

WHY DO WE NEED ROBOTIC & AI GOVERNANCE? AN ANALYSIS OF THE (SOCIO-) ECONOMIC IMPLICATIONS OF ROBOTICS AND ARTIFICIAL INTELLIGENCE

DOMINIK B. O. BOESL

*Technical University Munich (TUM), Munich, Germany; Vice President IEEE RAS
Industrial Activities; Chair IEEE TechEthics ad-hoc Committee; Senior Innovation
Manager & Vice President Consumer Driven Robotics with KUKA AG, Augsburg,
Germany*

MARTINA BODE

Junior Research Manager with KUKA AG, Augsburg, Germany

This paper illustrates the necessity of a Code of Conduct for Robotics and Artificial Intelligence – ethical and moral guidelines for the development and use of these technologies – by analyzing some of the potential (socio-) economic effects that might arise from technological advances in these fields. Contrary to other works in the field of roboethics, the authors did not strive to analyze individual cases that are of interests for ethicists and moral philosophers, but want to offer a broad, holistic and economic view on the effects of robotics and artificial intelligence on our future economy and, in the second instance therefore also society, as well as on the concept of Robotic & AI Governance for self-regulation of ethical, moral, sociocultural, socio-political and socio-economic questions, using a market mechanism.

1. Introduction

Robots are currently changing not only our manufacturing processes, but are gaining importance in other fields like service or consumer/household robotics. Europe's public view on these developments is rather critical and the fear that robots could replace jobs in the future is still present, even though studies suggest, that this is not the case (e.g. [1]) and that robots are currently creating more jobs than they are replacing. The European governments on the other hand are investing heavily in new technologies like service robotics, to create a competitive advantage for the region in the future. In countries like Japan, for example, robots are more accepted by the public, in general, due to cultural reasons.

Due to these changes, ethical and moral standards have to adapt. Researchers and scientists should critically ask themselves how their inventions will be used, who will profit from them and who will not as well as if they are sustainable [2] or if they can be misused, e.g. to commit crimes. In general, all these questions can be broken down to some general statements regarding for example the freedom of science and if it justifies the uncontrolled development of new and potentially dangerous technologies. To restrict research and development, on the other hand, means to hinder innovation.

Another one of these points is the missing “humanness” of machines. They cannot feel anything like compassion or sympathy and therefore it could be necessary to restrict the use of robots to fields where that lack of these qualities cannot lead to ethical conflicts. Asimov’s laws [3] of robotic for example do not allow the robot to harm humans, but is not considering this specific question regarding the lack of empathy. This leads to the idea if it could be possible to implant such a system of values into a robot. And could these values be compatible with Asimov’s laws? Many use cases, e.g. in service robotics would potentially imply a conflict of interest between the human’s freedom of choice and Asimov’s laws, for example if a robot is asked to serve alcoholic beverages to a customer, which are generally known to be harmful to health. How can a robot be enabled to make that decision? Can it distinguish between customers that are capable of taking that decision on their own (adults) and customers that are not (e.g. minors or mentally handicapped people)?

All these questions illustrate how complex the discussion regarding roboethics is. The goal has to be to prepare the generation of ‘Robotic Natives’ [4] to handle those autonomous robotic agents, to benefit from them and to co-exist with them.

2. Effects on the Labor Market

Thinking about the future of human labor, it has to be examined in which cases humans can, might or even should be replaced by machines and in which this seems impossible, improbable and undesirable. Many researchers think that the idea of an independent robot is science fiction. At the moment, automation of work is mostly restricted to industrial robotics, housed in factories and heavily constrained environments. The outside world is unpredictable and robots are in general better at pre-defined and narrow tasks and contexts. In addition, there are many legal questions yet unsolved [6].

Therefore, complex cognitive tasks, creative work as well as social-emotional intelligence can and will not be replaced by robots in the near future. Human labor will most likely experience a shift from physical work to more creative and complex jobs like artists, designers or public relation specialists. In addition, new, creative jobs might evolve in this process. The job market of the future will also see more part-time jobs as well as jobs where robots and humans work together. In addition, it is reasonable to expect that more people will work in jobs where social skills are important, such as e.g. elderly care, which could in addition help to solve the problem of labor shortage in these fields [7].

