Ethics and Technology: Some Issues

Author: Dr. Mirko Garasic, Dr. Steffen Steinert
Project Initiators and Designers:
Dr. Tina Ladwig, Axel Dürkop, Michael Heinemann

TEKETHICS

hoou.de

Ethics and Technology: Some Issues

A Booklet

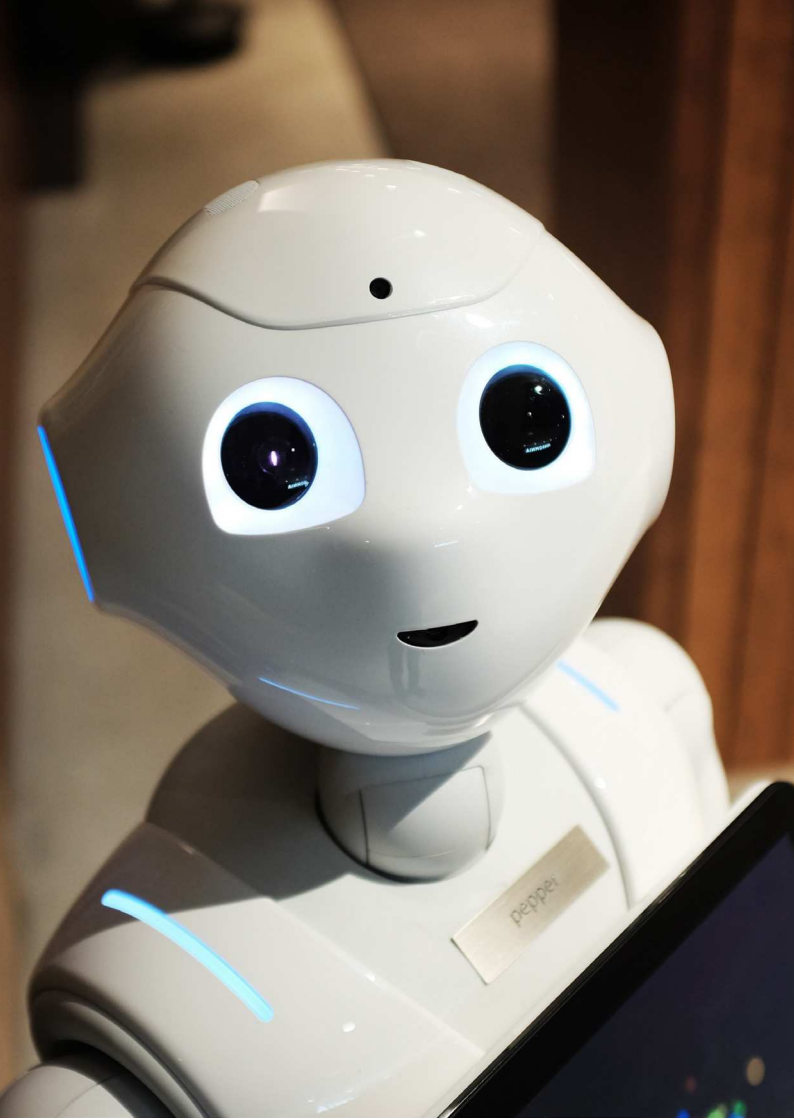
Dr. Mirko Garasic Dr. Steffen Steinert

Hamburg University of Technology
Open Discourse Series

Ethics and Technology: Some Issues - A Booklet. A Project by Hamburg University of Technology (TUHH) in Hamburg Open Online University (HOOU)

Authors: Dr. Mirko Garasic, Dr. Steffen Steinert
Editors: Prof. Dr. Sönke Knutzen, Dr. Tina Ladwig, Axel Dürkop, Michael Heinemann
Co-Editor: Sabrina Maaß
Cover Design: Johanna Mörke
Date of Publication: March 2018
DOI: https://doi.org/10.15480/882.1570

Contents


pepper

1 Intelligent machines are on the rise

Do we believe in technology as inherently representative of progress or should we fear that a Terminator is on its way?

AI and Robotics

The link between technology, artificial intelligence (AI) and robotics is probably so evident that it does not need much analysis. What should instead be at the centre of the present work is the ethical aspects that are peculiar to these two – often intertwined – spheres of inquiry: AI and robotics. Before proceedings further, we should thus clarify how we differentiate between the two.

AI refers to a more complex and theoretical concept (intelligence), and it is therefore in itself a more aleatory concept that can span through a number of fields of research. Most relevantly however, we could ascribe AI to belong to computer science – representing a branch with a particular emphasis on the creation of intelligent machines as human as possible. There are a number of standard tasks that modern, widespread machineries are already capable of doing thanks to their AI. For example, speech recognition allows our smartphones (remember that smart here means intelligent) to make a phone call or give us directions based on their capacity to recognize our voices, to process what

we are asking them and then to provide us with an answer. Or to make an action.

Hence, it is not surprising that society is currently pushing for the creation for a number of life changing innovations – such as self-driving cars for example – where there is a need for an enormous amount of data processed. Machines can interact intelligently only if they are provided with ways of categorizing the world that is commonly defined as knowledge engineering. Equally important is machine learning: without supervision, learning needs clear patterns in streams of inputs, while when supervision is available, numerical regressions and classification are involved.

Robotics is the other side of the coin of AI. Robots require (different degrees of) AI in order to successfully handle the job we assign them.

In some instances, however, robots can also be partially developed in terms of *independent thinking*, or not depending on AI but rather on the intelligence of humans.

For example, the use of a robotic arm in a study conducted at Brown University[1] aimed specifically at using the intelligence of the patient included in the study (in this cases for therapeutic reasons, as they were tetraplegic).

[1] https://news.brown.edu/articles/2012/05/braingate2

Although such examples are defined as Brain-Computer Interface (BCI), this type of use of a computer is more *mechanical*, and it assumes a supervision – if not a full engagement – of human intelligence. Hence making the distinction between AI and Robotics more evident.

What kind of ethical problems relate to AI and Robotics?

Improvements related to the implementation of AI and robots are various: from the use of intelligent technology in medical contexts to the increased safety and personalization of many of our everyday gadgets, it would be pointless to deny that life can be made easier by these innovations.

Yet, many ethical issues arise from these new discoveries as well, and we should look into some of the most prominent ones.

Self-driving cars: who should AI let die?

Self-driving cars (or Autonomous Vehicles) are in the making. From BMW to Uber, major transport companies are investing in projects aimed at providing a full-scale autonomous, intelligent car able to drive without anyone directing it.

This would represent a huge revolution for transports: we would not have to worry about parking somewhere downtown or if we had too much wine at dinner, the car will not be affected by it. What will affect the car behaviour however, is the surroundings along which it will ride. Aside from the interaction with other vehicles, the most impelling question is: how would the car react to a situation of foreseeable accident?

A very successful website put online by MIT[2] has been trying to create a number of statistical data showing how we would respond to relatively similar moral scenarios. Building on a famous philosophical thought experiment referred to as the *Trolley Problem* (we will explain more in depth this in the **Neuroscience and the Law** Chapter) the website puts the visitor in a position to choose what in her/his opinion would be the right behaviour (or else, what s/he would do in such a situation) for the car to have. Would you rather run over two old ladies or a healthy young runner? Should one's social status and past behaviour be taken into account when deciding whose life should be spared in such extreme cases?

The website is very interesting because: a) it shows how inconsistent humans' choices might be – raising suspects that a case to case moral assessment of the scenario could produce injustice; and b) for the opposing reasons, the experiment

[2] http://moralmachine.mit.edu/

shows how a human eye could be able to detect important, specific variables that could make the exception to the rule morally acceptable – or even morally required.

Robots and relationships

While waiting to disentangle how robots might be driving in our streets, much research has been put forward to use them in more and more intimate part of our lives, allowing robots to enter our homes with augmented centrality and freedom.

This is the case of care robots – robots programmed to help a number of people in society that can finally rely on constant helper by their side when needed. Groups include people with different forms of physical impairments and disabilities (be them temporary or permanent), as well as segments of the population that structurally need more attention – for now it is the elderly, but it is not that unthinkable to imagine a *kindergarten roboteacher* in the near future.

Though probably moved by good intentions the results of this trend in dealing with human relationships poses a number of questions that are everything but banal. We will get back to those in a second, but first it should be pointed out that a threat that often goes with technology has perhaps gone overlooked in the case of robots: they can be hacked

like any other technological gadget – posing serious concern on how this option could be used by criminals.[3]

For example, it would be sufficient to connect to the robot inside the house for a few moments to allow robbers into the house and empty (hopefully without the owner home). The increase in dependence from these new *collaborators* opens new potential worries that should not be underestimated for the sake of preserving the hype for innovation and *technology at all costs* we currently live in.

An even more specific ramification of robotics is represented by what are commonly referred to as sexbots (robots programmed to satisfy our sexual needs). Also here – if not perhaps more – there is room for the possibility for this technology to help some people in particular situations (for example replacing sex workers in addressing this specific need for people unable to engage in *normal* relationships due to a particularly dramatic accident or physical impairment).

Still, despite the social value that those robots could represent, there is concern on how this revolution in our sexual sphere could end up affecting society, and considerations on the impact of these intimate entertainments should abound.

