ARTIFICIAL INTELLIGENCE – SCIENCE FICTION AND LEGAL REALITY

Stefan Koos

Professor, Faculty of Economics and Organisational Sciences, University of the German Armed Forces Munich, Germany
Email: stefan.koos@unibw.de

ABSTRACT

Artificial Intelligence (AI) is being integrated into systems influencing decisions, analysing complex situations and conducting processes, which may be dangerous for humans. Yet AI is not strong enough to be able to feel or suffer and maybe it never will. However, the complexity of AI deciding will grow with the technical progress and it should be discussed, how far we are willing to integrate future AI systems into important social processes such as legal decisions or teaching. The article seeks to give an overview over constitutional problems of an integration of strong AI systems into social deciding processes. It analyses the ability of AI to have similar legal positions as natural persons and corporations in the light of the question, whether AI can be an agent for ethic. Furthermore, it presents options of a liability for damages caused by the activity of AI systems.

Keywords: Artificial Intelligence, Human dignity, Personification, Legal Person, Strict liability.

Introduction

Artificial Intelligence is an important topic of the Future Science. The inability of the humanity to understand its own existence and to perceive what we may call ‘the Soul’ leads to a deep fascination for AI as an image of the human being itself. In the US Television science fiction series ‘Star Trek, The Next Generation’ produced in the end-eighties an android named `Data´ which is possessing all logical capacities but lacking a chip for real emotions is seeking acknowledgement as self determined citizen of the ‘Federation of the United Planets’ and indeed wins the case. The philosophical and legal questions raised in this TV series are much more profound than we are used to in entertaining TV series: Will AI ever have the ability to have at least certain fundamental rights? If so, how would we legally and socially integrate AI units into our societies? Will there be a new kind of racism between organic and synthetic persons? However, that remains Science Fiction, as AI so far is not yet any sophisticated as androids in the Science Fiction culture. Recently all AI systems so far belong to the class of ‘weak’ or ‘narrow’ Artificial Intelligence, systems, which are only simulating intelligent human behaviour without any consciousness. ‘Strong’ or ‘general’ AI, which is acting – not only reacting - intelligent, totally autonomous and flexible does not yet exist. Whether it will ever have real empathy, and not only simulating empathy conducted behaviour, is not clear. Popular approaches like the Saudi-Arabian citizenship of the robot Sophia (The Jakarta Post, 2017) are rather marketing than the result of philosophical and legal considerations on AI.

If we ever see strong AI coming into the lives of the people, we should have elaborated an idea how to solve characteristically problems arising from this. Some of those considerations are belonging to rather futuristic topics, others are already now important to decide with the instruments of the recent law systems and again others are important as a basis for the decision, how far we want to accept the future development of AI. Talking about AI nowadays means not talking about knowledge-systems, which use swarm intelligence, such as Siri of Apple or Alexa of Amazon. We are also not talking about software simulating human behaviour using common software algorithms. We are talking about decision-making systems using highly developed computer-systems, which in the future may copy certain organic brain structures with neuronal nets and even may be constructed using organic cells.
coming closer to the status of a brain. Those systems may not only give answers to knowledge questions and filter information to the user, but they will have the ability to learn, to reproduce themselves or program new AI units and to use flexible experience, which lets them make independent decisions or develop complex acting recommendations. In the future, such systems may decide on the fate of humans or weighting life against life. We should ask whether we will ever be able to accept, that human existence is dependent on machine intelligence or to which extent we may accept it, as surely, we will not be able to stop the further development and the use of AI entirely.

Several years ago, a colleague told me in a discussion that in his opinion judges soon would be obsolete, as AI-systems would decide fully autonomous on law cases. They are logical acting, free from emotions and therefore better than human judges. In another conference on the future of higher education, an Australian colleague concluded, that in the future, law at universities may be taught by AI instead of human lecturers. From my point of view this statement shows exactly the problem – rather than the advantage – of the integration of AI into aspects of the social and existential situation of humans as we should be very clear about a crucial question: What does it mean to decide about human fate without emotions and without the ability to commit human faults? This basic problem refers to decisions on the life of citizens (i.e. in the case of the AI judgements). It refers also to the Civil Law liability situation, for example if AI in a car decides between a collision with a wall and a collision with a pedestrian to save the life of the driver.

