

Robotics and Criminal Law. Negligence, Diffusion of Liability and Electronic Personhood

*Susanne Beck**

“Robot kills worker at Volkswagen plant in Germany“¹ – even the Guardian and other foreign newspapers have reported on this tragic incident which took place two years ago. And although in this case the robot was a traditional one and the incident probably was caused by human error instead of malfunctioning of the machine, it has to be taken into account that such tragic accidents will contribute to the debate about interactions between robots and humans and the legal consequences of damages caused by machines. The phrasing itself is interesting: While in other contexts one would refer to an accident, here most papers talk about the robot as an active participant who “killed” the worker. This indicates an active role of the machine, a different perception of the inclusion of robots, a new fear of society.

Similarly, there has been a vivid discussion after a fatal accident of a Tesla self-driving car.² The unpredictability of the car was reflected as threatening and the debate was slowed down mainly by the realisation that the driver has been inattentive and not followed the instructions given by Tesla. Therefore, Tesla’s self-driving system was cleared by federal auto-safety regulators.³ As the autopilot required the driver’s attention at all times, the driver was regarded as liable for the accident. But still, the debate showed the existing scepticism towards this new technology.

Therefore, it is important to discuss the legal situation right now, before even more major incidents influence the public debate in a negative way

* Prof. Dr. iur. Susanne Beck, LL.M. (LSE), Leibniz Universität Hannover.

1 <http://www.theguardian.com/world/2015/jul/02/robot-kills-worker-at-volkswagen-plant-in-germany>.

2 <https://www.nytimes.com/2016/07/02/business/a-fatality-forces-tesla-to-confront-its-limits.html?action=click&contentCollection=Business%20Day&module=RelatedCoverage®ion=EndOfArticle&pgtype=article>.

3 <https://www.nytimes.com/2017/01/19/business/tesla-model-s-autopilot-fatal-crash.html>.

and hinder the technology unnecessarily. It also is crucial to determine the individual risks of the persons involved to be legally liable, because this might be another hindrance – if the risk of personal liability is too high it might lead to the technology not being developed at all or at least not in a riskful way. New technological developments often challenge society and its normative framework, thus, regulations have to be created to deal with new dangers such as robotics – and it is important to create these laws as soon as possible.⁴ But also, the existing laws have to be applied in a way that reflects the special dangers as well as advantages of these new technologies.

1. The Current Development of Robotics from a Legal Perspective

Robots will, in the years to come, play a more important role in many areas of life. They will work with us, support us when we are sick or immobile, drive us independently and maybe even educate our children, entertain us when we are lonely, give us advice when we are helpless. For performing these tasks, the machines have to become more and more „autonomous“⁵ because it is not possible to give detailed orders for all relevant situations beforehand. Therefore one has to create machines which are able to learn, to adapt (e.g. to the communication style of its user, to his eating habits, to his body functions) and to be trained to react in the best suited way for the user. It is quite possible that for some of these tasks, especially when speed is crucial, the decision of a machine might be quicker, more rational, more informed than a human decision.⁶ In general, machines will decide differently to humans – differently does not neces-

4 For specific regulation issues of robotics, see RoboLaw Group, Guidelines on Regulating Robotics: <http://www.robolaw.eu/>, 2014; see also Leroux et al., Suggestions for a Green Paper on Legal Issues in Robotics, 2012 and the suggestions of the Legal Affairs Committee of the European Parliament, <http://www.europarl.europa.eu/news/en/press-room/20170110IPR57613/robots-legal-affairs-committee-calls-for-eu-wide-rules>.

5 “Autonomous” is used in a broad sense here, meaning a certain space for decision-making for the machine. For a project working on different understandings of autonomy see <http://www.isi.fraunhofer.de/isi-de/v/projekte/WAK-MTI.php>.

