside of the area. It is a very simple thing to have been applied but it is transformative. Because efforts very often conservation or the environment are one example but this is broadly true of a number of Sustainable Development Goals where the resources to fight the problem are often quite lacking. We might be using these very linear approaches of catching someone and then enforcing the law in some ways if the law does exist. But those

don't scale very well. And though the malicious actors have the ability to circumvent a lot of these efforts quite broadly.

In wildlife trafficking, increasingly on the ground where you can tag animals, for example, or utilize microphones to identify where it is that illegal logging activity might be happening or poaching activities based on shotguns fired those are starting to be utilized to predict now where the illegal activity might lead. And then particularly excited by some folks that are looking at using that in a Black Market transaction, a financial transaction, very often illegal wildlife trade, using the same routes of laundering money as to human trafficking or narcotics and others. There is a lot of examples where tools used perhaps in larger market driven senses are being able to apply to things like conservation.

>> ANJA KASPERSEN: Thank you for that. It relates to Cindy's remark around the crime. If you watch X Men, the ultimate peacekeeper. You do the same thing, we are trying to preserve our planet. I am not sure if we will be sticking around or the optimization of goals the only way to preserve is get rid of the rest of us.

Going to the bit of the governance kind of right perspectives, so Peggy Hicks of the Office of Human Rights Commissioner, and also specializes in Human Rights. In this hyperconnected world around dealing with rights and data and who owns the data and forensics of data that goes in to particular cognitive technologies do we need a digital Avatar to help us protect ourselves or to help you protect us or to have a new set of rules?

>> PEGGY HICKS: Thanks very much and happy to be here. I have heard a number of things that then answer that question. I want to harken back to a comment that Robert made. One of the big challenges here is navigating a course between the risk of doing and the risk of not doing. And for me when you ask about a digital Avatar how we bring Human Rights or ethical values in it is about how we figure out those challenges because we know we need technology.

AI can be a huge resource as we heard in combating the world's problems and brings with it a host of challenges. I would like to make three points on that. Human Rights is really a useful tool in helping us navigate that space. We are not free floating in a world without any values or principles that have already been defined. We have worked on this project for many years. And we have ideas. And the universal declaration of Human Rights which was referred to earlier founds us and grounds us in principles like dignity and nondiscrimination that can be the watch words.

We also have really extensive experience that I think is



uniquely transferrable to some of the challenges we will face in looking at the risks of doing with Artificial Intelligence. And so Robert also spoke about soft law and how we develop it. So we have experience in things like the guiding principles in -- we are not sure what the space is bringing in the existing system of Human Rights and developing it in to standards and hard law that can help us navigate a more difficult space. We experience in both normative development and looking at how we monitor these spaces.

One of the things we heard from Brad Smith at Microsoft is there is a need for an international or independent place that will monitor and look at some of these questions. We can't leave it all to business or to Government. We need to find a separate space to do it. And I do think some of the analogies with the Human Rights sphere can be helpful tools. I wasn't to focus on, thought about as mind is the gap. In the British express, it is not a gap. It is a casism, which is the focus today. And so when we talk about applications and what we really need I think one of the main things I would like all of you to be thinking about in the next two days how do we close those gaps and bring those two spear spheres together in an effective way. Simple things like educating more people in both fields is a key step in doing. And then finally one of the principles of Human Rights that have to be brought in to this conversation is participation. And we know that the challenges of that in a field that is so complex that it is difficult to bring people along to even understand what we are doing much less support it. And Robert talked about the need for demonstration projects. It is more than just building trust. People fear what they don't understand. And getting people to understand this is going to be very, very hard. So we need to find ways to bring Civil Society and not just Civil Society focused in this particular space but a broader range of people in to the conversation.

>> ANJA KASPERSEN: Thank you. Important points. So Mr. Lee, ITU was established in 1860 to govern the electric traffic between states, Internet, mobile technologies. Where do we go from here? And I just want to refer to a recent Pew research report that was describing a situation looking at the U.S. audience but looking more globally where people are more confident in industries protecting us from some of these risks that Digital Age and new technologies bring on than Governments. And this can be a Government, organization and we need this kind of multi-stakeholder Government rooted as Peggy was saying both standards but also hard law. Where do we go from here.