**AI as Bearer of Ethic and Human Rights?**

From a philosophical point of view in a first step it could be discussed whether AI can act ethical or and morally. Closely connected with this is the legal question whether AI can be itself a bearer of human rights, namely of the human dignity. If we set the postulate, that being able to act morally means having dignity then it makes sense to ask whether a future strong AI can have human rights?1

In the Indonesian `Jakarta Post´ (Qusthan, 2018) a scientist from BINUS Business School in Jakarta recently published an excellent article on this from the perspective of the Philosophical Science and Ethics. He analysed the question whether AI is subject to ethics.2 The author comes to the conclusion that as soon as AI is simulating human intelligence to an extent that it gets a set of values this would mean that its decisions are value-conducted. As AI does not have natural rights but only artificial rights, AI should not be measured by the same ethical principles as human beings. As consequence of this findings AI would be regarded as a moral agent, which is slightly different from humans. If I understand the author right, this could mean the postulate of an ethic sui generis and therefore a splitting of the term of ethic.

I agree with the conclusion that AI, even if it may be developed to an extent that it is able to imitate a certain value calculation, cannot be measured by the same ethical standards as human beings if we see a connection between having human moral and the ability to have human dignity or something comparable to it. AI until now is nothing else than a mere simulation of human thinking. From the point of view of the natural law everything made by human is not human itself and therefore cannot be fully protected itself by human rights. If we talk about the human fundamental rights, the reason for this statement is in the human dignity as legal foundation for other fundamental rights. Human dignity is not necessarily understandable as result of a divine nature of the human being as in the theological Nature Law Theory. In secular constitutional systems, the human dignity generally can be connectable with a metaphysical reason for the constitutional acknowledgement of humanity.3

---

1 The interdependence between having fundamental rights and being able to act morally is asymmetric. It is not necessary to be able to act morally to have fundamental rights but if something can act morally it must be necessarily be protected by fundamental rights.

2 The author seems to understand the term ethic in the sense of moral, which is differing from the terms used in this article.

3 The German Constitution (Grundgesetz, abb. GG; https://www.gesetze-im-internet.de/englisch_gg/ [visited November 7, 2018]) in the preamble mentions God. This is not understood as a decision for a religious character of the German State, see Maunz-Dürig/Herdegen, *Grundgesetz-Kommentar* (83. EL April 2018), Präambel Rn. 33. Some scholars see this in connection with Art 1 GG and the Eternity Clause Art. 79 III GG as reference to the natural law character of the human dignity. However, it is discussed whether the GG is rather following a positivistic concept, see i.e. Maunz-
An entity, which is not a natural person, cannot possess the full canon of fundamental rights, which are finally constituting certain aspects of the fundamental dignity of the human being and it cannot possess human dignity or any analogue right. This does not say that such entities cannot have certain legal protection also based on the constitutional law but if so, this would only be a reflex of a human basis of the entity. Legal Persons such as corporations or foundations in the German law have the fundamental right of free speech and free political opinion. Those entities, however, have at least a human substance through their corporate members and they need those rights to participate as corporate entities in the society as part of their business activity. AI is not comparable to such entities as its acting is not just (collective or deducted) human acting through the organs of the legal entity. It is autonomous acting of the AI, without human decision, just simulating human decisions. This would be a different category of a legal entity. All dogmatic approaches, which exist in the historical legal discussion to explain the ‘personification’ of legal entities, whether they follow the ‘Fiction Theory’ by Friedrich Carl von Savigny or various sociological theories, inspired by the ‘Realist Theory’ by Otto von Gierke, are not claiming anything different from this. They only grant certain Civil Law positions to legal entities for legally pragmatical reasons or as result of the social reality of human groups or the manifest human will of the founder. This legal position as legal entity does not automatically imply moral duties or moral capacity. However, it may lead to a certain ‘legal ethic’ in the sense of responsibility for the acting organs arising from power in the market but in this case, we see human organs of the entity who are actually the agents for this Business Ethic. We would not find anything comparable as agent of the Business Ethic in the AI unless the AI itself, which is not human. Even if one follows modern approaches to acknowledge autonomous rights to animals or the nature, we would find in the nature or natural creatures rather a philosophical basis than in synthetic intelligent entities, at least as long there is no real feeling of the machine.

