

# Ethics in Artificial Intelligence

Sara Engardt sengardt@kth.se

As artificial intelligence technologies advance, more and more tasks and jobs are considered automatable. People have always tried to facilitate and automate processes, inventing everything from the steam engine to robots working in manufacturing. With the introduction of artificial intelligence, some tasks and occupation that for a long time have been considered as un-automatable seem to be done fine (or even better) by an AI.

Take Watson for example, an AI developed by IBM [1]. After beating the champions in the TV quiz game ‘Jeopardy!’, the manufacturers saw the possibility to apply the AI to medical purposes. Using a variety of AI-technologies, Watson is able to help, and even beat, human doctors in diagnosing patients. A doctor is maybe not the first occupation that comes to mind when thinking about jobs that can be replaced by computers.

As more and more professions are being automated, there is a lesser need for human labour. This potentially causes huge ethical problems around the entire AI research and industry. What will these people do instead? How will these people earn a living? How shall governments impose taxes on work?

The authors of the Stanford report “Artificial intelligence and life in 2030”, says the following [2]: “As labour becomes a less important factor in production as compared to owning intellectual capital, a majority of citizens may find the value of their labour insufficient to pay for a socially acceptable standard of living.”

One way of solving this could of course be to re-educate the part of the work force whose skills are replaceable. But according to the authors this will not be enough. They argue that we will need to expand the current social welfare, and they even propose a guaranteed basic income for all citizens.

Personally I believe that a guaranteed basic income will be necessary. According to Frey and Osborne [3], 47 % of all US employments are at risk. An adaption to Swedish conditions [4] arrive at almost the same conclusion, 50% of all jobs will be automated by 2035.

I see basic income as a positive thing. Instead of people doing jobs they dislike, or even jobs that can be dangerous, they will be able to focus on doing what they are actually interested in and that can benefit the entire society. The question is of course, who shall pay for it?

If 50% of all jobs are automated by 2035, this obviously means a huge loss of income tax for the state. It also means that big corporations and companies will be able to majorly reduce their costs for staff. The way I see it, governments and the AI industry itself are responsible for making sure that these profits do not only benefit a small elite, but the entire society.

The increase in efficiency and the decrease of human errors could increase the world's wealth significantly. If done properly, this wealth could be distributed to the entire population of the world. The possibilities of AI are endless, and as long as decision makers are responsible and have everyone's best in mind we can use it for everyone's good.

## References

- [1] <https://www.ibm.com/watson/health/>
- [2] Stone et al.: "Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016. Doc: <http://ai100.stanford.edu/2016-report>.
- [3] Carl B. Frey and Michael A. Osborne,: The Future of Employment: How Susceptible are Jobs to Computerisation? 2013
- [4] Stiftelsen för strategisk forskning: Vartannat jobb automatiseras inom 20 år - utmaningar för Sverige 2014