There are many studies regarding the amount of jobs that can be automated in general. Nevertheless, the utilization of new technologies is slowed down by economic, legal and social hurdles and the robotics industry is also creating new jobs. Taking heterogeneity of workers' tasks within occupations into account, across the 21 OECD countries, on average 9% of jobs are automatable, according to a report by Arntz et al. (2016) [8]. On the other hand, it is also possible to find reports that predict job losses of between 5 and 10 million jobs by 2020 [9]. According to a McKinsey report, technical work is more adaptable to automation. Jobs that require imagination, creativity, common sense, goal-setting and strategic thinking are harder to automate. Activities which involve managing and developing people for example have currently, according to McKinsey, only 9% automation potential, jobs that apply expertise to decision making, planning or creativity 18% [10]. The following figure illustrates the automation potential of jobs according to specific work activities:

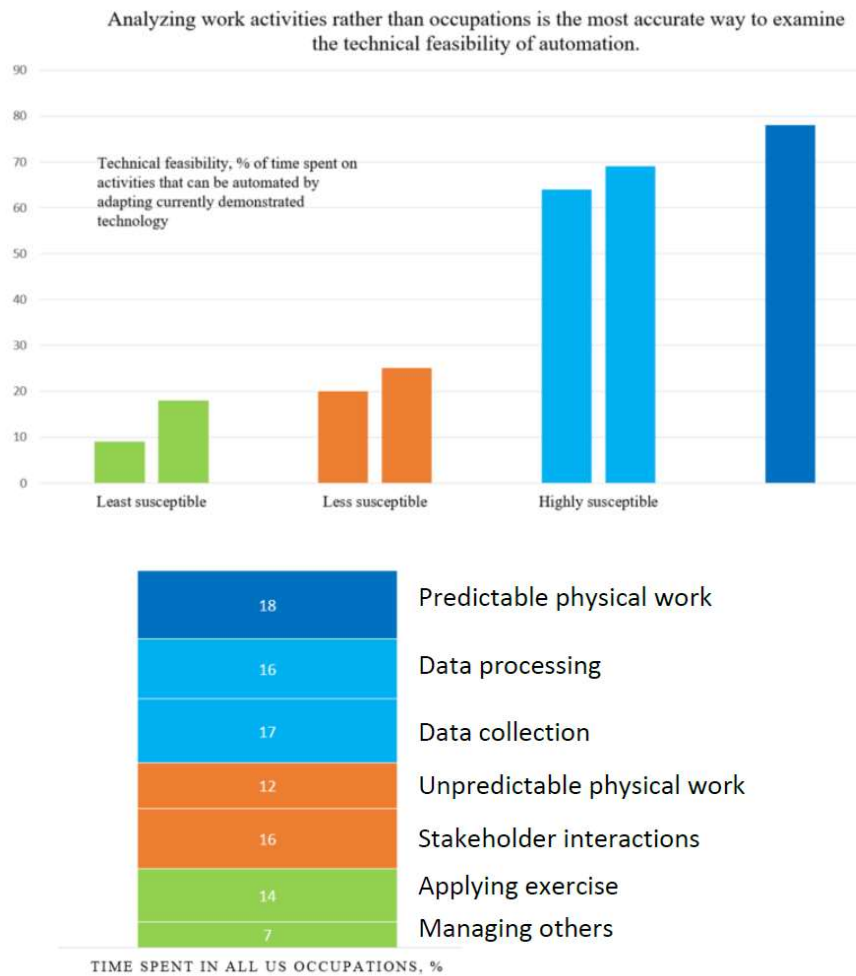


Figure 1. Work activities and automation potential [10]

There is much less research regarding the area of service robots and their effect on the labor market. A trend towards robots overtaking non-standardized tasks in this field of research can be noticed, due to technological advances. In addition, the prices for service robots are much lower than for industrial robots, which reduce the costs. Robots are not yet, but could become soon, a part of the value creation process in the service sector, even though this does not necessarily mean that they substitute human labor. They could also complement employees and make them more productive [11]. In a hospital for example, robots could

overtake transportation and logistics tasks and therefore the human workforce has more time available to spend with patients.