To begin with, the availability of passive, programmable sexual partners could create a number of societal dysfunctions:

[3] https://www.ft.com/content/1552b080-fe1c-11e6-8d8e-a5e3738f9ae4?mhq5j=e3 (registration required)

starting from our unidirectional way of interacting with a partner in an intimate and very relevant part of our lives, we could soon develop similar attitudes outside this relationship and behave accordingly when interacting with other human beings – within the sexual sphere, but more broadly still.

This repercussion could easily fuel an already problematic gradual detachment from *the other*, rendering our civilization poorer in terms of solidarity, empathy and other key values for a well-functioning society.

Robots and wars

In relation to the issue of preserving our human side, another version of robots needs analysis. In recent years, there has been a gradual implementation of unmanned airplanes (drones) in wars. Remotely directed from pilots, often sitting in sophisticated labs on the other side of the planet, drones are now widely used in military operations by – mostly – Western countries basing their moral legitimacy on the ground that they limit the number of casualties and help preserving human lives (in this case, the lives of soldiers). Leaving aside the political analysis concerning the possible discrepancy of value within equal lives in this instance, we need to investigate further into the implications of this technology as a more complex issue than so described,

as it would be too easy to see this only in a positive light if this was the case.

Somehow related to the concern raised above, one should question the impact that this detaching technology could (or does already) have on our way of interacting with other human beings and the way in which it can shape our behaviour beyond the supposed given functionality.

For example, those pilots that are manoeuvring the drones from their home countries, finish their shift (during which they might as well have killed some innocent babies by accident) and get catapulted back into their *normal* lives. Paradoxically this schizophrenic way of living their lives seems to produce more problems and imbalances than initially thought, and the effects need to be considered with attention.

Also important, is to consider the increasingly realistic dystopian scenario of war robots allowed to make a call *on their own* as to whether to shoot or not. There is wide consensus that we should not aim at programming a robot able to shoot a human being without the green light from another human being, but it is realistic to imagine contexts in which such an option would not be seen as particularly problematic. Even more relevantly, we should bear in mind the hackability of these technologies and wonder about the potentially catastrophic effects of leaving them in the wrong hands.

Robots and automation

Lastly, it is inescapable to point out the socio-economic impact of robots. Politicians, academics and policy makers have begun to engage with the phenomenon of increased industrial automation with more attention, as this represents for many a threat to the jobs of many.

In the attempt to clarify who is to blame in the case of an accident resulting from a mistake made by a robot the European parliament has recently proposed to grant some rights to machines, so to make them legal entities[4]. Not surprisingly, this has generated an intense discussion among scholars as well as the general public, as many see this move as the first step towards the creation of additional competitors in a world already short of jobs and overpopulated. In line with an ever-increasing automation of our chains of productions, sceptics of the positivity of the *roborevolution* see the increase of independence and consideration of robots as directly related to a decrease of value of human beings – workers or otherwise.

Technoenthusiasts instead, affirm that this is indeed the path towards greater social justice and individual growth: by allowing robots to independently deal with mechanical jobs and alienating jobs, we will ensure more opportunities for

[4] http://www.europarl.europa.eu/news/en/press-room/20170210IPR61808/robots-and-artificial-intelligence-meps-call-for-eu-wide-liability-rules

human beings to follow their own creativity and tailor themselves with a more unique profession. This optimistic view is of course very tempting, but the recent failure of the *internet experiment* (expected to guarantee a drastic increase in democracy and moral growth for humanity) demands from us to be careful in the assessment of how to move next.

Further readings

Pitsch, K. 2016. Limits and opportunities for mathematizing communicational conduct for social robotics in the real world? Toward enabling a robot to make use of the human's competences. *AI & Society* 31(4), pp. 587-593. https://doi.org/10.1007/s00146-015-0629-0

Scheutz, M. & T. Arnold. 2016. Feats without Heroes: Norms, Means, and Ideal Robotic Action. *Frontiers in Robotics and AI*. https://doi.org/10.3389/frobt.2016.00032

Veruggio, G., Solis, J. & M. Van der Loos. 2011. Roboethics: Ethics Applied to Robotics, *IEEE Robotics & Automation Magazine* 18(1), pp.21-22. https://doi.org/10.1109/MRA.2010.940149

Weiss, S. 2013. Raunchy robotics: the ethics of sexbots. https://hcri.brown.edu/2013/06/18/raunchy-robotics-the-ethics-of-sexbots/

Related videos

Mirko Garasic - What kind of ethical problems relate to AI and Robotics?
https://www.youtube.com/watch?v=jbL7F7kJRc8

Mirko Garasic - Robots and Automation
https://www.youtube.com/watch?v=6FOKx41e_y0

Related discussion

See also the discussion related to this chapter at

https://community.tuhh.de/t/1-chapter-intelligent-machines-are-on-the-rise-do-we-believe-in-technology-as-inherently-representative-of-progress-or-should-we-fear-that-a-terminator-is-on-its-way/525

2 How far should human enhancement go?

Human Enhancement[1]

Human enhancement (HE) has gained increased visibility and popularity in the last decades. This is as a result of a substantial increase in the number of fields interested in the topic, namely: nanotechnology, biotechnology, information technology, and cognitive science (NBIC). The rapid innovation in these fields has created a theoretical space for speculation over the moral acceptability of aiming to improve or expand human capacities through the implementation of genetic engineering, drugs and technology so to give rise to a number of subgroups such as cognitive enhancement (CE), moral enhancement (ME) and emotional enhancement (EE).

Many academics in the Western world, including both scientists and philosophers, now favour the enhancement of human beings with the tools of science, and the motto *if we can, we should* is sometimes used to summarise this position. The appeal of reducing human weakness and potentially improving a range of human abilities is obvious, but both moral and political questions remain as sources of dissent. We can

[1] This part of the chapter builds on some of my previous work on the topic. Garasic, M.D. 2012. Human Enhancement in the EU, *Australian and New Zealand Journal of European Studies* 4, pp.31-41.

make a general distinction between those arguing in favour of HE (described as *bioliberals*, *posthumanist* or *transhumanists* depending on the context) and those against interventions on the human condition that go beyond standard function (this latter group is that of the *bioconservatives*).

What do we mean when we talk about Human Enhancement?

Since the inception of this discussion, Nick Bostrom and Rebecca Roache have demonstrated that exponential developments in nanotechnologies, synthetic biology, neurology and genetic engineering have created the premises for a reconsideration of what it means to be human, as well as in which way(s) and to what extent we should interfere with humanity's natural disposition. Now not only can the most basic human capacities of an individual be restored to their initial levels before an injury or illness; these basic capabilites can also be enhanced.

A comparison between enhancement and therapy is instructive. While therapy aims to fix a problem or dysfunction and thereby allows the individual to regain a status of *normality*, enhancement aims to break the barriers of normality and go beyond the natural limits of mankind. This controversial distinction needs to be clarified. Here, therapeutic treatment is defined as: a) designed to re-establish a standard level of func-

tionality; b) reversible; c) less invasive and expensive than enhancement. In contrast, enhancing treatment is: a) designed to move the level of functionality above the standard; b) not easily reversible; c) more invasive and expensive than therapy. The following example elucidates the above definitions. The availability, purchase – and use – of a pair of glasses could be justified as part of the therapeutic treatment of poor vision, whereas, the process of neural implantation aimed at increasing one's memory by 20% would not. While the overlapping threshold between therapy and enhancement is open to debate, it is nonetheless important to have a frame of reference for the terms considered.

The famous example of the South-African runner Oscar Pistorius encapsulates the current difficulties in drawing a distinction between enhancement and therapy. Born without fibulae, he had his legs amputated below the knee at a very young age and replaced with prosthetic legs. Thanks to his commitment to training and to the implementation of a pair of Flex-Foot Cheetahs, in 2005 Pistorius managed to win the T44 200m gold medal.

When his results started to match and pass those of the *normal* runners however, the debate over therapy and enhancement turned into a polemic. In 2007, the International Association of Athletic Federations (IAAF) reported that "Pistorius was able to run at the same speed as able bodied athletes while using less energy and that his prosthetic limbs gave him

an advantage over able-bodied athletes."[2] By common standards, when Pistorius started using prosthetic legs as a child, he had merely undergone therapeutic treatment. Yet, in a situation where his artificial limb prostheses were identified to be a technical advantage, he was considered to have undergone enhancement. Of course, this makes his case even more germane to the ethical and political dimensions of the therapy versus enhancement debate: is it enhancement if the person in question is responding to a congenital condition over which s/he had no control?

Enhancements like those in the Pistorius case, are merely physical ones, hence – in a sense – less problematic (though part of the HE agenda and surely definable as physical, we will leave the more controversial issue of life extension for the chapter on environmental ethics and technology) than those enhancements that instead are aimed at improving (or simply changing) some of our less immediately visible characteristics such as predispositions and attitudes towards life and people, as well as our intellectual abilities and responses. Yet, before moving to analysing the various types of non-physical enhancements, we should briefly consider some of the arguments made for and against HE.