>> CHAESUB LEE: Thank you very much. On behalf of my staff I welcome you. I lost my chance because of another event just across the street. We had IoT week is another big event. IoT is

one of our events with the support of this AI. Let me challenge this, technology development is quite a control. Let me say this AI, AI is nothing new. AI is quite old tough. AI is coming from many individual parts. We call it maybe in our language is protocols. Protocols, they engage this development of AI. Development of this AI we have certain common elements like something data. So we have a situation of this vertical development and some of the common areas called horizon elements. This is a situation. It is the same as what we developed now UN systems. We develop our industries, verticals but we have more and more commonalities called the name of this convergence, whatever use of these ICTs. That's very good benefits but certain some challenges for us.

So as a UN agency of ITU we -- we look at this very carefully and we recognize these ICTs really common for that. And when we recognize the ICT also innovated by AI but okay, that's -- true innovator of this ICT we have to understand what is the impact of AI. So this -- right of the way. Listening of these sessions I'm thinking about this, we may single out futures how we can go forward. And try to analyze these front classifications, like associate issues, like privacy issues or Human Rights.

Another limit might be business perspectives. We need some business -- we have to facilitate this business, creation of this new business, new markets and final element might be technical elements. What kinds of technologies always impact this outline. Maybe some different ways how we can move forward and how we can collaborate. So this is -- I believe this is a challenging subject for us of this AI Summit because ITU is a part of this ICT. But other UN agencies, other stakeholders you have your own views. For example, we may think about this, the concepts of this essence by design, human centered. Someone highlighted human centered approaches. And we need to talk about associate subject. We tried many times by design concepts for announcing security and privacy, but unfortunately we are a little bit late. The security is already gone. The privacy is quite a difficult to say force of this by design, but now it is a good challenge again, ethics by design. Why not? Because this AI just start. This event is very important.

All stakeholders get together now. So it is good time to strengthen this ethics by design. Human centered approaches of this business. Some of -- business and technologies we have to think of how we can strengthen these verticals like, continue of this business development. But also how we can show of this horizon elements like data sharing. Data we know is AI should be produced result. We are based on input data. So if data is -- definitely -- so we have to have something, to make clean of this data to ensuring the data while we are protecting the

privacy as well. This kind of concept will be necessary I believe. And also we have to strengthen this many peoples like Cs, collaboration, coordination, cooperations with many stakeholders.

So we are ITU is a part of multi-stakeholder models. We are delighted to work with you and looking to the future. And we are expecting these next two days we can find some ways, what is directions. How we can do. If we can find some ways it might be very desirable event and very productive event.

>> ANJA KASPERSEN: Thank you, Mr. Lee. So a key role played by ITU here. So continuing on the design thinking which and design kind of relating to governance Professor Lan Xue worked in many different processes trying to bring to the forefront aligning signs with design and with governance and actually your Ph.D. was on engineering and public policy which is, you know, you are a silo buster on stage which is great. I would love to hear your views on this. How do we bring the design in to governance?

>> LAN XUE: I think a lot of colleagues have talked about various good wishes of how to harness the AI technology for the good of human beings and how to minimize the risks, but I think there are three major challenges that we have to overcome. The first one is what we call the regime complex issue. AI is so complicated that there are many so-called regimes that's related to AI. We have ITU and also many other international organizations that are related to many aspects. They are not hierarchical. So they each have some bearings to the AI, to the advantage, to the risks but they cannot actually dominate -- one regime cannot dominate the other.

How do you coordinate with these regimes? That's the first challenge. The second issue is what I call the cycle gap because of the complexity of these regimes. In order to find in a policy to guide the development of the technology the process is going to take much longer time than the technologies cycle. So in other words, technology moves much faster than the governance cycle. So I think that gap actually will not make, you know -- so actually the gap of the two make it very difficult to have an effective governance system. And the third outcome is the first two is fragmentation of efforts.