AI cannot act ‘morally’, so it has no human dignity because of this, if we assume that what can have moral must have human dignity. If we nevertheless concede certain own rights to AI, then we would do it because of pragmatical reasons, namely within the private law. Moral is not the same as Ethic as it is dependant on individual or on group convictions. In pluralistic democratic societies, Moral is not necessarily legitimised. The Moral of one group is not necessarily the Moral of other groups or individuals. Law in democratic societies is ethic – not necessarily moral – if it is the result of a pluralistic decision of the democratic basis of the society (Koos, 2018). It may be possible, that AI is able to act conform to those pluralistic decisions transferred into legal rules. Acting legally compliant to democratic legitimised rules may be part of programming. However, this would not be ‘moral acting’. The AI would only apply the rules, which are ‘loaded with values’ by the legislator. It is not applying the rules comparing them with own moral but it is merely executing; but deciding compliant to the law is more than applying legal rules. It needs interpretation, which is much more than just grammatical. Legal rules may not be applicable due to a teleologic reduction or may be extended for certain, before not foreseeable reasons. It is essential that legal rules, in any case, must not be applied if they are violating human dignity, for example in cases of essential infringements of the principle of proportionality. To measure legal application exceptions for the sake of human dignity it is more necessary than the ability to apply a rule. How authentic the decisions of AI can ever be on the question whether a legal rule is ‘just’, ‘unjust’ or moreover non-conform with basic rules of humanity if there is no real, not only simulating, feeling of the entity, may we call it ‘Soul’?

It may sound paradox that exactly the individual moral in differentiation from legal ethic, which is part of all human acting, must indisputable remain the basis of social, not only legal, decisions if we do not want to leave the constitutional basis of the human dignity as highest and indispensable rule of all democratic acting. We would give up this rule if we systematically allow AI to take final decisions on existential questions of the personal and social existence of human individuals.

Dürig/Herdegen, Grundgesetz-Kommentar (83. EL April 2018), Art. 1 Abs. 2 Rn. 17 ff. Even then, the human dignity is not disposable by discretion of the legislator due to the Eternity Clause and the expressive intangibility of the human dignity.

4 See Art. 10, 71-74 of the Constitution of Ecuador (nature as legal person); however, this is a pure positivistic legal fiction as Art. 10 II says, that the nature is legal subjects for the rights which are given to it by the constitution; see further examples cited by Fischer-Lescano, Natur als Rechtsperson, ZUR 2018, 205 (206).
Civil Law Consequences of AI Decisions

AI as Legal Subject

Which are the consequences of the categorial difference between AI and Human from the Constitutional and Civil Law Perspective? The discussion before belongs to the question of moral or ethic acting. If we take it as possible, that AI can at least have ‘ethical’, though in a rather functional legal sense as acting compliant to (pluralistic legitimized) legal rules without integrating own ‘moral’ valuing, then it has to be decided, whether AI can at least be subject itself to liability if its decisions infringe rights of legal persons. If not, then who is liable for damages caused by AI decisions?

We must conclude first that AI is not able to be ‘guilty´ in the Criminal Law as it cannot be ‘moral´. Criminal Law guilt is a personal guilt concept, which needs the ability to feel (expiation) and to be afraid (individual prevention). AI is not a person with owns autonomous moral control, it is not feeling fear or sorrow and if we ever should come to the point that it can be afraid of being deleted, then the humanity would be in severe danger, as the AI may not longer accept the supremacy of the humans over its existence. Furthermore, personal guilt of AI is not possible, as AI has not the freedom of own will. Its decisions are rational calculated following algorithms; AI is determined.