[2] http://oscarpistorius.com/

Fears from the past

As enhancements (especially the psychical ones) could also be genetic, once that the political dimension enters the debate, it is unescapable to think of eugenics. Even though the ideology had its genesis in Great Britain and found much success in the United States, the field of eugenics is associated with Nazi Germany, and the atrocious consequences of their projects. It is no surprise that the terrible and recent experience of the National Socialist era makes discussion of HE in Europe an especially difficult topic. Among others, the great German sociologist and philosopher Jürgen Habermas, in his book **The Future of Humanity** has brought forward the following argument against eugenics: the application of any form of genetic enhancement to our offspring would undermine the autonomy of those yet-to-be-born individuals.

Among other critiques against Habermas' view, two are particularly relevant and worth of consideration in this context. First, one could argue that the limitations of physical autonomy that a severely disabled person might have to suffer from as a result of a lack of implementation of accessible genetic enhancement tools will be a much greater destabilising factor on the individual's freedom to pursue her/his own priorities in life than the knowledge that her/his parents chose her/his genes scientifically. Second, we could apply the famous argument put forward by Nicholas Agar in his work **Liberal**

Eugenics: if we accept – as we do – environmental enhancement (e.g. private schools, music lessons, exclusive sport facilities and so on) as acceptable parental behaviour, why should we not think of genetic enhancement as a variant of the same principles?

Of course, the counterargument to this assessment of what should be permissible could also be made. Namely, we could affirm that, given that environmental enhancement is producing an ever-growing gap between the *haves* and the *have-nots*, and in light of the negative moral and political results that this entails, it would probably be wise not to implement genetic enhancement either.

Politics and Enhancements

Sarah Chan and John Harris accurately describe "*an enhancement (as we are using the term) is something of benefit to the individual.*"[3] This commonly shared transhumanist definition of enhancement could be seen as sufficient proof that distributive justice is not truly a central consideration of HE ideology. After all, at the root of HE, there is a Hegelian vision of progress. Relevantly, the same Harris draws a parallel between HE and the use of candles, affirming that we should

[3] Chan S. & J. Harris. 2007. In Support of Human Enhancement, *Studies in Ethics, Law, and Technology*, 1(1), pp.1-3. https://doi.org/10.2202/1941-6008.1007

not ban the former in light of the benefits that we can perceive to have achieved through the use of the latter. The only things we have to work on, Harris says, are working hours, minimum wages and so on.

In response to this historical analysis (and, as a result of the potential outcomes of future implementations of HE), it might be worth applying a Left-Hegelian perspective to consider that the example proposed by Harris may demonstrate that the advent of candles may have been the point at which the level of exploitation of workers reached a tipping point. The industrial revolution has certainly resulted in certain benefits and progresses for a part of humanity, but it is equally true that it represented a turning point in delineating once and for all the gap between rich and poor countries. If such a gap has been consistently increasing since its inception, are we sure that we want to start another revolution that might run along the same tracks? Most importantly, could we convincingly affirm that this process would constitute an enhancement for humanity as a whole rather than only for a select or lucky few?

It appears clear that there are a number of issues related to HE, but before proceedings in analysing the specificity of the biochemical enhancers, we should first look at recent techniques that instead use electricity to re-calibre our interactions with other human beings.

Non-Invasive Brain Stimulation[4]

Non-Invasive Brain Stimulation (NIBS) is divided into a number of subgroups.

Transcranial magnetic stimulation (TMS) uses magnetic induction to produce electric current through the scalp and the skull allowing for both neurostimulation and neuromodulation. The main uses of repetitive TMS (rTMS) are currently linked to the diagnosis and treatment of certain disorders, as well as the study of brain functioning, thanks to the ability to also block the activity of certain brain areas in a selective and reversible manner (polarizing neuronal membranes).

Transcranial current stimulation (tCS) uses electrodes placed on the scalp to deliver a weak current (1-2 mA) to the brain. There are several techniques of tCS – tDCS (transcranial direct current stimulation), tACS (transcranial alternating current stimulation) and tRNS (transcranial random noise stimulation) – but most studies are focused on tDCS, which is mainly used to modulate excitation and inhibition and to alter and improve cognitive functioning.

[4] Andrea Lavazza is to be thanked for his contribution in our co-authored paper to which I refer extensively in this part. Lavazza, A. & M.D. Garasic. 2017. How Non-invasive Brain Stimulation Might Invade Our Sphere of Justice, *Journal of Cognitive Enhancement* 1, pp.1-8. https://doi.org/10.1007/s41465-017-0008-5

The use of NIBS is the combination of some of its unique characteristics with the spread of its use for enhancement purposes in both cognition and sport performance. Not surprisingly, the gradual spread of the use of NIBS – and tDCS in particular – was first motivated by its therapeutic value.

The insight that the application of electric current can modify the function of the nervous system has found progressive confirmations and today we know that TMS can be used to normalize areas of abnormal activity due to illness. Indeed, several studies have shown that rTMS can have a positive effect on mood in patients with depression, while tDCS can have an apparent therapeutic potential for chronic neuropathic pain, parkinsonism, stroke recovery, tinnitus, traumatic spinal cord injury, depression, and drug addiction, even though tDCS has not been approved by the FDA for any therapeutic applications and these treatments are presently considered an *off-label* application.

Among other areas of enhancement, those are worth of special attention perhaps: memory, reading, mood, learning, perception, mathematical cognition, decision-making, motor skills, creativity, motivation, and moral reasoning.

On the practical side, it is interesting to take into account that athletes of the national ski-jumping team of the United States and some sprinters of various nations that took part in the Rio Olympics allegedly experimented with the hand-

set for tDCS developed by Halo Neuroscience so to improve their athletic performance. In particular, stimulating motor areas during training would improve coordination, strength and fatigue resistance. This has also attracted the interest of the World Anti-Doping Agency, which is monitoring the use of tDCS to improve athletic performance to evaluate whether to include it in the list of treatments forbidden by the International Olympic Committee. The main criteria to be included in the list are risks for the athlete's health and potential violation of sportsmanship.

Other important issues have arisen from the debate on the ethical guidelines that should govern this technology. It has been documented that relatively low cost and easiness of manufacturing of tDCS could led to a DIY-tDC phenomena. Indeed, equipment is cheap, easily available, and apparently very simple to use. Hence, some private companies have put on the market devices for NIBS, such as Foc.us, which was advertised as being able to improve attention and memory so as to make users better at video games. Actually, a scientific study has found that the effects are not at all what had been promised; indeed, stimulation with the device in question seems to worsen the user's memory. However, Do-It-Yourself brain stimulation seems to be growing, as testified by various studies.

As a response to this social risk deriving from an unregulated use of NIBS, a group of leading neuroscientists in the field

has recently published an open letter warning about the risks of stimulation if not monitored by experts. In the letter, the scholars note that: (1) stimulation affects more of the brain than a user may think; (2) stimulation interacts with ongoing brain activity so that the specific activities carried out during stimulation modify the effects of tDCS itself; (3) enhancement of some cognitive abilities may come at the cost of others; (4) changes in brain activity (intended or not) may last longer than a user may think; and (5) small differences in tDCS parameters can have a big effect. In short, safety considerations seem therefore very important and also require, in addition to ethical ones, attention from the scientific community and the social and political authorities in view of choices of regulation, given the growing spread of self-administered NIBS. The use of tDCS was also discouraged in military and security services.

Finally, it should be taken into consideration that, when it comes to neurocognitive enhancement, there seem to be reasons to set limits to total personal autonomy. In particular, in competitive-selective contexts – such as job interviews – issues of fairness and overall social efficiency are at stake and should be addressed.

Cognitive Enhancement

Developments in neuroscience are gradually exposing more and more the way in which our brain responds to substances – shaping our interaction with the world in accordance. In the literature this is usually referred to as CE. For example, amphetamines such as Ritalin or Adderall (initially meant to be used only for therapeutic use by people affected by ADHD or narcolepsy) are now widely used in academic and military contexts as ways of boosting one's attention, responsiveness and ability to focus beyond normal level. They can also significantly reduce fatigue and hunger, making their use extremely appealing for students stressing out for the submission of a paper or armies for having more efficient soldiers serving.

Aside from obvious concerns related to the safety of these drugs, probably the most problematic issue specifically related to CE (as in the case of NIBS that also represent a form of CE, though not biochemically based) is its interaction with distributive justice – more specifically still with the disclosure of its use in competitive-selective contexts. Would we consider to be fair for a person to get a job over another if we knew that s/he took advantage of a pill that allowed her to be more concentrated during the entrance test? Probably not.

Moral Enhancement

One possible *internal* way of solving this issue – and others – of unfairness, some authors claim is to take the biochemical enhancement to the next level: instead of only improving our technical skills in achieving a certain goal, we should address more directly our (faulty) morality, so to ensure a greater good for humanity. Along those lines, Ingmar Persson and Julian Savulescu argue that if we are to survive as a species, we need to morally enhance ourselves (so to alter our empathy through an intake of oxytocin for the sake of being less aggressive for example). The idea derives from the reading of our biological history somehow not as evolved as our technological one: a single individual is able to create enormous damage with a dirty bomb say, while a lack of coordination (possibly moved by more or less conscious selfish interests) is not allowing us to face vital issues such as global warming. Calibrating ourselves towards a more cooperative way of interacting with society would produce benefits for society that we should seek.