6 Arkin, Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative / Reactive Robot Architecture, GIT-GVU-07-11, 2008, p. 2, <https://smartechnology.gatech.edu/jspui/bitstream/1853/22715/1/formalizationv35.pdf>.

sarily mean better or worse, though. In all probability, their decisions will be oriented on rationality and efficiency and less on empathy or emotionality, but, as mentioned, for some tasks this actually might be valuable.

One relevant feature of these new kinds of machines is that when programming the machine one cannot predict how it will act in a specific situation. It will become almost impossible to reconstruct *ex post* why the machine reacted in a specific way.

The common denominator of these kinds of machines is that their function is to disburden humans of making decisions of one kind or another. Humans might only decide beforehand where and when to use autonomous machines and give them guidelines but leave the assessment of the situation to the machines, sometimes even with the ability to learn from former experiences to advance their decision making process. That such developments carry risks and side effects and mistakes in the decision-making-process leading to damages cannot be doubted. As already mentioned, machines will decide differently, and therefore it is even thinkable that machines make „right“ – from the perspective of the machine – but morally questionable decisions. One could even fear the dehumanisation of society in different social concepts when machines take over more and more of our human tasks.

Because of this, it often is postulated that there always, or at least in some contexts, should be a “human in the loop” of the decision making process. Thus, the decision would still be based on human morals, empathy, and potential liability of the human involved. One has to realise, though, that in many situations, this might lead to excessive demand and responsibility of the human in question. The driver of an autonomous vehicle, for example, does need at least 6 seconds to overtake – too long for most traffic situations. In other contexts, the mental influence of the suggestions by the machines is high as well, thus the decision by the human in the loop is determined in a way that might lead to doubts about his responsibility. At least, one should be aware that even when there is a human in the loop one cannot speak of a human decision anymore, but a decision made by human and machine collaboratively.

2. Legal Questions – Overview

The deployment of autonomous machines will lead to questions in different legal areas. The functioning of autonomous machines will require col-

lecting and processing enormous amounts of data – of course, this could collide with the existing data protection regime.⁷ There will be questions not just about the intellectual property rights of programmer, producer and user but also about these rights if the robot produces something by itself. It will have to be discussed – in labour law – if one can be forced to co-work with robots. In international law, one is debating about the legality of Autonomous Weapon Systems.

In the following, I want to focus on the law dealing with the risks of robots instead of the just mentioned, very specific legal questions. Risks can be dealt with in public law (here one can discuss the conditions for usage of these machines and the areas of life in which they can be used), in civil law and in criminal law. Concerning legal handling of risks, the debate resembles other debates in which modern risks are analysed from a legal perspective, e.g. the Internet, Biotechnology, the importance of cooperation and thus the responsibility diffusion in collectives. All this can be summarised as „Risk Society“ – and how this kind of society can be regulated, controlled or governed has been discussed over the last decades.

Adapting the legal system to the development in robotics can mean to enact laws in the area of public law, civil law and criminal law, to interpret existing laws in a specific way, to take non-state regulations into account.

2.1. Public Law: Controlling the Risks

In Public Law one discusses, inter alia, if the administrative laws in certain areas have to be adapted to the usage of autonomous machines – be it the laws about medical devices, the traffic laws, laws about the conduct of research in private or public areas, etc. Additionally, the security standards are introduced by non-governmental institutions such as International Organization for Standardization (ISO). The interaction of state and social norms, of government and governance, is challenged by the development of robotics, because no social standards for adequate behaviour exist until now and one is challenged by having to develop such standards from scratch. While in other areas standards such as ISO-norms are generally well integrated into the legal system (with exceptions, of course), this is

7 Leroux et al., Suggestions for a Green Paper on Legal Issues in Robotics, 2012, pp. 46 et seqq.

sometimes questioned in the area of robotics – it seems to be an area in which the intransparency of these standardizations has become more obvious and problematic, maybe because of the new developments, the missing discussion in society, the impossibility of orientating the standards to other areas of life. Later on, we will come back to the debate that circles around how to develop socially acceptable security standards for such an important, dangerous and unpredictable new technology as the development of “autonomous” machines.