So you have a lot of people, a lot of institutions trying to work out different things but at the same time it is not going to make, you know, a huge impact on the major trend of the technology.

So how do you address those two -- those three obstacles? I think the first one I think that we can think about is the -- I think we can learn from the climate change. Another sort of major global challenge we are facing. I think the IPCC mechanism

is one that we can learn from. In other words, creating sort of a knowledge, a global knowledge base doing kind of analysis, comprehensive analysis so that after providing a knowledge platform for policy for action.

So I think, you know, so one proposal I would make here, there is probably a need for global knowledge alliance for AI that at least we have a clear sense of where the technology is. What is the potential impact. Both positive and negative. So that actually, you know, that would serve as a guidance for the industry and also for the Government.

And the second thing that we can also think about is to think about sort of taking the adaptive governance approach. I think when we designed the policy we always wanted to be comprehensive and wanted to be, you know, covering everything and also being very effective, but at the same time I mean given the complexity and given the uncertainty related to the technology I think actually we probably need to think about adaptive governance which has been used in many other areas to allow some ambiguity to allow some sort of uncertainty in designing where it is policy and guiding the development of the technology. So that actually designing a learning process, so that in a way you can move from one step to the other and then learning and adjusting the policy that we can move from there.

>> ANJA KASPERSEN: Very profound ideas. I am going to open up to the audience. Please introduce yourself and keep your questions succinct or comments to the point. I will start over here.

>> Thank you so much. My name is Dr. David Mebnu. I am working with a startup named Alfred in psychiatry machine learning. Two questions. One is a serious one and one is a thought provoking one. We heard a lot today about the potential for autonomous weapons system and the other way that AI could end up directly killing people. I would like to ask something about already killing people is targeted advertising through AI. We heard about the challenges of chronic disease. Chronic disease is often informed by lifestyle choices that are also often exacerbated by smoking, alcohol, poor food choices, much of which is driven by advertising in which now has this added weapon of Artificial Intelligence.

So are we going to be answering to a place where AI is used by companies is going to be in competition with AI used by social good organizations like the UN one doing education, one doing advertising? And how do you look at corporations and have them resolve that conflict while also having an incentive to profit from targeted advertising? For example, selling alcohol to somebody you can tell from the press and social media.

Second question we have heard a lot about human centeredness.

But what about AI rights when they get complex enough to deserve them, if they do?

>> ANJA KASPERSEN: Over here there was a lady in the back. Yes.

>> Yep. Thank you very much. My name is Kristy Nann. I am the Chair of the European group. My -- thank you very much for these talks. Artificial Intelligence can do a lot about understanding problems and analyzing them, doing diagnosis of every kind. But is there a power gap, not only a cycle gap but a power gap in solving the problems? So what's about diagnosing a disease but not being able to treat them? That's my question to the panel. So Artificial Intelligence seems to be very powerful but not in solving the problems so far. Sorry.

>> ANJA KASPERSEN: (Off microphone).

>> Yes. Hello. My name is Altan from the school of hygiene in the UK. And I am working on chronic noncommunicable disease analytics, but my question related to my previous life where I was an advisor to the former Prime Minister David Cameron and I have a big question. And it is unfair to ask the panel at the end of the first day, but I wonder if we can land the question and it can be tracked during the course of the next two days. How does this dialogue fit in to current experiences around global trade, patent, and commercial legal practices? And what I am trying to bring out here if we are saying there is clear and very specific social good related to this technology, which, in fact, is a commodity, consists of products and service as continue highly likely it is going to fall in to our frameworks for trade. If there is a strong social value being place around this agenda.