Whether AI can be liable under Private Law, for example in cases of ‘wrong´ decisions of AI implementing technical systems such as cars, robots or internet system bots, is more difficult to decide. One way for a construction of a liability of AI may be the acceptance of a legal personality or at least a legal subjectivity\(^5\) of AI. This could be the basis for a liability in the tort law. There are two main theories to explain, why legal capable subjects can exist beyond the natural persons, the ‘Fiction Theory’ (Savigny, 1840) by Friedrich Carl von Savigny and several sociological theories, similar to or deduced from the ‘Group Theory´ by Otto von Gierke, such as the Realist Theory. The Fiction Theory states, that something is a legal subject whenever the legislator decides it should be a legal subject. This decision of the legislator leads to a mere fiction of a legal personality (Hassan et al, 2012). It does not need any real or sociological background. Following the ‘Realist’ or ‘Natural-Entity Theory’, certain entities can have real legal personality if they possess an own life and own will. This is namely applicable to partnerships: If a group of natural persons wants to act as a group and if it is recognised as a socially acting group then it has to be recognised as a legal subject by the law.\(^6\) The German Corporate Law follows mostly the Realist Theory and the Group Theory, actualised in the mid of the twentieth century by Hans Flume (1972) and the German Federal High Court applying this theory to partnerships and non-registered associations.\(^7\)

What do those theories of personification mean for the question whether AI can be a legal subject? Under the Fiction Theory, it would be rather easy to argue, that AI can be a legal subject, if the legislator decides so. Under the Realist Theories AI could only be acknowledged as ‘real´ legal subject, if there is any ‘substrate´ for a recognition as an entity with own life and own will analogous to natural persons. Therefore, we come to a similar question as before when we asked, whether AI can have Fundamental Rights. The natural law concept of the realist theories refers to a human substrate of the entity as a

---

\(^5\) In the German Corporation Law exists a differentiation between entities which possess legal personality (\textit{Juristische Personen, i.e. Gesellschaft mit beschränkter Haftung, Aktiengesellschaft and Stiftung}) and entities which have legal capacity but not legal personality (\textit{Rechtsubjekte ohne Rechtspersonlichkeit, i.e. partnerships like Offene Handelsgesellschaft and Kommanditgesellschaft or non-registered associations}). The legal rule giving legal capacity to commercial partnerships (§ 124 I Commercial Code \url{https://www.gesetze-im-internet.de/englisch_hgb/index.html} (visited November 7, 2018)) was acknowledged by the Federal High Court as expression of a general principle, transferring it to other entities beyond the Commercial Code, such as the partnership under the Civil Code (\textit{BGB-Gesellschaft}), see BGH, decision dated January 29, 2001, Case II ZR 331/00 Wafers Ross, BGHZ 146, 341 and \url{http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=en&skid=9e7d52c6c50b10e9f33ab62e6f6ad022&nr=22085&pos=1&anz=2} (visited November 7, 2018). Dogmatic background for this decision is the Group Theory of Otto von Gierke further developed by Hans Flume.

\(^6\) See v.Gierke, \textit{Die Genossenschaftstheorie} (1887) and \textit{Deutsches Privatrecht} I, 470. In this case the legal subjectivity is not as far going as legal personality – in the German law doctrine this is called ‘Teilrechtsfähigkeit´ (partially legal capacity), see fn 5.

\(^7\) See fn. 3.
personification-substrate’. Corporations, associations and partnerships have members and organs which are agents of the own will of the entity. In the case of foundations as legal persons, there are no members but there is the perpetuated will of the founder conducting the foundation constituting the ‘personality’ (Koos, 2002). In the case of AI we do not find any personal substrate: there is no human will of organs and not even a somehow ‘materialised’ will of a creator of the AI. AI is not personal itself, consequently it is difficult to argue, that there is any real basis for the recognition of it as legal subject. The real basis has to be reasoned with a certain similarity to human will and acting which must be more than merely a slight simulation of human behaviour. Narrow AI does not have this similarity at all and it must be concluded, that intelligent animals such as apes or dolphins rather would be candidates for a legal personification as they have emotions and will.