Of course, this approach is far from unquestionable. Among other possible critiques, two seem particularly powerful.

On the one hand, there is a *bioconservative* worry. Such a tool is positive if we give for granted that the morality enhanced is universal and unquestionable, but surely we have plenty of examples were moral disagreement is in place. Scarier still,

what would happen if those in charge of the moral enhancers would push for questionable – or even immoral – goals?

On the other hand, the critique arrives from within the *posthumanist* tradition itself. According to Harris for example, there is no moral justification in seeking to alter our *freedom to fall*[5] . We can enhance our cognitive capacities so to be more likely what is best for us and others, but we should not alter our core structure that guarantees that we make a certain (moral or immoral) choice in accordance to our fallible – yet unique and autonomous way of seeing the world.

Emotional Enhancement

Finally, EE is the most recent one of the subgroups within the cluster of biochemical enhancers, and it is possibly the most controversial of all. The idea at the bottom of this specific category of enhancers, is that why could and should use the knowledge that we are getting from studies on our brain to help us understand more what triggers our emotional responses and how we could control it.

The case is particularly powerful when considering instances of romantic love that creates unhealthy behaviours within

[5] Harris, J. 2010. Moral enhancement and freedom. https://doi.org/10.1111/j.1467-8519.2010.01854.x

ourselves (dependence) or from our partners (domestic violence). Hence, the portrayed scenario of having a pill that could help an abused lover to detach her-/himself from an abusive and violent partner seems very tempting of course. Yet, the possibility of chemically redirecting our emotions according to what we rationally decide is not free of worries. Could we not get rid of love altogether then? What would that leave us with? Also troubling, is the way in which this emotional self-creation could affect sexual minorities: if provided with the tools to *cure* themselves, many homosexuals might feel pressured by their conservative environment to change their *sexual tastes* and this scenario seems like restricting our freedom rather than broadening it.

Further readings

Agar, N. 2005. Liberal eugenics: in defence of human enhancement. Malden: Blackwell Pub.

Bostrom, N. 2010. Letter from Utopia, *Studies in Ethics, Law, and Technology* 2(1), pp.1-7. http://www.nickbostrom.com/utopia.pdf

Garasic, M. D. & A. Lavazza. 2016. Moral and social reasons to acknowledge the use of cognitive enhancers in competitive-selective contexts, *BMC Medical Ethics* 17(1) https://doi.org/10.1186/s12910-016-0102-8

Garasic, M. D. 2017. Enhancements 2.0: self-creation might not be as lovely as some think, *Topoi* https://doi.org/10.1007/s11245-017-9469-5

Habermas, J. 2006. The future of human nature. Cambridge: Polity Press. http://doi.org/10.1080/15017410500246111

Sandel, M.J. 2002. What's Wrong with Enhancement https://bioethicsarchive.georgetown.edu/pcbe/background/sandelpaper.html

Savulescu, J. & I. Persson. 2012. Moral Enhancement, *Philosophy Now*, Issue 91 https://philosophynow.org/issues/91/Moral_Enhancement

Related videos

Mirko Garasic - What do we mean when we talk about Human Enhancement?
https://www.youtube.com/watch?v=nDIAigxw2oM

Related discussion

See also the discussion related to this chapter at

https://community.tuhh.de/t/2-chapter-how-far-should-human-enhancement-go/541

3 What kind of duties do we have towards nature, and why?

Engineering and Environment

Most people would perhaps not imagine environmental ethics (EE) to belong in a module on technology and ethics – possibly by an approximate reference of environment meant as an extension of nature, with this latter term seen as antithetical to technology. Yet, this would be a misplaced assumption, and we should look at how environmental ethics is deeply connected with advancement in biotechnology and what kind of problem this connection poses. To begin with, we should look at the broader definition of EE, to then move to analyse how an increased sensitivity towards an awareness of the limits that we should pose to ourselves if we are not to damage the ecosystem beyond a point of no return. Hence, we will consider what current argument seeking life extension (or even immortality) would imply, what genetic engineering allows us to do to both humans and nonhumans (animals and plants) alike and in which way we could use engineering more broadly to help us facing global warming.

Environmental ethics

EE is concerned with the relationship of human beings with the natural environment. As a result of the advancements in industry and technology, population growth and economic expansion, since the 1960s the impact of human beings on nature became gradually more the focus of studies due to ethical concerns towards wildlife.

The drastic increase of human population in the last three decades, pollution and the depletion of natural resources have added worries in regards to the preservation of the earth for future generations, the loss of wilderness and the issues of public health, yet we could say that EE evolves around two main questions: what duties towards the environment do we as human beings, and why do we have them? In answering the latter one (in a sense, the most crucial one), other more philosophical sub-questions emerge, such as: would it make sense to care about the environment if there were no human beings left?

Without entering the deep dimension that an attempt to answer this question would require, we must nonetheless point out that a negative response would stress our (surely Western but not only) anthropocentric approach towards nature.

For some, this attitude is a problem. Namely, the fact that in some tradition human beings are the only ones granted

with a moral standing, has been seen by some environmentalists as a structural problem that has not helped us becoming sensitive enough towards animals and plants – contributing in making us consider nature as something external, thus exploitable and worth of true care.

Whether we use anthropocentric or non-anthropocentric reasoning to determine what can be considered to be ethical behaviour in relation to the environment, it would appear as if we have reached a stage where we need to act against pollution and global warming anyhow. However, our view could instead shape our evaluation of other ways to implement technology in (broadly speaking) environmental issues.

Genetic engineering

Genetic engineering, also called genetic modification, is the application of techniques from biotechnology and bioengineering to modify the genetic makeup of an organism. Instead, by transgenics we mean a specific process of genetic engineering that aims at removing genetic material from one species of animal or plant (it is still largely considered unethical to even conceive the idea of applying this technique to humans) and add it to a different species previously deprived of such a characteristic.

Certainly transgenics and genetic engineering have enormous potential – both commercially and scientifically. Hence, though worth of attention, care should abound in assessing the limits of the implementation of these techniques, as this form of engineering carries with itself some specific ethical considerations that we need to address through some examples.

GMOs and the environment

Genetic engineering has already been a reality that most of the readers will be familiar with under the name of *GMOs* (Genetically Modified Organisms). Although inclusive of a larger group of GMOs, we usually refer to this group of organisms as fruit and vegetable that have been modified (although the most accurate nomenclature should be GM food), but only by removing some weak characteristics (e.g. a tendency to become ripe too fast) rather than including new genetic variables from other species (e.g. a fluorescent tomato might be useful to find in the dark but it is still not worth the risk of being tasted according to the scientific community). More recently, GM food have also included insertion of synthetic genes, but even a less invasive way of altering food does not escape various questions. For example, should these new versions of food have priority over old ones?

GM food such as crop for instance, are extremely aggressive towards *old fashion* cuisine: in a relative short time, growing two types of crop next to each other will show that the GM one will attack the biological version, creating at least two problems. On the one hand, farmers willing to preserve the biological products will have to be guaranteed a sufficiently large buffer zone, but where would this land be? On whose field? On the other hand, this awareness should make us question more broadly what are the risks related to GMO: will we eventually lose all non-GM food?

Transgenics as chimeras

The combination of DNA of two different species, could also result into chimeras of course. Though not yet accepted by most of the scientific community – and politico-legal systems – chimeras have been part of the Western culture (as well as other cultures with other names but same features) for a long time.

As a sort of extension of the Posthumanist ideology we referred to in the chapter on enhancement, human-animal chimeras are seen by their supporters as a way forward for humanity, as a possibility to enhance our evolution with genetic traits the we do not (yet?) possess.

Should this scenario become reality though, some questions concerning the outcomes would become apparent. For ex-

ample, chimeras could also be *mostly* animal but with human DNA, should this new non-fully-human animals be granted special protections and rights? Could we eat these chimeras? In which way will this affect our scale of priority in assessing what life is more worthy?

Xenotransplantion

A sort of functional chimeras could be represented by those individuals using organs or living tissues from other animals. This type of operation is called xenotransplantation, and – due to the shortage of human hearts and kidneys for transplants – it is gained popularity as a feasible way to contrast the shortage of organs we face regularly.

Among other animals, pigs have been found to be particularly ideal as candidate donors as they have a very similar physiology (recent advancements in gene editing techniques appear to have increased the level of compatibility even further) and organ size.

Hence, this might seem as an uncontroversial issue perhaps, but some questions can be raised here too. Should we prioritize our anthropocentrism (kill a young pig to save an elderly human being)? Also important is the social impact that the increased (endless?) availability of organs would entail: will people act less wisely (e.g. binge drinking) on the assumption

that their organs are easily replaceable and therefore in need of less care?

Three Parents Baby

Another technique that has recently been accepted in Mexico and the UK is Mitochondrial Replacement Therapy (MRT). As MRT requires the fusion of the DNA of three parents (although of a minimal percentage in the case of one of the two female genitors) into an embryo, it has been often – perhaps questionably – referred to as the way to a *three parents baby*.