Additionally, robots can be used as assistance against risks, be it in natural disasters, in war or as assistants of police and security organisations. Here one will have to discuss in future in which cases it is proportional to use until now unpredictable and maybe not fully controllable machines. In general: As far as autonomous machines are probably faster and can guarantee better risk management than human assistants, and as far as in these cases one can avoid risking the lives of human helpers, it is more than plausible to use machines.

2.2. *Civil Law: Liability for Damages*

In Civil Law it is, for example, questionable who is liable for contracts closed by autonomous machines or for damages caused by these machines. Differently to traditional machines, it is not plausible to regard electronic agents as mere tool of the user – the decision making range of these machines is too broad for such a categorisation. The existing regimes of liability for damages are also not applicable, at least not without adaptation. With

One also needs to discuss how to deal with the necessary insurances, if it might, e.g., be possible to force insurances to contract with the users of such machines, which categories they belong to, if these machines can only be used if insured, etc.

With regard for some of these problems it also is debated if and how electronic agents can interact as legal actors.⁸ A new legal actor⁹ might be necessary because the traditional liability concepts (e.g. negligence, prod-

8 Hanisch, *Haftung für Automation*, 2010.

9 Leroux et al., *Suggestions for a Green Paper on Legal Issues in Robotics*, 2012, pp. 58 et seqq.

uct liability or strict liability) are difficult to apply in the context of autonomous machines.¹⁰

2.3. *Criminal Law: Responsibility for the Robot's Action?*

Robots will participate in traffic and be used as tools, and in all these contexts it is possible that they will be used to commit crimes – this will lead to specific debates if the action in question fulfils a specific criminal law (traffic laws, trespass, etc.).

But more relevant in criminal law will be the question of criminal responsibility for the robot's damaging a third party. Criminal Law generally is based on the damnable conduct of the offender, on his intent or negligence about the violation of the goods of a third party. All this is challenged by the usage of "autonomous" machines. Even if we are talking merely (for the moment) about machines acting in a dynamic and unstructured environment based on feedback information¹¹, it is almost impossible to pinpoint one individual which is criminally responsible if the machine has violated the rights / goods of a third party.¹²

3. *Focus: Robotics and Criminal Law*

Public law mainly addresses the conditions to balance the interests of the individuals involved or potentially restricted or violated by this new technology, and civil law mainly discusses the contracts and the financial balancing in case of damages. In these areas of law, the main threat for the people developing and producing robots is to be financially liable, a threat that can be dealt with beforehand, by insurance or by collective payments of the parties involved in the development of the technology, for example.

10 Beck, Dealing with the diffusion of legal responsibility: the case of robotics, in: Nida-Rümelin / Bisol (eds.), *Technical Options and Ethical-Legal Responsibility*, 2014, pp. 167 et seqq.

11 Jain, Autonomous weapon systems: new frameworks for individual responsibility, in: Bhuta et al. (eds.), *Autonomous Weapons Systems – Law, Ethics, Policy*, 2016, pp. 303 et seqq.

12 *ibid.*

It is differently in the context of criminal law: Here it is not possible to avoid individual responsibility, on the contrary: Each party is responsible for its own actions.

Several individuals could be considered as perpetrator: the producer, the programmer, the seller or the user of the robot. In most cases the humans involved will not have intent about the specific action of the machine. Still, the violation of another human being could lead to criminal liability arising from negligence. This kind of criminal liability can be connected to every stage of the production process and usage, including research and development.