Liability of an AI-Legal Subject

In the case, that AI can be a partially legally capable subject it may be liable in private law similar to corporations. However, this alone would not solve the problem. A claim for damages in tort law would need the test of culpability of the infringer (The German Civil Code § 823 I BGB) and we just stated that AI has not personal guilt in criminal law. However, in certain legal systems the concept of culpability in the private law is different from the concept of guilt in the criminal law. Since civil tort law here has not the purpose to impose expiation to the infringer and only secondary the purpose of a certain prevention but to compensate damages, the private law culpability is a more objectivised concept. It is nevertheless in my opinion impossible to accept the possibility of culpability of AI. An AI acts conform to programming. Learning processes and flexible adapting its behaviour to experiences and collected knowledge may lead to a development of the capacity of AI but this does not mean that AI is negligent if the adapting process had deficits due to i.e. not enough determining factors within the knowledge of the AI. We can say that AI can make mistakes but the mistakes are determined by the nature of the programmed process. There is no free choice for the AI, as it cannot decide free for the wrong way, knowing or having to know but ignoring, that to follow the other right way is its duty. A culpability therefore can only be reasoned analogue to an attribution of the culpability of the organs, which act for the legal person (The German Civil Code § 31 BGB). However, there is no adequate basis for such an analogy, as the AI does not possess any immediate personal decider for its acting.

A second problem to solve is, that an AI legal subject would need own liability assets or a concept of sharing liability assets with other legal entities for example the user of the AI-system. Furthermore there must be some kind of publicity of the existence of the AI legal entity, namely in cases of contracts made by AI. Contracting with AI, which would be a possible consequence of the acknowledgement of legal capacity of AI, bears certain legal risks for contractual partners and it would be an insecurity factor.

Strict Liability

The liability problem could be solved by the legal figure of a strict or objective liability. Here the user of a system or an item, which is potentially dangerous or risky to cause damages, is liable even if he is not culpable. This can be the case i.e. with cars (The German Law § 7 I Straßenverkehrsgesetz StVG), airplanes, industrial facilities (The German Law § 1 Umwelthaftungsgesetz), products (The German Law § 1 Produkthaftungsgesetz) or animals (The German Civil Code § 833 BGB). At least as long as AI is not yet fully autonomous this seems to be a fitting legal instrument. AI is a dangerous system or a risk raising part system in dangerous systems for example if it is a part of a car. Socially the use of AI in cars or other systems seems accepted. Nevertheless, it has to be decided whether the risk of damages arising of the use of the system should be on the society, which accepts or even supports the integration of AI or on the user of the AI or the AI integrating system who has the personal advantage of the use of the system. Given, that the user is able to insure the specific risk of damages there is no reason for

---

8 See the principle of joint ownerships liability (Gesamthandhaftung) for partially legally capable partnerships in the German Law. The partnership as owner of the property is liable with its assets but additionally also the members of the partnership with their private property. This liability is primary, immediate and joint (solidary) in the relation of the partners. The creditor is free to claim in the same time against the partnership and against any partner of the partnership.
imposing the social risk of the use of AI to the society.\textsuperscript{9} Not in all legal systems the rules for strict liability are capable for analogy, in this case they would have to be developed by the legislators. As in the German traffic law with § 7 I Straßenverkehrsgesetz (StVG) exists a legal rule for a strict liability of the registered user of a motor vehicle, damages caused by integrated AI-systems within the car would already \textit{de lege lata} be covered by this rule.\textsuperscript{10}

Another concept \textit{de lege ferenda} may be the development of a new category of a collective liability of the creators of AI, such as producers or especially programmers. This again would need a legislative initiative leading to the duty of constituting a collective liability funding system by the groups subject to collective liability.