The energy needed by any cell in an organism to function is provided by structures contained in the fluid surrounding the cell nucleus. Cells can have different numbers of mitochondria, but each will contain specific sequences of mitochondrial DNA (mtDNA), that comprises of 37 genes – each of which purely aimed at maintaining mitochondrial function. In fact, more than 99% of the DNA of a cell is encountered in the nucleus. Nuclear DNA (nDNA) contains over 20,000 genes, with at least 1,100 of them playing an active role in mitochondria. Mitochondrial diseases (most of which are extremely serious and life-threatening) can be caused when mutations occur in either mtDNA or nDNA. However, while nDNA is inherited from both parents, mtDNA is only heritable from the female genitor, therefore any kind of mutation present in this woman's mtDNA might be inherited by her

child. Hence, it is understandable if a woman suffering from mitochondrial disease – or discovers to be a carrier – would want to remove her mtDNA if possible.

Yet, questions abound in this scenario as well. Some concerning the social construction of parenting, others the limits and obligations towards having a genetic link with our offspring (in the face of adoption for example, that is drastically diminishing in rich countries).

Life extension

The possibility to intervene directly on our cells and DNA and hence *adjust* our body at a micro level unthinkable until just a few decades back, has opened the door to opportunities to tackle problems such as that of aging. Aubrey de Grey for example, defines death as an illness and himself as a "crusader against aging" (See related videos).

Although the idea of extending our lives (perhaps even endlessly) might appear tempting to many at first glance, there are plenty problems related to this way of conceptualizing aging and death. First of all, in a more existential sense, this view could lead us to take value away from considering some aspects of life assessed as particularly important because they are finite. Shifting our approach on the matter would have a larger impact on how and what we consider worth of pursuing in life.

Secondly, such an approach would have a huge impact on the issue of overpopulation that seems culpably undervalued. Should we all live longer (say 300 years old for example), how much bigger would our carbon footprint be? And more crucially perhaps, who will decide who is going to access this extended life, will it be fairly spread across the globe or will we face situations where wealthy individuals will live five, six times longer than other poorer fellow human beings?

Geoengineering

Aside from some sporadic exceptions, global warming has been one of the most discussed issues by the international community in the past two decades, and with that the prospect of greenhouse gas emissions reduction. However, as too often is the case when it comes down to coordinating human beings across the globe towards a praiseworthy course of action, the plan has not been successful in its implementation – and we are now facing a situation that appears to be bound to only get worse.

In this state of affairs, the idea of geoengineering began to gain visibility and approval in the last years, possibly also because of impressive advancements in other fields (e.g. medicine) thanks to technology – hence providing hope that technology could be the solution in this context as well.

Scientists have started to really engage with this vision and we now have many enterprises that are actively looking into this option, that also encompasses *green energy* and *green economy*. From sponges that can clean the sea from oil, to solar panels with increased power, there are more and more ways to limit our impact on this planet and make our carbon footprints smaller.

Yet, such examples are not sufficient for more drastic changes and that is, what geoengineering really is about. Let us focus on two specific techniques of this sort.

The first technique we should look at is the one that is most widely discussed among the circles of geoengineers and that is considered to have the best chances of success if implemented: the technology of spraying sulphate particles into the stratosphere – also known under the definition of *Solar Radiation Management* (SRM).

Here the idea is the following: by spraying sulphates particles into the stratosphere, we ensure that those very particles will reflect part of the solar radiation back into space, allowing for our atmosphere to cool off a little.

The other geoengineering technology that we should take into account is the so called *Carbon Dioxide Removal* (CDR).

This technique instead aims at removing greenhouse gases from the atmosphere through a number of means (from di-

rect air capture to ocean fertilization by growing specific seaweeds). It should be stressed that the big difference between CDR and other forms of prevention against pollution (such as removing CO2 from the emissions) is that here we can tackle also the carbon dioxide that is already in the atmosphere.

Thanks to this innovation, we can conclude this short journey on two positive notes: not only links between technology and responsible, ethical approaches abound and are necessary, but we might also have started to understand how to use those to limit ourselves instead of constantly chasing more – even when not needed.

Further readings

Brennan, A. & Y. S. Lo. 2016. Environmental Ethics *The Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/entries/ethics-environmental/

Corner, A. & N. Pidgeon. 2014. Geoengineering, climate change scepticism and the "moral hazard" argument: an experimental study of UK public perceptions *Phil. Trans. R. Soc.* A 372: 20140063 https://doi.org/10.1098/rsta.2014.0063

Garasic, M. D. & D. Sperling. 2015. Mitochondrial replacement therapy and parenthood *Global Bioethics* (26(3-4),

pp. 198-205. https://doi.org/10.1080/11287462.2015.1066082

Palacios-González, C. 2015. Human dignity and the creation of human–nonhuman chimeras *Med Health Care* Philos 18(4), pp. 487–499. https://doi.org/10.1007/s11019-015-9644-7

Pasztor, J. & S. Sy. 2016. The Ethics and Governance of Geoengineering *Ethics Matter* https://www.carnegiecouncil.org/studio/multimedia/20161212-the-ethics-and-governance-of-geoengineering

Related videos

Mirko Garasic - Engineering and Environment
https://www.youtube.com/watch?v=QyhFV27g4IM

Mirko Garasic - Environmental ethics
https://www.youtube.com/watch?v=5xBzpL_UGas

Aubrey de Grey - Life extension
https://www.ted.com/talks/aubrey_de_grey_says_we_can_avoid_aging

Playing God With the Planet: The Ethics & Politics of Geoengineering
https://www.youtube.com/watch?v=sTwqJtPFizM

Related discussion

See also the discussion related to this chapter at https://community.tuhh.de/t/3-chapter-what-kind-of-duties-do-we-have-towards-nature-and-why/547

4 Neuroscience and the Law

What do neuroscience and the law have in common, and why should we consider such a relationship so specific to grant it a whole chapter? The answer is probably to be seen in the fact that their relationship – with its limits and ongoing changes – is perfectly in line with trajectory of this module, with the willing spirit to do more and understand better how technology and ethics could and should interact in the future.

Two major points make the connection between the two fields inescapable. On the one hand, lawyers will always seek to best represent their clients' interests. Hence, any technological advancement that could help them support their arguments or provide proof against an accusation against their clients can only be seen positively.

On the other hand, legal systems tend to weigh evidence about why and how person X behaved in a certain given way when attempting to regulate society and apply a coherent code of justice.

Yet, the application of neuroscience to real case scenarios gives or will give raise to a number of questions that need to be analyzed with attention, as their impact could lead to a drastic reshape of our society and our way of conceptualizing law and punishment. We shall look into some of those instances in what follows.

Free will

The concept of free will is fundamental in a retributivist legal system, as the possibility of having the choice to act otherwise is part of the assessment of the gravity of the crime in certain instances, and the evaluation of the punishment is defined to be backward looking – meaning to be taking into account the specific variables of that given individual for having broken the law.

Through a number of experiments (most notably Benjamin Libet's[1]), in recent years, some neuroscientists and philosophers have questioned the legitimacy of this *illusion*, affirming that a forward looking punishing system – a purely consequentialist one – would guarantee much better outcomes for society both socially and politically.

Aside from the usual risks associated with some perhaps overoptimistic ways of portraying the results in our hands, one thing appears clear and beyond doubt: should authorities officially accept the absence of free will, we would most probably enter an era of chaos as no one will feel guilty or culpable for their actions and their unethical behavior any longer as we will all feel justified by the fact that the crime was bound to happen anyway and we could not have acted otherwise. Perhaps then, the only way to avoid such a dystopian future from occurring would be to

[1] https://www.youtube.com/watch?v=OjCt-L0Ph5o

morally enhance ourselves (not surprisingly, as mentioned in the chapter on enhancement, Posthumanists do share a consequentialist approach to life) prior to the official acceptance of the absence of free will?

Though not strictly concerning criminal law, it is certainly interesting to consider what theological implications such a revolution would have. After all, most religions in the world believe that each one of us should make an active effort to follow a certain – often challenging – path towards moral nobility. Should we put ourselves in a condition where our freedom to act otherwise would disappear, in which way could we prove our commitment to such a mission? And – probably more importantly – how could we continue to make sense of such religions that put us (single, free, autonomous individuals so central the specific ramification of a certain theology) in the hands of already given events and choices? How could we be unlawful (morally speaking) sinners, if there would be no real sin to commit?

Reestablishing competence to ensure suffering

In October 2003 the Supreme Court of the United States allowed Arkansas officials to force Charles Laverne Singleton, a schizophrenic prisoner convicted of murder, to take drugs that would render him sane enough to be executed. On Jan-

uary 6 2004 he was killed by lethal injection, raising many ethical questions.

Among others, the most relevant one refers to the fact that he was forced to take some psychotropic drugs that would have allowed him to experience the execution with full competence.[2] This might as well be an interpretation of the law that is sound (after all he was a fully competent individual at both the time of the crime and the time of the sentence), but this interchange of knowledge and vision between law and neuroscience is certainly controversial.

Why do we need him to suffer competently from his past actions? If that is the argument used to justify the enforced medication, should we expect an increase in case to case sentences based on the neurological map and history of each criminal? Neuroscience has already been used to reduce the sentence in many instances, perhaps in the future it will be used directly to shape the sentence.