Maybe we do not necessarily have to be afraid to lose our jobs, due to an ongoing automation. Nevertheless, we will see a shift on the labor market and maybe not every employee is suitable for a job which requires creativity or leadership competence. The shift towards more service and social occupations will also mean that it could be necessary to re-educate employees and to change the perception and status of these jobs in our society. On the other hand, this could be a solution to labor shortage in these fields as well as an answer on how to deal with demographic change and an overageing society. Not only will there be more humans available to care for elderly people, less physically demanding jobs could also mean a chance for elderly people to be available on the labor market longer. Therefore, a structured and planned automation of work seems to be the reasonable solution to ensure, that these positive effects can be exploited without risking a de-stabilization of the labor market.

3. Military Robotics

New emerging technologies do not only change the labor market but have also the potential to change military and defense industry as well. Artificial intelligence and autonomous aerial and ground robots open up new ways of warfare. This also ethical raises questions like [12]

- Will autonomous robots be able to follow established guidelines of the Laws of War and Rules of Engagement, as specified in the Geneva Conventions?
- Will robots know the difference between military and civilian personnel?
- Will they recognize a wounded soldier and refrain from shooting?
- The ethical debate on military robotics is multi-layered. On the one hand, using robots in the battlefield could mean to preserve the lives of soldiers, who could have been harmed or killed otherwise.

The United Nations on the other hand urge to ban so-called “lethal autonomous weapons systems” before they are even developed [13]. Because of the increasing autonomy of robots, the prospect that systems could be used more flexible and therefore encounter influences which were not anticipated, as well as the complexity of technology, operational morality could not be sufficient for military robotics. Operational morality means that the actions of a robot are

completely in the hands of their designers or users and therefore the robot itself does not have to evaluate its actions regarding ethical concerns [12].

The effects of an unregulated development and use of these technologies can be illustrated using game theory. It could result in an AI arms race between superpowers like the US, Russia, the EU and China. Therefore, it seems necessary to prevent the development of such technologies before they exist, if this is not already too late by establishing regulations and not to found research in this field of expertise. Using robotics and artificial intelligence in military fields could mean advantages for the country or region regarding for example reaction time. This could result in a future where it is rational to ban human control almost entirely in military scenarios or applications [14].

The situation can be described by a social prisoner’s dilemma scenario. If we look at two parties, e.g. the US and Russia, it is best for every single party to invest in military robotics and artificial intelligence. Nevertheless, the overall welfare is higher, if no one invests in these technologies [15]. The problem in real world scenarios is, that there is imperfect information and therefore it is difficult or even impossible to determine if the other party is collaborating, before it is too late. Therefore, it is “safer” for every single party to invest in these technologies, just in case the other party does so too. The following payoff matrix illustrates this by showing an example for possible outcomes of the individual country as well as overall outcomes, resulting from the different decisions:

Table 1. Example for a payoff matrix - prisoner’s dilemma.

	Country B Not investing		Country B Investing	
	Country A Not investing	A: -2	B: -2	A: -6
	Overall: -4		Overall: -7	
Country A Investing	A: -1	B: -6	A: -5	B: -5
	Overall: -7		Overall: -10	

Regulation, e.g. in the form of a code of conduct could help to eliminate information asymmetries and therefore make it more profitable for all involved parties not to invest in the development and use of artificial intelligence and military robotics in these fields. Nevertheless, in order to make the mechanism work, commitment of all parties as well as monitoring are crucial factors to build trust in these standards or guidelines [5].

On the other hand, not stopping these developments could result in a so-called “race to the bottom” regarding safety standards, as already illustrated in the case of military armament and private possession of weapons [14], [16].