\textbf{Social Decisions of AI and the Human Dignity of Persons Submitted to them}

In order to understand the problem of AI decisions for example as legally judging systems replacing human judges in the context with the basic rule of Human Dignity, the classical Moral-Dilemma problem may be helpful. A decision, whether someone should kill his own child to save hundreds of lives of other people in the case of AI would lead or to a non-decision of the AI (if it is programmed to not actively kill people) or to a rational decision equilibrating numbers of saved lives. Both options may be part of legal rules but the missing of feeling, of compassion and of experiencing love will influence the decision, making it necessarily rational and determined. In other words, the decision of the AI will be necessarily ‘soul-less’.

However, to submit the life of human beings under the decision of a machine, which is not in an authentic way able to feel the dilemma between the saving of beloved persons or of suffering individuals on one side and of a number of strangers on the other side may seem on first view not the problem in question. Socially we should prefer decisions of neutral courts rather than self-justice decisions of involved people of a criminal case. Personal nearness to victims or delinquents can be an obstacle to just legal decisions. However, exactly the missing of the ability to feel the dilemma between the killing of a child and the death of several adult persons, the lack of freedom to choose beyond rational considerations – the Determinism of the Machine - makes the categorial difference between the AI-decider and the human decide. The respective decision of the human finding himself within the dilemma can fall to one side or another side, there is no ‘right’ or ‘wrong’ within the dilemma, but this openness of the decision is originally ‘human’ and to be subject under this openness is the right of the human being as member of the society. Human Dignity means also not to be subject to logical or calculating decisions of intelligent, ‘soul-less’, machines. It means exactly to be subject to those in a certain way unsure, sometimes faulty human decisions, may they be in favour or not in favour to the individual. AI cannot replace human judges without infringing the supreme rule of the intangibility of Human Dignity.

\textbf{Conclusion}

The postulate of a separate ethic (in the sense of a morally flexible acting) without having developed practicable legal concepts on liability, responsibilities of creators and users and the scope of the acceptance of the integration of AI-systems in state and society seems socially dangerous. Without analysis of the human rights implications of the use of AI, this could open the door to a systemic misuse of AI for human rights infringements as those violations can be always declared as committed under the ‘separate ethic’ of the machine system. Only as long as AI stays an instrument or tool without being considered an agent of ethic \textit{sui generis} the responsibility for damages by the use of it remains with humans. This is necessary for serveral reasons, especially because for the time being there is no social

\textsuperscript{9} A difference to the strict liability for damages caused by other items such as industrial facilities, animals or products is that a user of AI may not be able to control the behaviour of the system and therefor has very restricted means to limit the likeliness of damages by the AI.

\textsuperscript{10} However, the registered user can argue that the damage was caused by \textit{force majeure} (höhere Gewalt), see § 7 II StVG. Whether the acting of the AI can be considered as being \textit{force majeure}, is certainly questionable.
concept for handling autonomous acting of machines independent of human responsibility. In the Science Fiction of Isaac Asimov we find the ‘Robot Rules’, which are indispensable connected to the supreme rule of ‘Robot must not harm Human’. The nearer we come to the development of general AI the more an acceptance of AI as autonomous moral agent would necessarily require a socially accepted basic rule system for the acting of machine intelligence in the society. At this moment, this is still pure Science Fiction. The solution to the use of weak AI in the recent reality of the society lies in an intelligent use of the civil tort law and in the development of liability asset principles.

Acknowledgement

The article is a revised version of a paper presented at Syariah and Law Public Lecture organised by the Faculty of Syariah and Law, Universiti Sains Islam Malaysia (USIM) on the 25th October 2018.

REFERENCES:


BIOGRAPHY:

Prof. Dr. jur. Stefan Koos is a University Professor for Private Law and Private Business Law at the University of the German Armed Forces in Munich. He is teaching as guest professor at several Indonesian Universities such as Universitas Airlangga (Surabaya), Universitas Islam Riau (Pekanbaru) and Universitas Bina Nusantara BINUS (Jakarta). His research areas include Unfair Competition Law, Consumer Protection Law, Intellectual Property Law and Cyber Law.