Social duties and enhanced responsibilities

With the intention of providing useful groundwork for public policy, the challenge that performance enhancing drugs

[2] Garasic, M. D. 2013. The Singleton case: enforcing medical treatment to put a person to death Medicine, Health Care and Philosophy 16(4), pp. 795-806. https://doi.org/10.1007/s11019-013-9462-8

(PEDs) could pose to our concept of responsibility has recently been discussed by various scholars.

Particular attention has been given to the possible rise of new duties related to the sensitivity of certain professions and the expected socially beneficial impact that some PEDs could have in such contexts. By focusing on certain professions such as surgeons or pilots, some have adopted a *socially responsible innovation* approach that supports the idea that, in the future, new duties may be expected for some professionals.

Within a more legally oriented and specific framework, some scholars have speculated over the legal obligations that surgeons might be expected to comply with. Their conclusion is that, at the moment, the use of PEDs cannot be imposed on surgeons because of the uncertain levels of safety related to their use. Should we manage to achieve PEDs with no side effects (as the IDF study tried to do) however, the legal scenario might change. This approach, cannot push to question what such a change would imply for society at large. Will we eventually expect everyone to hyper-perform? To what end?

Thou Shalt Not Lie

Although other methods for extracting information from individuals (for example through the use of polygraphs) have been questioned, this has not stopped the neurobehavioral

scientific community from continuing being engaged with this endeavor. Even if fascinating in many respects, the actual effect that such findings could imply for society should not go unquestioned.

One historic urge that scientist have had is that of discovering a reliable way to assess whether or not a person is lying. Aside from the fact that – as in other contexts where neuroscience is called into action – the interpretation of scientific data will require training and responsible behavior (the fact that I might be lying might not clearly state on what I am lying about, nor perhaps the intensity of my lie). Should we, as a society, not sufficiently stress this passage in the implementation of new technologies, we risk to create a dysfunctional future as a result of our misrepresentation of data.

In legal terms, the achievement of such technique would have direct consequences on some of our – normally guaranteed – rights. Most notably, how would such a scenario see our right to silence (or Fifth Amendment in the US context)? Will we be forced to undergo an interrogation while scrutinized by machines that will not allow us to remain silent to certain question? Or perhaps even extract information from our silence?

An even more intrusive way of dealing with this passive role of individuals, is represented by memories. MRI scans to extract memories that can help gathering sensitive information

about a case have already been implemented in a number of courts across the globe, and this opens the door for many ethical questions concerning the use of this technology. Should we pose limits to its use?

As the accuracy is constantly improving, soon enough we will be able to extract the face of a third person from the memories of an accused person or prisoner – or even a free person. Even against their will. Should we see this favorably as a way of ensuring truth under all circumstances, or should we grant the *right to cognitive liberty* and its deriving siblings to not enter certain spheres of our brain without explicit permission?

Further readings

De Caro, M. & A. Lavazza. 2014. Free Will as an Illusion: Ethical and Epistemological Consequences of an Alleged Revolutionary Truth *Social Epistemology Review and Reply Collective* 3(12), pp. 40-50. https://social-epistemology.com/2014/11/03/free-will-as-an-illusion-ethical-and-epistemological-consequences-of-an-alleged-revolutionary-truth-mario-de-caro-and-andrea-lavazza/

Ienca, M. & R. Andorno. 2017. Towards new human rights in the age of neuroscience and neurotechnology *Life Sciences, Society and Policy* 13(5) https://doi.org/10.1186/s40504-017-0050-1

Langleben, D. D. & J. Campbell Moriarty. 2013. Using Brain Imaging for Lie Detection: Where Science, Law and Research Policy Collide *Psychol Public Policy Law* 19(2), pp. 222-234. https://dx.doi.org/10.1037/a0028841

Santoni de Sio F., Faulmüller, N. & N. A. Vincent. 2014. How cognitive enhancement can change our duties *Front. Syst. Neurosci.* 8(131) https://doi.org/10.3389/fnsys.2014.00131

Related videos

Mirko Garasic - Neuroscience and the Law
https://www.youtube.com/watch?v=FUZaVVJ_c1w

Related discussion

See also the discussion related to this chapter at

https://community.tuhh.de/t/4-chapter-neuroscience-and-the-law/548

5 From Military Robots to Self-driving Pizza Delivery

This chapter will give you a short introduction to roboethics and address some ethical issues of recent robot technology. The intention here is not to give definitive answers but to present some ethical issues of current robot technology.

> There are a lot of excellent introductions to the field of roboethics. The book 'Robot ethics: the ethical and social implications of robotics' (2012, Eds. Lin, Abney & Bekey) is an excellent introduction to the ethical issues of robotics. Recently, Spyros Tzaphestas (2016) has published an introduction to robot ethics that covers a lot of relevant issues and is accessible to the general public.

Now, without further ado, here is what we will do in this chapter: (1) In the first section, to get things going, we will have a brief cultural-historical look at our obsession with artificial creatures. (2) Then, we will turn to roboethics and what it is concerned about. (3) Next, we will address some ethical issues regarding current robotic technology. Particularly, military robots, companion robots, care robots and self-driving cars.

(1) A little history to begin with

Humans have been obsessed with artificial creatures for a long time now. Just consider Talos, from Greek mythology,

which is a giant bronze automaton that is supposedly good at crushing your enemies (as early as 400BC). Then, of course, there is the Golem of the Jewish tradition, a creature made out of non-organic material, such as clay, and that comes to life through magic. Another example that is closer to robots comes to us from Leonardo da Vinci, who devised a mechanical humanoid knight (around 1490). Our obsession with artificial creatures generally, and mechanical automatons in particular, is nowhere more evident than in movies and literature. To name just two historic examples here: There is E.T.A Hoffman's famous story Der Sandmann (1816) that features an artificial woman named Olimpia and there is the classical movie Metropolis (1927) by Fritz Lang, where the artificial creature Maria stirs unrest. Of course, we could continue this list of examples until we arrive at the latest instalments in pop culture ranging from cute little robots like Wall-E (2008) to cunning murder machines like in the movie Ex Machina (2014). So, taking into account our obsession with artificial creatures, it may not come as a surprise that we are at a stage of technical development where vacuum robots like Roomba clean our apartments, self-driving cars are likely to hit the streets in the near future, and care robots are deployed in hospitals and retirement homes.

If you want to delve deeper into the history of automatons (what we today call robots), Kang in his book 'Sublime Dreams of Living Ma-

chines: The Automaton in the European Imagination' (2011) provides an intellectual history of mechanical beings.

The Czech author Karel Čapek was the first to introduce the term 'robot' in his play Rossum's Universal Robots (1920). Interestingly, in this play the robots are trying to overpower its human masters. This is another example, like Olimpia in Fritz Lang's movie or the humanoid robot in Ex Machina, of both our obsession but also fear of our own creations.

(2) Roboethics

Before we will come to roboethics, a quick word on the classification of robot technology. As one may expect, there are many ways to classify robot technology. Here is one example taken from Kopacek (2013):

For the purpose of the chapter, however, let us use a simpler classification that divides robots into industrial (which we will not address here) and non-industrial robots and then divides these kinds further. Here is a visualization of this simple classification:

Accordingly then, roboethics can be split up into assistive roboethics, military roboethics, and so forth.

Now, what is roboethics? To answer this, we will first take a look at ethics generally and then shift to roboethics. Although in ordinary contexts ethics and morality are used interchangeably it is customary, at least in philosophy, to dis-

tinguish the two. Morality refers to the collection of norms and values that people hold, whereas ethics is the investigation and reflection on morality. Simply put, ethics reflects on right and wrong conduct, so please keep in mind that ethics is also concerned with the justification of our conduct. That means giving reasons for or against something. So, for example, when we say "Hitting a child is wrong", we pass a normative judgment. In turn, judging that something is good or bad, right or wrong, is not enough. We also have to provide reasons (that is, a justification) for why we think that this is the right or the wrong conduct.

> If you want to know more about the distinction between morality and ethics, the BBC has a homepage devoted to the question 'What is ethics?'[a]. If you want more in-depth material on ethics and ethical theories please visit the Internet Encyclopedia of Philosophy on ethics[b].
>
> [a] http://www.bbc.co.uk/ethics/introduction/
> [b] https://www.iep.utm.edu/ethics/

Traditionally, ethics is concerned with the proper conduct towards other human beings and towards non-human living beings. However, ethics nowadays also includes reflecting on the right and wrong actions regarding the environment, and recently it has come to include the reflection on how we should treat our robots. We will come back to the behavior

towards our own creation in the last section. For now, let us turn to roboethics.

Generally speaking, roboethics is concerned with the examination and analysis of the ethical issues associated with the design and use of robots. For example, whether some robots should not bee used because they are a threat to human well-being or whether some robots infringe on values and human interests like privacy. And, again, ethical examination here is also concerned with providing reasons for positions and for certain actions. Please note that in this chapter we will focus on robots that possess a certain level of autonomy. Why the talk about autonomy here? Well, admittedly all robots raise some ethical questions but as a general rule the more autonomy the robot has, the more moral sensibility and more scrutinizing is required. So, the focus in in the next section is on robot technology that exhibits a certain amount of autonomy or "intelligence", that is to say they are able to carry out certain tasks without human intervention. Keep in mind that autonomy of robots should not be confused with autonomy in humans, where it usually means to conduct one's life according to one's own reasons.