4. Regional Effects

An ideal scenario would be to establish worldwide binding laws that will cover all critical questions and adapt to new developments without any time delay. But of course, this cannot be realized in practice. Therefore, Robotic Governance strives to provide and establish an ethical and moral framework for the new fields of robotics and artificial intelligence and therefore fill the gap that is necessarily emerging until such values will develop in society. This should empower individuals to make moral decisions as well as stakeholders and the society to identify parties that are not complying to these rules.

Nevertheless, as the Robot Manifesto is not binding, it is not realistic to expect it to be established worldwide at the same time. But if, in the first step, countries and regions like Europe, which is often a pioneer in the field of regulation and standards, or some important manufacturers and research institutes will voluntarily commit to these values, this can lead to so-called California Effect [17], known from the environmental sector, and the adoption of the code of conduct in other regions and companies. Here, the influence of the regions or stakeholders complying are important factors as well as, of course, the pressure from possible image damage in the public.

Another possible effect is the emergence of clusters. Again, two scenarios seem possible, looking at the field of environmental regulation, as a benchmark. On the one hand, it is reasonable to expect, that restrictions in research and development will lead to a competitive disadvantage, if other regions will invest in these technologies. The results would be robotic clusters in economic regions, which do not comply to these frameworks. Clusters are defined as a regional concentration of organizations, which are all connected by the same industry or field of activity. Due to the geographical proximity externalities and a potential for competitive advantages arise [18]. The so-called pollution haven hypothesis implies that companies from countries with a strict regulation will shift their production facilities abroad, in order to avoid these requirements [19]. This raises the question, if Robotic Governance will lead to robotic or technology havens. But here, the soft-regulation or self-regulation has advantages compared to a legal solution. The companies and organizations comply to these rules either because they are convinced that it is the right thing to do, e.g. because they were involved

in the process of establishing these guidelines, or they expect a competitive advantage from complying or at least a competitive disadvantage from not participating because of strong market pressure.

On the contrary, the Porter hypothesis suggests that regulation can also foster innovation - contrary to common neoclassical theories, which are based on the assumption, that a normal, profit orientated company will already have exhausted all available possibilities to increase efficiency and that additional regulation therefore would only involve additional costs [20]. Porter and van der Linde (1995) believe, that regulation can make companies more productive and encourage investments and technical progress in new directions. In addition, responsible organizations could gain a competitive advantage [21].

One might argue that especially in this context a variety of regulations could be beneficial for single states and regions. When Brandeis (1932) described the United States as laboratories of democracy, he acknowledged the benefits of allowing differing regulations, which can be adapted to the different needs and wishes of the population and therefore allow a greater freedom of choice [22]. Nevertheless, since 1932, Globalization changed a lot. The population is not always the customer base of the economy in a country or region. The effect would be, most likely, that some regions in developed countries offer sustainable products for a niche market that is very informed and could even gather a competitive advantage in this market. On the other hand, in other countries, where regulation is less strict, companies could produce e.g. technologies which do not meet certain ethical or moral standards and therefore be cheaper or more attractive for certain markets.

5. Conclusion

Of course, this is only an overview of some of the potential economic effects of technological advances in the fields of robotics and artificial intelligence. A complete analysis is almost impossible. Nevertheless, the examples illustrate why Robotic Governance and a Robot Manifesto – a Code of Conduct for the robotics community – might be more efficient than legislation, from an economic perspective, in order to solve the ethical and moral questions currently arising from the development of new technologies like artificial intelligence and autonomous machines.

The authors believe that, due to the fast pace of these developments, we have to create a uniform set of values regarding the development as well as use of robot and artificial intelligence, not to be overrun by the rapid technological progress in this fields, like this was the case with digitalization and the internet. We should

enable individuals to make moral decisions in particular cases, based on these accepted and trusted values and guidelines, instead of trying to regulate every single, ethically questionable case that might arise during the development of these new technologies. Especially because it is very hard to predict future innovations and technological breakthroughs as well as their effects in advance. In addition, if the technologies are developed, there is no way to undo them again.