The first requirement of negligence is that the person whose liability is discussed acted without “reasonable care”¹³. The standard of care is usually determined by a person’s expected form of behaviour in a given situation. As indicators one can refer to non-legal standards, such as ISO and DIN standards.¹⁴ As we have already heard, developing these standards is difficult in the area of robotics. When determining the standard of care for people involved in research and production of robots, there are especially two important things to note:

First of all, at the moment, only few standards exist for the here relevant areas of robotics.¹⁵ One reason for the slow development of standards is that the machines these standards would be relevant to are still in development and the knowledge about possible risks (kind and intensity) is still low. Standardising institutions are challenged not just by determining how to avoid inadequate risks but also by deciding which risks actually are inadequate. In such cases, the general-social standard of rationality is applied additionally: How would a rational person have acted to avoid damage in a similar situation? This vague evaluation, though, offers only little help in complex technological fields such as robotics.¹⁶

Secondly, non-legal norms only are indicators for whether the actions of a person were consistent with the legal standard of care. They also are, generally, developed with regard to civil liability instead of criminal law. Criminal law is not simply an accessory to the regulations of non-govern-

13 Kudlich in: Heintschel-Heinegg B (ed.), Beck-OK StGB, § 15 para 35 et seqq.

14 See, e.g., BGHSt 4, 182 (185); sceptical: Duttge, in: Joecks / Miebach (eds.) *Münchener Kommentar StGB*, § 15 para 114 et seqq.

15 See, e.g.: ISO 10218-1: 2006; ISO 8737: 2011; ISO 10218-2: 2011; ISO 13482: 2014.

16 Duttge, in: Joecks / Miebach (eds.) *Münchener Kommentar StGB*, § 15 para 114.

mental groups, thus one must always additionally consider overall social morality¹⁷. If certain internal rules do not contradict social expectations and standards of rationality, and if any party in question has recognized this deficiency, liability for negligence must be included in the determination of criminal liability.

For considerations on the few already existing standards in robotics as well as on the process of developing such standards it is necessary to consider the two relevant perspectives: The perspective of standardising institutions can probably best be shown by quoting the German DIN-Institute itself (my own translation): "Standards foster global trade and serve rationalisation, securing of quality, protection of society as well as safety and communication. Economic growth is influenced stronger by standards than by patents or licences. Standards are strategic instruments in competition."¹⁸ Even if the protection of society is mentioned, it becomes clear that the standardising actors are also aiming for economic advantages.¹⁹

This has to be contrasted with the perspective of criminal law: Criminal law does not only serve to minimise risks and prevent danger. It also stabilises the normative consciousness of society concerning actions that are regarded as socially inadequate. Thus the danger of a certain action is not sufficient to penalise it; it also is necessary that it violates social-moral rules²⁰. These rules have to be – in theory – accepted by every member of society, which could be an indicator for specific norms based on singular interests (of specific groups) not fitting the criteria for enacting criminal laws. One has to be aware, though, that society accepts – and actually

17 Lackner / Kühl StGB, § 15 para 39.

18 „Normen fördern den weltweiten Handel und dienen der Rationalisierung, der Qualitätssicherung, dem Schutz der Gesellschaft sowie der Sicherheit und Verständigung. Das Wirtschaftswachstum wird durch Normen stärker beeinflusst als durch Patente oder Lizenzen. Normung ist ein strategisches Instrument im Wettbewerb.“ (http://www.din.de/cmd?level=tpl-bereich&languageid=de&cmsareaid=erfolg_durch_normung).

19 According to Gusy, "Antizipierte Sachverständigengutachten" in Verwaltungs- und Verwaltungsgerichtsverfahren. *Natur und Recht* 9 (4) 1987, 164, empirical analysis show the following order: interests of the market leader before others; interests of the industry before others; interests of the providers before interests of the consumers; private interests before public interests; etc.

20 The (criminal) law giver is obviously also influenced by the interest of different lobby groups but still democratically controlled; Burkatzki E (2011) Legalität und Legitimität im Marktkontext. *Zeitschrift für Internationales Strafrecht* 3, 2011, 162.

needs – specific subsystems such as research, economy, and the health system. It would be inconsistent to rely on these systems on one side and not to accept their specific norms which regulate these subsystems and the interests of its parties on the other²¹. Thus the inclusion of economic interests in standardising procedures does not necessarily lead to their irrelevance for criminal law. Obviously, this acceptance has its limits if the values of the subsystem outweigh society's interests, but the turning point for such specialised norms becoming irrelevant for criminal law is difficult to locate.