A short summary of what we have addressed so far. First, we briefly looked at our obsession with artificial creatures and robots. Then, we introduced a simple way of classifying robots. Most importantly, we addressed ethics and roboethics. In the next section, we will look at some

ethical issues that arise in connection to particular robot technologies. Specifically, we concentrate on 4 types of robot technologies: military robots, companion robots, assistive robots and last but not least, autonomous vehicles.

(3) Robots and ethics

The most natural question on a lot of people's mind when it comes to robots is: How do we get robots to behave in a way that we deem appropriate? In his novels, the author Isaac Asimov presents an answer to this question. He puts forth the idea that robots may be programmed to behave according to moral rules or laws. So, for example, the robot could be programmed to do x, but not do y. The rules that he introduced have come to be known as "Asimov's laws of robotics" and they are as follows:

1. First law: A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. Second law: A robot must obey orders given to it by human beings, except where such orders would conflict with the First Law
3. Third law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

4. Law Zero (added later): No robot may harm humanity or, through inaction, allow humanity to come to harm.

Now, on a first glance, the idea to program robots to behave according to a set of rules seems like a very reasonable thing to do. However, there are some well-known problems with this approach (For more on the shortcomings of Asimov's laws and an alternative see Murphy and Woods 2009). Asimov was well aware of these problems and used them as a device to propel the narrative of his science-fiction stories. One problem concerns the vagueness of the terms used in the laws. For example, it is not clear what the term "human" means in the first law, or what "robot" and "doing harm" means precisely. Further, there is the issue of a bloat of rules. That means that the world is a messy place and we need a lot of rules and rules for the exception to the rules in order to address all the circumstances that a robot may find itself in. This, however, seems to be an impossible task. The most obvious problem, though, is that there are a lot of situations where one rule will conflict with another rule. Consider the well-known trolley scenario, where an out of control trolley runs along a track on which there are 5 people. Yet, the trolley can be diverted to run along on another track. Unfortunately, there is a person on this other track. So, a decision needs to be made between diverting the trolley to the track where it will run over, and presumably kill, the one person,

or letting the trolley stay on track and letting it run over the group of five. How should a robot in this situation behave, given that it is supposed to save human lives? (For that matter, how are humans supposed to act in such a situation?). Last but not least, another problem is that Asimov's rules may not be feasible in some contexts. There may be contexts in where we expect the robot to harm a human being for example. This brings us to the first robot technology that we will take a closer look at: military robots.

Military robots

Not surprisingly, the military is at the forefront when it comes to robot technology. Military robots are here and they are here to stay. For example, in 2005 the New York Times reported plans of the Pentagon to replace soldiers with robots, and only 5 countries backed a UN resolution to ban killer robots[1]. It is worth pointing out here that fully autonomous weapons already exist (Please recall, that autonomous here means that the robot goes about its task without human intervention). South Korea has an automatic machine gun that can identify and shoot targets without human commands[2]. Another example comes from

[1] https://www.theverge.com/2014/5/16/5724538/what-happened-at-the-un-killer-robot-debate

[2] http://www.dailymail.co.uk/sciencetech/article-2756847/Who-goes-Samsung-reveals-robot-sentry-set-eye-North-Korea.html

Russia, where the military uses autonomous tanks to patrol sensitive areas[3].

Just for the fun of it, here are two more examples: Dubai recently showcased one of its new Robocops that is supposed to patrol the streets in the near future[4]. Russia is developing a Terminator-look-alike that can actually fire a gun! (Or two guns, if necessary)[5].

Now, despite their ability to shoot people and their occasional intimidating looks, using military robots could have some beneficial consequences that may be taken as reasons to ethically justify their deployment. For example, military robots may reduce casualties because you need fewer humans to fight your war. Of course, obviously, this advantage only applies to the side that has military robots. Further, robots are not subject to psychological stress like human beings. Given that a lot of soldiers suffer from PTSD (posttraumatic stress disorder) after returning from the battlefield, it seems to be a good idea to reduce this kind of suffering by using robots instead of humans. Another advantage is that robots do not give in to emotions and rage and, unlike human soldiers, blindly obey the commands given to them.

[3] https://www.newscientist.com/article/mg22229664-400-armed-russian-robocops-to-defend-missile-bases/

[4] http://www.dailymail.co.uk/sciencetech/article-4530260/Real-life-Robocop-starts-work-Dubai-tomorrow-onward.html

[5] https://www.popularmechanics.com/military/research/news/a26140/russia-robot-gunslinger/

Despite these (potential) advantages there are some crucial ethical concerns that need to be addressed: One of the pressing issues is whether military robots should be given the authority to fire at humans without a human in the loop. This is particularly important, because we need to make sure that robots are sufficiently able to distinguish between combatants and civilians. Further, the availability of military robot may decrease the threshold of armed conflicts. After all, if you have a bunch of robots that can fight for you without human losses on your side (!), then the motivation to start an armed conflict may be higher. A related issue is that the potential ease of using robots may foster an attitude that takes military robots to be a "technical fix" to problems, so that other, more peaceful, solutions drop out of sight. Also, there is the question of how responsibility is to be distributed, especially when the military robot harms people that it was not supposed to harm. How do we determine and distribute who is responsible for the behavior of military robots, particularly when they are autonomous? This issue is very complex because we have to take into account the multitude of players that are involved: the creators of the robot (including IT companies that provide the software, and other research institutions), the military (for example the people in the chain of command like commanders and soldiers). Or maybe we can attribute responsibility to the robot itself? Now, it is not surprising that philosophers have a lot to say about this issue. Some authors have argued that it is impossible to at-

tribute responsibility to any of the players when it comes to military robots (Sparrow 2007), whereas other authors have suggested a way of attributing responsibility (e.g., Schulzke 2013).

> Because of the risks and moral dilemmas involved in military robots, some people, including Stephen Hawking and Elon Musk, have called for a ban of 'killer robots'[a][b].
>
> [a] https://www.technologyreview.com/s/539876/military-robots-armed-but-how-dangerous/
> [b] http://hir.harvard.edu/article/?a=14494

Companion robots

After the rather bleak topic of killer machines, let us now turn to more uplifting machines: companion robots. Usually, these robots are set up to allow some kind of interaction, such as speech or gestures. In short, companion robots are robots that, as one would expect from the name, keep people company at home, at work, in hospitals and retirement homes. The classic example here is Paro the fluffy robot seal that can be used in retirement homes to cognitively stimulate people with dementia or calm them down. Two more recent companion robots are Kuri and Buddy. These two are supposed to be all-round companions that can play music, remind people of tasks and duties, and, with the build in camera, you

can send it to specific places in our house to check something out[6].[7]

There are some things that speak in favor or having companion robots. There is some indication that companion robots increase interaction and communication of autistic children (Scassellati, Admoni, Matarić 2012). Also, companion robots may also ameliorate loneliness in some people, especially when they are elderly or socially isolated (Bemelmans et al. 2012). However, the cuteness and cuddliness of companion robots should not blind us to the ethical issues that need to be addressed. One of the problems concerns attachment and deception: Should we really create things that have a high potential for attachment on part of the user but where this attachment ultimately rests on a deception? After all, the robot pretends to be something that he is not: a friend or companion. In other words, do the benefits that a companion robot may bring outweigh the cost that said benefit is achieved by deceiving a human into thinking that he or she has a reciprocal relationship with it? (Sparrow & Sparrow 2006). Another ethically relevant issue is data security because people interact and talk to these companion robots in intimate settings like their home. The information gathered in these interactions should be protected and stored

[6] https://www.wired.com/story/companion-robots-are-here/
[7] https://www.technologyreview.com/s/539356/personal-robots-artificial-friends-with-limited-benefits/

securely, so as not to allow access from unauthorized third parties. Also, it is worthwhile to think about the ownership of the data that are gathered in these intimate contexts. Should the ownership of the data reside with the person that interacts with the companion robot, or is it legitimate that the company that produced these robots has ownership? (A similar concern can be raised regarding other technologies as well. For example, think of devices and services like Amazon's Alexa or Microsoft's Cortana). Another ethical issue concerns the level of authority and autonomy that we give to our companion robots. Should a companion robot that is "tasked" with keeping a young child company be able to intervene, when the child is about to do something that she is not supposed to do; eating candy for example? Some of these ethical issues just addressed also apply to assistive or care robots, to which we will turn next.

Care Robots

Care robots are robots that fulfill crucial tasks in the care for other people, primarily the elderly or bodily disabled. Such tasks may include grasping and lifting objects, or carrying and feeding people. An example for a state of the art care bot is the so-called Care-O-bot developed by the Fraunhofer Institute that is equipped with a tray for bringing things and a tablet interface for displaying websites. Further, the robot

can remind its user to take medicine or call help when the user has fallen and cannot get up[8].

There are clear advantages of care robots. Obviously, they can support elderly and ill people in their home, which increases their independence and quality of life. Care robots could also promote mental welfare in that they may prevent feelings of loneliness. Further, they potentially prevent danger and save lives when they are equipped with the capability to monitor the health and behavior of people. Lastly, the introduction of care robots may be a way to address the so-called care gap in an aging society, in that they take some burden off of care personnel.