How do you keep this social commitment and do it in a way that's effective in terms of trade processes? Think about frontier markets. Frontier markets are places where people are looking to move in to. And also Developing Countries and, of course, we know from our experiences that that's also about shaping markets. And I think the question is essentially that, how do we dialogue with global tariffs and trades around this. Should AI be, in fact, a special economic zone which is free of trade and tariffs and taxes? That's my question. And could this question be put to the WTO and other multilateral institutions with that responsibility? Thank you.

>> ANJA KASPERSEN: As you said a big question and it can simmer and mature over the next two days. First a gentleman all the way in the back and then you.

>> Thank you. Richard Hill, Association for Property Internet Governance. I'm a Civil Society activist, although I wear camouflage. And I was very pleased because at this conference just in the first day I have heard every single theme

that I think is important, both in general and specifically. But I wanted to drill down on one specific, which Ms. Hicks, you mentioned one of the initiatives from Microsoft in terms of cyber attacks, but perhaps you could comment on another which is the call for Geneva Digital Convention whereby states would voluntarily agree to refrain from certain activities. Thank you.

>> I am Ragesh from Germany. I'm first of all very much thankful and I am very much pleased to attend this conference. And I only think that this conference will grow bigger and bigger in the coming years and centuries. I would like to say that -- quote Albert Einstein, he said technology is not helping us solve the problems in the world because we have not solved the problem of international relations. So he meant the United Nations should be further empowered. And we should come out of nationalism.

So from World War I we move from the League of Nations and World War II we moved to United Nations. I wish we don't have any further catastrophic event for World Government. One common plan. My question is how are we going to make sure all the regulations are going to be followed towards the world because we have so much partnership on AI, but who is going to and how are we going to make sure that these regulations are followed throughout the world? And I potentially feel empowering United Nations is the only way. And further such discussions should be made in the future.

>> ANJA KASPERSEN: Thank you. You should go to Yoshua's panel on AI for prosperity tomorrow which we want interesting discussion around this. And then was two more questions and then I'm closing for now to leave it back to the panel. Sorry, it was -- yes, you sir and then here. Yes.

>> We spoke about exponential growth and this links very well with AI. But exponential growth means that whoever comes first using the AI can get such an incredible success that it destroys everyone else. If I start selling ice cream and I have a very powerful AI system to advertise my ice cream, I can become the best ice cream seller in the world. But any other guy selling ice cream will be out of business. I'm concerned about the people being left behind by the exponential effect of AI and the fact that someone will be incredibly successful. Of course, is the lead story that will come to the newspaper. But all the millions that be will left behind probably should be remembered, too, by us. My name is Masimo. I am an intellectual property lawyer from Italy.

>> ANJA KASPERSEN: Thank you. Last question here.

>> So yeah. Thomas. I am from Fraunhofer from Berlin and from the Technical University in Berlin. So my question is how do you deal with the fact that a human is a human and a machine

is a machine. That's obvious. But if you think about it, a machine at its core can never be guaranteed to respect human values because it doesn't have the neurological necessities, right? So everything can be hacked. And, of course, humans turn on each other but that percentage is small and that's why society is stable. So if you think about it, then one step further when some of the great talks we heard before when AI becomes smart enough to matter, don't we then need a security check on all AI systems all the time? So when you enter an airplane and you go through security because no human can be trusted for some reason and no AI system can be trusted. There needs to be a security check for self-driving cars. And we also need to address another fact which is the fact that networks and a malicious AI system can have once it is networked and maliciously hacked it could, you know, have a mass destructive effect although it is designed a completely different way. These are two questions. And I am sure there will be lots of other questions and which organization will be regulating AI systems in the future and how do the panelists see that.

