AI introduction/abstract

> *"One can imagine such technology (AI red) outsmarting financial markets, out-inventing human researchers, out-manipulating human leaders, and developing weapons we cannot even understand. Whereas the short-term impact of AI depends on who controls it, the long-term impact depends on whether it can be controlled at all."*

Stephen Hawkin

Today we cannot escape Artificial Intelligence (AI) in the media, in research, in policy and politics, when using Google, Facebook or Alexa, AI is there, and it touches all aspects of our lives. The attention for AI is expanding rapidly. AI is also more than a new trend or buzzword, it is here and it is here to stay, and due to its enormous impact on society, requires all concerned to critically asses the relationship between AI and our society and human kind.

Yet, AI is far from new. Already in the mid 50´s of the previous century, AI started to move from imagination and theory to implementation, with the improvement of computing power and storage. Although ever since, significant amounts of money have been invested in R&D, AI, despite its promise, has not been in the forefront of attention, and progress has been steady but slow.

The major barrier for real (practical) advancements in the application of AI has been its need for massive amounts of (digital) data, a commodity not readily at hand in the previous decades, as well as insufficient computing power.

Over the last two decades, the amount of digital data has exploded. It is increasing at an ever-higher speed: each second of the day massive amounts of new data are generated and stored. In the early 2000´s the term Big Data has been introduced to label this phenomenon. Developments of an omnipresent internet with billions of users (private ad professional) worldwide, highly digitized systems in industry and business, the internet of Things, but also hugely increased storage capacity against ever dropping prices, and major advances in computing power are some of the key factors driving this development, which really is only at the beginning.

In fact, we are only at the early stages of developing useful applications for so called narrow AI, aimed at performing a specific task (speech driven applications, self-driving cars, online personalised services and advertising through social and other media, intelligent robots, to name a few), and still very remote from general AI where most if not all cognitive functions could be performed by AI systems.

While most people agree on the potential benefits of large scale AI application, such as badly needed improvement of productivity, improvement of quality of products, services and processes while

reducing cost, improvement and support to everyday tasks and activities, also major concerns are voiced, not only by advocacy groups but also by leading scientists and technologists.

Elon Musk is one of the most outspoken technologists on the risks of AI: *"I think we should be very careful about artificial intelligence. If I were to guess like what our biggest existential threat is, it's probably that. So we need to be very careful with the artificial intelligence. Increasingly scientists think there should be some regulatory oversight maybe at the national and international level, just to make sure that we don't do something very foolish. With artificial intelligence we are summoning the demon. In all those stories where there's the guy with the pentagram and the holy water, it's like yeah, he's sure he can control the demon. Didn't work out."*

AI implies a paradigm shift from our traditional human centric perspective on society to a perspective in which AI equals or outperforms our human cognitive functions, which leads to existential concerns. The key questions raised by most focus on the preservation of human values. How can we safeguard our human values while deploying AI, now and in the future?

Other concerns typically deal with concerns on *who* applies AI and big data, and especially *how*. Biased or flawed data and algorithms will produce biased or flawed results, irrespective if this is intended or not. Next to genuine and informed concerns, there are also many myths about AI that blur the discussion.

The understanding of the potential of AI as well as the concerns are global. AI is an important, maybe even key development in the global competitiveness, and the balance between the different geopolitical and economic influence spheres.

As AI developments are still in their infancy, it is timely to discuss all aspects of an AI future, balancing a favourable environment for the advances and contributions and AI can and should make to our society, with the protection of basic human as well as social and economic values.

The new president of the European Commission Ursula Von der Leyen has promised to propose legislation to address the human and ethical implications of AI during her first 100 days in office. The challenge for Von der Leyen, according the European Commission itself, will be *developing legislation that reflects European norms and values while also avoiding overregulation that might hinder innovation and investment*, in other words finding the balance.

Some argue that there are too many unknowns about AI to be able to regulate, *you cannot regulate what you don't know.* While it is correct that there are many impacts and effects that today are entirely unknown, it would be unwise to regulate the technology itself, or not to start regulating; instead a framework that regulates on principles and values, could capture currently known and unknown effects of AI application, as well as providing a needed degree of certainty to those developing and deploying AI, as well as investors in AI.

The High-Level Expert Group on Artificial Intelligence has developed a framework for trustworthy AI n the EU, including

- Ethics Guidelines on Artificial Intelligence: The Guidelines put forward a human-centric approach on AI and list 7 key requirements that AI systems should meet in order to be trustworthy.
- Policy and Investment Recommendations: towards sustainability, growth and competitiveness, as well as inclusion – while empowering, benefiting and protecting human beings.

The needed regulation should reach further than ethical and core human values, to make AI work in the future. All regulation we know today is based on the paradigm of human as the actor, originator, decision maker etc. The paradigm shift from this human centric perspective should be reflected in all relevant regulation including domains as liability, intellectual property, competition, privacy to name a few. Only a comprehensive framework can secure a thriving and trustworthy AI deployment, in which the benefits can be reaped, and the human values protected. While it is the current ambitions of the new European Commission is highly commendable, it is highly ambitious and perhaps doubtful that such a complex and comprehensive framework can be achieved in hundred days.

Debates as in the Knowledge4Innovation AI & Big Data for Innovation Summit are aimed to contribute to a balanced result of the process.