However, we should no be so careless as to neglect some crucial ethical issues when it comes to care robots. One of the most pressing issues is the potential conflict between the values of autonomy and freedom of choice on part of the user and the level of interference in the life of the elderly. For example, how persistent should the robot be if a person refuses to take the medicine? Another obvious issue concerns data security. Care robots are used in a sensitive environment and may also have access to medical and other personal data of the owner, so it needs to be ensured that the data is safe and that they do not get into the hands of people that exploit these data. Further, care robots may lead to a decrease in social

[8] http://www.care-o-bot.de/en/care-o-bot-3.html

contact on part of the elderly because relatives may choose to deploy a robot instead of a human caretaker or visit less frequently because grandma has a robot companion. Also, people that are cared for by robots may feel objectified by being handled by a machine. Further, as with companion robots above, the issue of deception lurks. It may be argued that care robots create the illusion of relationship because they "deceive" the user or patient by pretending to be a companion or friend although in reality they do not care. Ultimately, when it comes to care robots, there are also some broader societal issues that we have to take into account. We should ask ourselves in what kind of society we want to live. Do we want to give our most vulnerable members of society over into the care of robots and if so, to what extent exactly? The answer to questions like this should concern everyone and should not be left exclusively to the people that drive technological development. Speaking of driving, the last robot technology that we will have a closer look at is self-driving cars.

Autonomous vehicles

If you follow the media, you will be familiar with both Tesla's and Google's self-driving cars. However, given the price of a Tesla car, maybe a more relatable example is the self-driving pizza car that is being tested in a collaboration between Ford

and the pizza chain Dominos[9]. This is how the self-driving pizza car is supposed to work: You order the pizza and an employee puts the pizza into the self-driving pizza delivery vehicle. Then, the car finds its way to your house autonomously. When the car with the pizza arrives at your place, you take out the pizza and the car drives off to the pizza place again. It is likely that we will actually see self-driving pizza cars in the not far future because other companies have entered the race. Recently, Pizza Hut has teamed up with Toyota to work on its own version of an autonomous pizza delivery vehicle[10].

Having your delicious pizza pie delivered by an autonomous vehicle has some well-known advantages that also apply to self-driving cars in general. Most traffic accidents are due to human error. There are some estimates that self-driving cars could reduce traffic death by 90 percent[11]. Saving lives is valuable, so that speaks in favor of self-driving cars. Also, self-driving cars potentially lead to fewer cars on the road and a better traffic flow because of the potential capability of these cars to connect to each other and communicate traffic data. This will benefit cities, the environment, and individu-

[9] https://medium.com/self-driven/how-pizza-is-helping-us-design-our-self-driving-future-a78720818e99

[10] https://www.eater.com/2018/1/8/16865982/pizza-hut-toyota-self-driving-truck

[11] https://www.theatlantic.com/technology/archive/2015/09/self-driving-cars-could-save-300000-lives-per-decade-in-america/407956/?utm_source=SFTwitter

als because it ultimately means less traffic related pollutants that are one of the culprits in such ailments like asthma.

Nevertheless, despite the advantages of self-driving cars, some ethical issues need to be discussed. Similar to the military robot technology that we have addressed above, there is the issue of responsibility ascription and distribution. Who should we hold responsible when a self-driving car caused an accident? A related issue concerns what kind of decision capabilities we want in a self-driving car. Think about a critical traffic situation, for example a version of the trolley scenario that we have looked at in the section on Asimimov's laws. Imagine there is a group of people ahead, and a choice needs to be made between running over the group of people, steering to the left and running over one person or steering to the right and crashing into a wall, possible injuring the people in the car. Here the question naturally arises, based on which criteria the autonomous car is supposed to decide. One option is to have no decision power in these situations and leave it up to the driver. However, what if the driver is not attentive? Should the car then be allowed to decide on an option? Ultimately, we have to ask ourselves what risk we want to take as a society and whether the benefits of having self-driving cars on the street outweigh the dangers and risks. Another crucial and not to be neglected ethical issue is the potential loss of jobs that comes with self-driving cars. According to the American Trucking Associations

there are 3,5 million truck drivers in the US[12]. You would not need them anymore if trucks could drive autonomously. The same goes for our self-driving pizza delivery vehicle because it eliminates the human element in pizza delivery. In the concluding section, we will see that robots may not only come for our jobs but also for your rights.

Ethical treatment of robots?

Remember, at the beginning of this chapter we said that ethics not only deals with the justifiable conduct regarding other people and non-human animals but that ethics nowadays is also concerned with the right conduct towards artificial products. Consider this example: In October 2017, Saudi Arabia granted citizen rights to the sophisticated humanoid robot called Sophia. This is the first robot to receive citizenship in the world[13]. This incident suggests that we may want to start thinking about how we treat robots and what part they will play in our social world. Should we regard them as persons and grant them rights? After all, we regard companies as persons and grant them certain rights. Further, is it possible to treat robots in an unethical way (e.g., by harming them)? We will likely be confronted with these and similar questions in the future.

[12] http://www.trucking.org/News_and_Information_Reports_Industry_Data.aspx

[13] http://www.hansonrobotics.com/robot/sophia/

Even more so, because robots will likely reach a level of sophistication that will prompt us to rethink what it is that distinguishes us from them. So, we better get a head start in thinking about these issues instead of trying to catch up with the technical development.

Further readings

Bemelmans, R. et al. (2012). Socially Assistive Robots in Elderly Care: A Systematic Review into Effects and Effectiveness, Journal of the American Medical Directors Association , 13 (2), 114 - 120.

Kang, M. (2011). Sublime dreams of living machines: the automaton in the European imagination. Cambridge, Mass: Harvard University Press.

Kopacek, P. (2013). Development trends in robotics, Elektrotechnik und Informationstechnik "e&i", 2, 42-47.

Lin, P., Abney, K., & Bekey, G. A. (2012). Robot ethics: the ethical and social implications of robotics. Cambridge, Mass.: MIT Press.

Murphy, R., & Woods, D. D. (2009). Beyond Asimov: The Three Laws of Responsible Robotics. IEEE Intelligent Systems, 24(4), 14–20. https://doi.org/10.1109/MIS.2009.69

Scassellati, B., Admoni, H., Matarić, M. (2012). Robots for Use in Autism Research, Annual Review of Biomedical Engineering 14 (1), 275-294.

Schulzke, M. (2013). Autonomous Weapons and Distributed Responsibility. Philosophy & Technology, 26(2), 203–219. https://doi.org/10.1007/s13347-012-0089-0

Sparrow, R. (2007), Killer Robots. Journal of Applied Philosophy, 24: 62–77.

Sparrow, R., & Sparrow, L. (2006). In the hands of machines? The future of aged care. Mind and Machine, 16, 141–161.

Tzaphestas, S. G. (2016). Roboethics: a navigating overview. Springer.

Related videos

Dr. Steffen Steinert - Roomba, Drones and Terminator - The ethical implications of robotic technology https://www.youtube.com/watch?v=5tTEEGRAHsI&t=11s

6 Imprint

This work "Ethics and Technology: Some Issues - A Booklet" is part of tekethics, a project of Hamburg Open Online University (HOOU) and Hamburg University of Technology. It was created by Dr. Mirko Garasic, Dr. Steffen Steinert, Dr. Tina Ladwig, Axel Dürkop and Michael Heinemann. The website of the main project **tekethics** can be found at https://tekethics.rz.tuhh.de/

This booklet was generated using a stack of fine free and open source software. We use Atom and Markdown to write the text, Pandoc and LaTeX to render the PDF, and GitLab as well as Docker to build the whole thing continuously online. Thanks to the communities that keep these tools up and running. The booklet is an Open Educational Resource (OER) and can be forked, remixed and redistributed at https://collaborating.tuhh.de/hoou/tekethics-booklet.

You will find a documentation of the production process of this booklet at https://collaborating.tuhh.de/hoou/tekethics-booklet/wikis/How-We-Did-It.

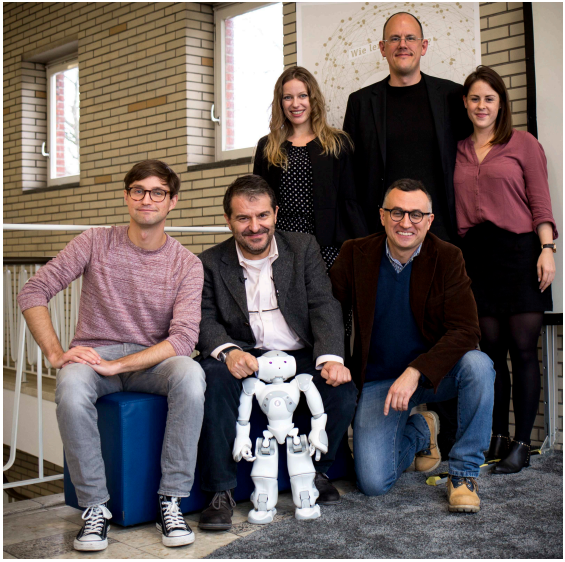


Figure 6.1: The tekethics team. Photo: Stephan Dublasky

Version: 675b96da556cc80e97dfccaeeffd7d309774d544
Job-ID: 53439
Tag: v1.0en