>> ANJA KASPERSEN: Final question. So given the time we have left what I will do is I will leave to the panel, you have all the questions and let me just quickly summarize the questions, around targeted advertising and the danger of it. It was about AI power gaps and solving versus treating. It was dealing with diseases and, you know, thinking around practices and social good. And it was AI as a commodity. AI is in a trade relation and data is already part of the trade discussions, will the refined version when it turns in to algorithm or cognitive systems be part of that equation as well. And then it was the comment on the call for a digital Geneva Convention which is a thing, attribution of cyberspace and how do we make the discussion truly mobile and make it truly global. And what is the right platform. And it has to be the UN which was the suggestion. And then the last question if I get this correct now was -- second-to-last question was exponential growth and ensure we are all on the same playing field and have the cognitive augmented capabilities to not put all other ones aside. And the last one to be more of the adversarial question and how do we vet and validate and verify how the AI has been coded and come to determine our lives or decision making process. Those are pretty profound questions that spans over a wide area. I will just give the word for -- to each one of you in reverse order. And would ask you also to think about the how and think about what you want to leave all the participants with in terms of thinking about how to prepare for tomorrow's workshops. We will dive in to some of these questions. So we will start with you.

>> LAN XUE: I think there is several issues related to the

governing of potential risks, technology to go through the market and being widely used. And then we address the risks and the problems at the end cycle. I think with AI technology I think what all people raised in terms of concern we need to start at the beginning. We need to think about what system are we going to adopt to address the risk issue. And so that actually we can really embed the system and the principles in to the designing and the -- of the system, in to the application. I think that's probably the principle I would follow.

>> ANJA KASPERSEN: Mr. Lee.

>> CHAESUB LEE: Yeah. Those questions are really profound questions and not easy to answer. One thing is we assume AI implemented, completed just tomorrow. That's not going to happen. AI is -- continuously growing up. And now it looks very close to us. But it is still far away to complete. I don't know when it will complete but AI is continuously grown up. Even our society is grown up. Since we use this mobile phone, how many years now we use this Smartphone? Around 15 years, 10 years.

>> ANJA KASPERSEN: Nine.

>> CHAESUB LEE: We have changed many parts but still we have the change in process. So we can raise these questions for the target goals, target goal times, raise the question in this today. Might be something to struggle ourselves. Rather than I want to say this as a step by step, like someone talking of this exponential growth. We may think of that -- that possibility. But we believe, I believe it is old -- during that development of this AI process, AI enabled ice cream, AI enabled ice cream controls, our society, our business, our communities they develop their own ways. So will be relatively shorter gaps. But that question would be very important for us because we now want more fair, human centric means. We look at these fields. There is 17 goals of this. We are all human beings. Now want to realize, implement these 17 goals, which means that ensure this business about profit. That's also important but it is time to think about how to meet this SDG.

Business insurance is very important. I don't know regulator of this but it is time to get different way of balance. So winner take all, we can continue these ways or we can think about how we can show the benefits. This is another element. This is a good time to think about how we can really share these benefits to make people have a more sharing of these human beings, well-beings. How we can make it happen. This is the important part. And in this regard some of the regulatory aspects, so can control this, I believe each verticals, especially our colleagues from the UN organizations, each vertical has some -- UN organizations are responsible for each vertical areas. And we have certain responsibility to address



that.

Some of our groups, I believe ITU in terms of this ICT, we have certain roles of these horizon areas. We have our responsibility to address this horizontal part. If we can have some good agreement of such directions, we have more deep investigations how we can practically implement these areas. This is my view.

>> ANJA KASPERSEN: Thank you. So Peggy.

>> PEGGY HICKS: Thanks. I will take up the question on Geneva Digital Convention. The UN Human Rights Office doesn't have a position. Important things, too, I think the analogy that Brad Smith has brought to this is that this is like a weapon in a sense. And we have to deal with it as seriously as we do deal with other weapons. Is it really an important way to garner attention to and to think about how we take up this challenge? It is also important if you believe there needs to be a Convention to realize that we can't wait for a Convention. Get more immediate action because, you know, this isn't a process that can wait five, ten years before we can have a full-fledged Convention. That's why this question of a monitoring facility, inform more legal or normative process is an interesting starting point for the conversation.

