

Remarks of Fabrizio Hochschild
Artificial Intelligence: Technology to serve humankind, setting legal standards,
Republic of Slovenia, Council of Europe and UNESCO

New York, 28 October 2019

First, I would like to congratulate Slovenia on hosting this event. The topic - developing AI that can serve humankind, is timely and of critical importance. How we manage this is as important as how the world came together in past decades to steer nuclear or bio technology to good uses and restrict malicious use or unintended consequences.

In my remarks today I will briefly address the speed at which the AI revolution is happening, and the benefits it brings

I will then highlight some of the risks that AI technologies entail around:

- Surveillance
- Agency and Accountability
- Predatory beta testing of AI technologies

I will then describe how risks and benefits may be unequally distributed, both geographically and across our societies.

Finally, I will then conclude with possible solutions from the High-level Panel on Digital Cooperation

The AI Revolution – according to a landmark McKinsey study - is happening “ten times faster and at 300 times the scale” of the Industrial Revolution. Research and development in Artificial Intelligence - the capability of a machine to imitate intelligent human behaviour, is experiencing exponential growth. AI delivers powerful tools to convert massive amounts of raw data into actionable knowledge to address complex challenges.

With the amount of data produced today AI is a game changer and a necessity as much as simply a new technological advance.

AI holds the potential for new concepts, higher accuracy, effectiveness, and lower cost in applications ranging from energy efficiency and personalized medicine, to transportation and financial services. AI is already transforming how we communicate, consume news, drive our cars, carry out commercial transactions, and in some cases, are educated.

The very power of AI applications carry a series of risks.

One is an enhanced capacity for Surveillance and curtailment of the right to privacy

It has been said that authoritarianism is in AI’s DNA. AI backed surveillance systems, facial recognition, digital identity, or social media analytics can be used for legitimate law enforcement activities or, the same tools can be used to track political dissent follow human rights activists. The same AI applications can be abused by non state actors to manipulate consumer or voting behavior.

The collection, sharing and use of data for AI applications such as facial recognition or digital identity systems raise challenges related to adequate data privacy law, application of human rights / anti-discrimination law, as well as interpretability, security and oversight.

Legal and ethical frameworks around AI will differ significantly between different governmental regimes, balancing the trade-off between economic / social stability and population autonomy and individual privacy.

A second set of risks relate to Accountability and Agency

The use of AI for predictive policing and court sentencing highlights the benefits of resource efficiency, and at the same time when decision-making is delegated to machines, who is accountable ? Autonomous vehicles are less accident prone, and at the same time surface questions of accountability, liability and the blurring of machine and human agency. Here, the risks of the erosion of human moral decision making is also of concern.

A third set of risk relates to what is called Predatory beta-testing

Our ethics and legal discussions must also consider the ways in which AI technologies are being developed and tested. Predatory beta-testing and ethics dumping - where risks are exported to marginalised and vulnerable populations or to low and middle income countries, are increasingly of concern. There are many specific examples of poorer countries or marginalized communities being used to test AI applications.

The counter-point to 'ethics dumping' is what is often not done to protect people - or what might be called 'ethics shirking'. Consider one large digital platform's rapid response in the face of a law in a Western European country obliging it to limit hate speech on its platform - which has seen it recruit over a thousand moderators in that country alone - versus its slow response to violence fueled on its platform against against an ethnic minority in a developing, where only a few dozen moderators were only belatedly appointed.

This highlights my next point that of potential unequal distribution of risks and benefits of AI.

Critical reflection on how AI technologies can best serve humankind is a necessary exercise. It is not a given that the transformative benefits of AI technologies will be equitably distributed across all of humanity.

In fact, AI has yet to offer significant positive transformation to the 3 billion individuals living in poverty. If we do not do a better job of steering it the benefits of AI developments will likely be distributed along pre-existing dimensions of digital inequality.

Of course, there are examples of AI being used for development efforts - AI is being used for example to analyze satellite imagery for crop yield forecasts or famine prediction, and to deliver insight for effective humanitarian response. But this is not where the majority of AI investment in applications is occurring as the financial pay-off is not always obvious.

And then there are the side effects: AI will also play a huge role in the automation of labour, with resulting displacement and unemployment likely to be felt more acutely in emerging economies. On

the peace and security front, cyber attacks using AI may also hurt developing countries disproportionately, given limited infrastructure to safeguard against such attacks.

A 2017 Harvard Summit on AI and Inclusion identified a related problem: Actors influencing the regulatory discourse, remain firmly grounded in data mature and affluent geographies, despite the Global South arguably holding the most to gain and the most to lose from AI, according to the ITU.

The Global South is clearly underrepresented in this debate around legal standards and regulatory frameworks to maximise the benefits and mitigate the risks of AI deployment - with many areas of Africa, Asia, Latin America and the Caribbean underrepresented in terms of governments and policymakers engaged in this vital conversation.

We need to collectively and inclusively think of legal and ethical frameworks in which AI could fuel economic, political and social improvement in less privileged countries and communities, and mitigate the reinforcement of existing digital and economic inequalities.

There are many Initiatives that seek to promote a truly global approach. such as the IEEE technical standards organisation's Global Initiative on Ethics of Autonomous and Intelligent Systems which involved 'hundreds of participants over six continents' to inform their Ethically Aligned Design Report. In addition, the IEEE have set up a Classical Ethics Committee, to consider the cross cultural overlaps and differences of ethical issues surrounding AI systems.

UNESCO has also initiated a global discussion on the Ethics of AI.

These kinds of initiatives will be critical to formulate relevant legal standards that ensure the compliance of AI tools with human rights and fundamental freedoms, in ways that support economic and political stability.

In this same spirit of ensuring broad distribution of benefits of AI technologies, the SG High-level Panel on Digital Cooperation calls for AI engineering and ethical standards (such as explainability, accountability, and transparency) to be developed using multi-stakeholder and multilateral approaches. This is means representation across geographies and involving stakeholder expertise from industry to civil society to academia. The Report also calls for enhanced digital cooperation with multiple stakeholders to think through the design and application of standards and principles such as transparency and non-bias, in consideration of different social settings.

To conclude, the AI revolution holds many benefits for humankind, but also risks, ranging from enhanced surveillance, challenges for agency and accountability, to predatory beta-testing of technologies. These risks are not equally distributed, both geographically and across our societies, and in order to resolve this, the restricted representation in AI Policy setting needs to be addressed. In light of this, the High-level Panel hopes to bring together fragmented geographic and stakeholder groups.

I have no doubt that the deliberations today and **the leadership of Slovenia** will help to shed light on how to bring about a collective means to develop artificial intelligence technologies, and the legal standards necessary to ensure the benefits are indeed in service of humankind in our entirety.

Thank you