Another aspect that could help transferring standards into principles relevant for criminal law is the procedure of developing external standards by non-government institutions. As mentioned, standardising institutions often lack democratic legitimation and transparency.²²

Why are these considerations important for robotics?

First of all, because there is a very strong activity of standardising institutions in robotics at the moment, thus it seems, from a legal perspective, important to analyse these activities and relate them with legal evaluation. One might even have to consider interaction with the standardising institutions to secure plausible normative premises and processes.

Secondly, the reliance on these standards is also very high: Most researchers and producers are convinced to have acted legally when complying with the existing standards, even if they are somehow vague, not covering all relevant (dangerous) aspects of their activities and normatively questionable. It is necessary to discuss how to connect this strong conviction, supported not just by the official impression of standardising institutions but by the general custom in the actors community, with negligence liability; it might be worth to consider its relevance for the subjective aspects of negligence (guilt). The (potential) "sense of right and wrong" is part of liability for negligence as well.²³ Unavoidable mistake in the lawfulness of the action can therefore lead to negation of negligence. This is the case especially for the parties not directly involved in and profiting

21 Steinmann, Unternehmensethik und Recht. Zeitschrift für Internationales Strafrecht (3) 2011, 100-109.

22 For an in depth analysis, see: Gusy, "Antizipierte Sachverständigengutachten" in Verwaltungs- und Verwaltungsgerichtsverfahren. *Natur und Recht* 9 (4) 1987, 156-165.

23 OLG Karlsruhe NJW 1967, 2167, 2168; OLG Düsseldorf NJW 1990, 2264 f.; Sternberg-Lieben in Schönke / Schröder (eds.), StGB, § 15 Rn. 193.

from the usage of the robot (researcher, programmer) who are surrounded by a community in which everyone is convinced that fulfilling the requirements of standards is sufficient to act lawfully.

Another condition of criminal negligence is foreseeability of the damage.²⁴ The more autonomous and potentially dangerous a machine is, the more it can be – generally – foreseen during the research phase that it may, later on, bring harm to humans. The usage of robots for military purposes and the usage of autonomous cars in everyday traffic are plausible examples: It almost seems unavoidable that thereby human beings are (for warfare: unjustifiably) violated. On the other hand: The foreseeability is only connected to the general possibility of harming; the specific conditions and situations become more and more unforeseeable²⁵. Robotics is therefore an opportunity to discuss how specific the foreseeability has to be: Does it have to be directed towards specific circumstances, causalities, harms, or is it sufficient to foresee the possibility of violating humans as such?

4. Responsibility – Challenged by Robotics?

The technological development of robotics could even be understood as part of this normative change. Overwhelmed by over complex situations, by everyday life entailing endless risks of damaging third parties, by unforeseeability of already small decisions, we react technologically. In some ways the transfer of responsibility might be the point of these machines: The over-complexity of modern society, in which one has to make numerous decisions every day and knows that many of decisions bear the potential to harm others, leads to building machines not just to decide how to best find our way in traffic or to get our car into a parking spot, not just to remind us about our medicine or buying food – we are building machines to decide about life and death of other human beings. The transfer of the decision only makes sense if the human parties involved are not fully responsible for the decisions. This development has to have consequences for the concept of responsibility as such.

24 Lackner /Kühl StGB, § 15 para 46 et seq.

25 Sternberg-Lieben in Schönke / Schröder (eds.), StGB, § 15 Rn. 125.

The adaptive and learning ability²⁶ of robots necessarily involve a certain degree of unpredictability in their behaviour: Because of the increase of experience made by the robot on its own, the robot's conduct cannot entirely be planned anymore. It also gives more control to the user of the robot than in the case of other products. This leads to the question if every „mistake“ by the robot is necessarily caused by a wrongful act of one of the parties in the legal sense²⁷. If robots with adaptive and learning capabilities are let free to interact with humans in a non-supervised environment, they could react to new inputs received in an unpredictable way. If a robot then causes damage because of these reactions it is hardly plausible that it was caused by a wrongful act of the programmer, producer or even the user²⁸.