The goal of the robotics community should be to develop technologies that will enhance the quality of life and serve humans, instead of posing a potential threat. Robotics and automation technologies have the potential to solve many of the problems like labor shortage and an overaging society, the economies of industrialized countries are currently facing. Robots could take on jobs that are dangerous or unpleasant for humans. Nevertheless, we should shape our future actively, instead of reacting to changes and trying to solve issues, after they emerged.

References

1. M. Arntz, T. Gregory, and U. Zierahn, *Digitalisierung und die Zukunft der Arbeit: Makroökonomische Auswirkungen auf Beschäftigung, Arbeitslosigkeit und Löhne von morgen*, ZEW (Zentrum für Europäische Wirtschaftsförderung GmbH), Germany, 04/04/2018 (2018).
2. G. H., Brundtland, *Our common future*, Report of the World Commission on environment and development, United Nations (1987).
3. I. Asimov, *I Robot*, New York: Doubleday & Company (1950).
4. D. Boesl, and M. Bode, *Generation 'R': Why our grandchildren will grow up as the first Generation of "Robotic Natives"*, Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech), IEEE, 417-420 (2016).
5. D. Boesl, and M. Bode, *Roboethics and Robotic Governance – A Literature Review and Research Agenda*, In: A. Ollero et al. (eds.), *ROBOT 2017: Third Iberian Robotics Conference, Advances in Intelligent Systems and Computing* 693 (2018).
6. A. Lyyra, *The idea of robots as independent machines is science fiction*, London School of Economics (2015).
7. J. J. Hughes, *What Is the Job Creation Potential of New Technologies?*, In: *Surviving the Machine Age*, 131-145 (2017).
8. M. Arntz, T. Gregory, and U. Zierahn, *The Risk of Automation for Jobs in OECD Countries: a comparative analysis*, OECD Social, Employment and Migration Working Papers, **No. 189**, OECD Publishing (2016).
9. J. Pistrui, *The Future of Human Work Is Imagination, Creativity, and Strategy*, Harvard Business Review, 01/18/2018 (2018).

10. M. Chui, J. Manyika, and M. Miremadi, *Where machines could replace humans—and where they can't (yet)*, In: McKinsey Quarterly, 07/2016 (2016).
11. M. Decker, M. Fischer, and I. Ott, *Service robotics and human labor: A first technology assessment of substitution and cooperation*, In: Robotics and Autonomous Systems, **87** (2017).
12. P. Lin, G. Bekey, and K. Abney, *Autonomous military robotics: Risk, ethics, and design*, California Polytechnic State University San Luis Obispo (2008).
13. O. Bowcott, *UN urged to ban 'killer robots' before they can be developed*, In: the Guardian, 04/09/2015 (2015).
14. T. Metzinger, *Towards a Global Artificial Intelligence Charter*, STOA - Science and Technology Options Assessment (2018).
15. P. D. Straffin, *The Mathematics of Tucker – A Sampler*, In: The Two-Year College Mathematics Journal, **14 (3)**, 228–232 (1983).
16. J. Heath, and A. Potter, *The Rebel Sell*, Harper Collins, Toronto (2005).
17. D. Vogel, *Trading Up: Consumer and Environmental regulation in a global economy*, Harvard University Press (1995).
18. M. E. Porter, *Location, Competition, and Economic Development: Local Clusters in a Global Economy*, In: Economic Development Quarterly, **14(1)**, 15–34 (2000).
19. A. Levinson, and M. S. Taylor, *Unmasking the Pollution Haven Effect*, International Economic Review, **49 (1)**, 223–54 (2008).
20. S. Ambec, M. A. Cohen, S. Elgie, and P. Lanoie, *The Porter hypothesis at 20: can environmental regulation enhance innovation and competitiveness?* Review of environmental economics and policy, **7(1)**, 2-22 (2013).
21. M. E. Porter, and C. van der Linde, *Toward a New Conception of the Environment-Competitiveness Relationship*, 99–101 (1995).
22. New State Ice Co. v. Liebmann, 285 U.S. 262 (1932).