Could I also just go back to the question that started us off about advertising and social good versus those that are using corporate, promoting corporate goals in a different way? I think it is a really important question that sort of should underline some of the conversations in the next few days. The reality is we are talking about how AI can contribute to discrimination or inequality globally, but we already live in a very unequal world. So we are starting not from a level platform or we are starting with huge deficits as I know that Salil Shetty talked about earlier today. And we are to figure how can we not only invest in the side of trying to make sure we don't make it worse but using technology to improve and address those gaps. And to me that comes down to a fundamental question of investment. We will be as societies willing to see this chance to meet it and to invest in using Artificial Intelligence and other technologies for social good and initiatives like that one just described by the Zuckerberg Foundation are a critical starting point, but it is going to have to be much broader and much deeper.

>> ANJA KASPERSEN: All those are very vulnerable contacts. Paul.

>> PAUL BUNJE: Thanks. I think I'm going to build a bit on what folks are saying and I picked up on both those questions throughout today on a couple of themes that are relevant and how we codify those values in a program for developing AI and AI for

good in particular. And the other being is exponentials. Something like targeted advertising or the use of Artificial Intelligence to more effectively make our weapons kill people may also be considered human values. Not all values that humans maintain are positive in nature. The point about exponentials and this notion of being able to digitize puts the potential in it a small group of individuals, the power that one took large linear institutions. You are decoupled with technologies like AI. Be they ones that we all might share or ones that vary widely. So what does that mean? Where do we go when you asked how do we deal with this? Let's pick up on this notion of Artificial Intelligence, like any technology can be used for good or for ill. And we often look to specific institutions to represent some sort of codified social norms or social contract and then regulate and to place what we want to see happen.

But I think also picking up on the idea of how institutions work and the theme of this panel, collaboration, we need to recognize that the world is changing so dramatically that things like states are undermined by the very technologies we are promoting. Have the ability to utilize these things across state boundaries and implement whatever use they might value notionally at that given time.

So we need to think about how it is not we just use our institutions to regulate but we have an expression at XPRIZE that you get what you incentivize. We need to start thinking about the other tools, carrots in the mix that will allow us to demonstrate what are more likely to be shared human values and incentives. And I think that's where this -- two sort of themes that are coming out of this. One why leave -- we exist right now at a very opportune time. We heard about the growth, exponential growth and heard about some similar shared principles or at least notional general ideas that AI can be used for good but also heard that it is not there yet. AI is still not very well developed but puts in a good position to think about this roadmap for the future.

And you heard Mr. Katsumi Emura talking about as well as Professor Russell how we think about a vision of the future and building out a meaning plan to achieve it. And over the next couple of days we have the opportunity to put some of those very beginning steps on to a roadmap that we can execute against. That's a rare thing in human history to have that ability of foresight and ability to plan. And we heard about this need for collaboration. And there are a lot more conversations like this to be had. Who the participants are, what the sort of rules of the road that might look like. There is some notions here, but it is not going to be easy to come up with I think the sort of firm next steps that are happening. It is only going to happen

through further dialogue.

>> ANJA KASPERSEN: And AI transcends and Europe borders permeates. Everything we do, our mind set and responses need to correlate. What you --

>> PAUL BUNJE: We need to build our institutions and our society to recognize that very fact.

>> ANJA KASPERSEN: It goes to your point who is then the right place to do so. So this is a starting point also to discuss that. Sam.

>> SAM MOLYNEUX: Thanks so much. So I think what I would say cutting across a couple of the questions is that there is a big new tool available and Governments, not-for-profits, philanthropies are in sort of an unprecedented position to turn their dollars in to social good. Cases where you can basically find a problem that's an information challenge or translated in to an information challenge where to drive for it an AI project and then make that available to everyone through the Web. And I think, you know, that could be fighting dangerous advertising or it could be new health challenges. And I think this is a really big deal.

And then the second thing that I would say is related to the question over here about sort of the asymmetrical advantage for first movers with AI in a market. A lot of new advances that come from machine learning models that are reduced to practice become commoditized and a lot of the advantage comes through unique datasets to train.

>> ANJA KASPERSEN: Thank you. I am waiting for your next stint of mobilizing the audience as well before the drinks. So you can close off here.