As mentioned, in cases in which an autonomous robot makes a mistake and thereby damages a third party the traditional negligence regime is – besides missing standards – already confronted with different parties interacting and their interaction probably resulting in future in a machine that acts partly autonomous and can learn from experience; thus the different causes are difficult to impute to one of the parties.

From a general legal point of view, this conflict can, i.a., be solved in the following ways²⁹:

- One of the human parties is regarded as generally liable, e.g. the user.³⁰
- Only the human party is liable who, provably, made a mistake.

26 Günther et al., Issues of Privacy and Electronic Personhood in Robotics, Proceedings of 2012 IEEE International Symposium on Robot and Human Interactive Communication, 2012, pp. 815-820.

27 Boscarato, Who is responsible for a robot's actions? In: B van der Berg, L Klamming (eds) Technologies on the stand: Legal and ethical questions in neuroscience and robotics, 2011, pp. 383-402.

28 Generally about these problems see Leroux et al., Suggestion for a green paper on legal issues in robotics, euRobotics, The European Robotics Coordination Action, 2012.

29 See also Beck, Dealing with the diffusion of legal responsibility: the case of robotics, in: Nida-Rümelin / Bisol (eds.), Technical Options and Ethical-Legal Responsibility, 2014, 167 et seqq.

30 This is how the law handles, at the moment, park distance control systems; Amtsgericht München, Urteil vom 19.7.2007 – Az.: 275 C 15658/07, NJW RR 2008, 40.

- All human parties “behind” the robot can be transformed to a new legal entity.³¹
- One could even, e.g. for social useful robots, transfer the damages onto society itself.

All of these solutions are based on premises about who is profiting from the usage of robots, who should be “punished” financially for its mistake, who is thought to be in control or stay in control even if the machine overtakes some of the originally human decision-making.

But risks in the context of robotics do not only include damages or mistakes: There are also risks of unwanted side effects. Of course, every new technology is accompanied by discussions about slippery slopes. That this is intensively debated in the case of robotics is not surprising: The imagination of robots nursing the elderly or baby-sitting, taking over our everyday communication, giving psychological advice or waging our wars obviously threatens our accustomed perception of the “social”³². The probability of change does not necessarily imply that a development has to be restricted or even forbidden. When discussing robotics one has to be aware of the responsibility for these potential side effects, though.³³

This view onto responsibility problems robotics shall be completed by the already mentioned aspect of the responsibility transfer onto machines³⁴. Machines overtaking responsibilities even on the stage of decision making, can, as mentioned, be characterised as technological reaction to the over-complexity of modern society. Behind this development I suspect, besides the hope that machines by having more information and reacting faster than human beings might make less mistakes, the need to hand over these decisions because we feel overwhelmed by the responsibility for them. But this development leads to new questions: Who, then, is the responding entity? Can the machine respond in a way that is neces-

31 This obviously does not, by itself, solve all problems and not necessarily exclude the other solutions but gives the third party a kind of addressee, at least for its financial claims. Wettig / Zehendner, *The electronic agent: a legal personality under German Law*, Workshop on the Law and Electronic Agent 2003, p. 9.

32 Fitzi, *Roboter als 'legale Personen' mit begrenzter Haftung. Eine soziologische Sicht*, in: Hilgendorf/ Günther (eds.), *Robotik und Gesetzgebung*, 2013, 377-398.

33 Beck, *Dealing with the diffusion of legal responsibility: the case of robotics*, in: Nida-Rümelin / Bisol (eds.), *Technical Options and Ethical-Legal Responsibility*, 2014, 167 et seqq.

34 *ibid.*

sary for the social and legal construct of responsibility? As we have seen, without such response it will be difficult to establish a new normative structure that will be fully accepted by society.