>> CHRIS FABIAN: Never give away the Moderator mic. Thank you. So I started sort of being funny but I think we -- I would like to end on something quite serious. And hopefully this is something that underscores the discussions in the next two days. It is easy to think about all the great stuff we have heard today and all the technology that we have seen, and think about the wonderful future that awaits many of us. And that's true for many of us. If we are wealthy and if we live in urban areas, we are fine. But the world is entering a time of great darkness. And if we are thinking about the next 30 years, if we listen to what Jack Ma said, the next 30 years holds more sorrow and pain for humanity than they do joy.

Answer these seven questions with new numbers and facts and we will end this. There are 55 million kids who are on the move today because of war and violence. 55 million is a number that I cannot make anybody understand. These are kids who don't have a state or sovereign Government or institution to fall back on. There are 500 million kids who are at risk of not getting the

right nutrition because they do not have access to the right food because they don't have water, not enough anymore or they have too much water and these are because of climate change. If you are a kid who is moderately or acutely malnourished, your brain does not develop. And that creates an asymmetric world.

Three, I talked recently to the head of a large airline company and he shut down his pilot training academy. They are not training human pilots anymore. Why? They have enough humans to last until they can do the piloting. They felt most acutely in poor places, so you have a country like Malawi where you got 18 million people and of them 50% of them are kids. 10 million kids. In the next 15 or 20 years the population grows. It will grow to 40 million, and that's 25 million under 20 years old. And there are not enough jobs. And how do you answer them? There is very little that we can do.

We heard about a post-sovereign world and there are more people on Facebook than people in China. Those services that platforms are providing sovereign services, cash transfer, education, identity, and we as the UN do not know how to negotiate with them, do not know how to work with them and how to set policy with them.

It is a new world and a very scary world. We as UNICEF have a mandate and a mission to advocate for the rights of most vulnerable children. And the 55 million kids who are on the move don't have Government that we can relate to. So I think that as we look to the next five years or ten years and I challenge everyone to think about this for the next few days. We can think of AI solving problems. How do we solve the problems for all of us? With Facebook and Google and IBM and these other big tech partners who want to make money but also want to do good. How do we create a world of fairness and understand our position? These are the sort of questions that I think I would actually raise as secondary questions to some of the ones that we heard. And I very much look forward to discussing them with all of you in the next two days. Thank you.

(Applause.)

>> ANJA KASPERSEN: Thank you so much for those final points. Just to add my organization also works with people on the move in a very vulnerable context. Other forms of serious violence or crime. And we in addition to what you are saying about the people suffering there is also the reality of that an increasing number of these are connected and they generate data, and how do we deal with the protection of data when you are at your most vulnerable. Every single one of these kids growing and they are going to apply for the next bank loan and data may have followed them for 15 years. So there is a lot of issues that come out of this increasing number of vulnerabilities, different

levels of literacy, understanding and also being able to say no, you can't have my data.

You are standing there in a refugee camp and you need to give away your data to get the food you need. Certainly to what will eventually feed an AI in these contexts.

So I want to thank the speakers, the panelists, and Nick had to leave to catch a plane, the organizers for putting this on. And just kind of again encourage everyone to when you go back home, you sleep, like really let this discussion today simmer and come back prepared tomorrow. And let's really start working on the how.

We have two more days to figure out the how and who governs and how adapted we can be. I'm really looking forward to that.

Just an FYI, housekeeping things, outside is going to be a reception and I look forward to engaging with people there. And there is going to be drinks and some refreshments and other things, right? Fred?

>> (Off microphone). Drinks. The reception will be at the cafeteria in Montbrilliant. So you have to go up two levels of stairs and down a narrow hallway.

>> ANJA KASPERSEN: Thank you so much, everyone. And sorry for keeping you late.

(Applause.)

(Session concluded at 6:27 p.m. CET)

\*\*\*

This is being provided in rough-draft format. Communication Access Realtime Translation (CART) is provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.

\*\*\*