5. *Potential Legal Solutions and Their Consequences for Concepts*

General Adaptations in the Risk Society could be to less focus on foreseeability, but more on social adequacy of the action as such, to focus less on external regulations for developing the “standard of care” in a certain area of life but on the legal construction of “admissible risk”. This means to negotiate in each area of life if and under which conditions the usage of robots is regarded as such “admissible risk” and if one does act in the adequate framework, one cannot be responsible for the consequences hereof. This also leads to a restriction of the usage of criminal law. In general, it is plausible to not use criminal law too strictly in cases of modern technologies having potential side effects. One also has to be aware that the individual who would be criminally responsible could be heavily overburdened by it. One could pick the driver, the doctor, the user – but he might be the one having to use the autonomous machine (because, for example, his job as taxi driver depending on it), not understanding it sufficiently and being determined in his situative decision, e.g. in traffic, because he could not be as concentrated as if driving himself, and therefore needing more time to react which one does not have in traffic. To be criminally responsible in such situations could be unjust and overburdening.

It also will be necessary to make some adaptations because of the responsibility transfer onto machines. In general, one increasingly focuses on the „principle of reliance“: If different parties cooperate, generally, only the party is criminally liable who provably made a mistake. The other parties can rely on the lawfulness of the other’s actions. It is questionable, though, if this principle can be adapted for the responsibility transfer onto machines. These entities are still, in many ways, unknown, unpredictable and uncontrollable. Thus, arguing that one relies on a specific course of action of these machines is hardly plausible. As we have already discussed, it also is not convincing that the user is fully responsible because of his decision to use the machine, because this would render the machines pointless in many ways.

Another solution which is discussed more and more frequently is to transfer the responsibility onto the machine in the legal sense as well, con-

structing a new legal entity, often called the “electronic person”. From an inner perspective, it does not pose a big problem for the legal system to reduce individual responsibility in the cases of robots making decisions, to create new legal entities with specific legal responsibilities and to support these changes by strengthening of institutional responsibility in the background, because institutions will decide about the direction of robotics – by financing research, giving out licences, insuring under conditions, etc. But one has to be aware, that by constructing machines who make decisions for us, we give away part of our (social) identity – or maybe better, we reconstruct our identity in a way that it includes machines because we have beforehand decided to use them for a specific part of our autonomy-space.

It has to be discussed further if and how machines or human-machine-hybrids can “respond” for mistakes in a socially acceptable way. It socially is necessary that the entity one makes responsible – morally and legally – has to be able to „respond“. This is important not just for the counterpart, the victim, who experiences the human response to its violation and thus might be able to process it in a better way; it also is important for society that there is someone responding to violations normally attributed to humans. This requires, inter alia, some kind of „freedom“ – at least from external force, and other normative attributions constructed on the moral and legal premises of each society.

Thus, before adapting the legal system, one has to consider the (potential) changes of fundamental social concepts such as identity, autonomy and personhood. One has to be aware that changing the legal system has as much interdependence with society as social changes do have concerning the regulating laws. Thus, responsibility in the context of robotics includes caution in constructing new entities and changing our normative concepts; this does, of course, not mean that changes are impossible and should be avoided. But they have to be implemented consciously and in awareness of their consequences.

6. Conclusion: What are we discussing?

Discussing responsibility in the context of robotics means more than distributing the financial risks or creating insurances that cover the usage of robots in different contexts. It means to discuss – including society – in which areas of life the advantages of robots outweigh the risks and how

the usage should be created. Thus it is possible to create a so-called “admissible risk”, allowing certain actions without being responsible for each unwanted consequence and – most importantly – without overburdening powerless individuals who might have to use the machine without having a choice, who are made the “human in the loop” without being able to make meaningful decisions.

It also means to understand what happens if we intentionally hand over decision making onto machines. It means to legally react on changing fundamental concepts and consciously create the space for these changes. Finally, it means to leave room for decisions against machines taking over responsibility in specific contexts and it means to strengthen the awareness of the relevant institutions who will decide about the development of robotics: They do not only decide about the future of one new technology – in my opinion, they decide about the future of our very basic social concepts, of our understanding of ourselves.

