Ömer Oktay Gültekin

ID: 21901413

Section: 18

SHOULD PEOPLE TREAT AIs (ROBOTS) LIKE HUMANS ON LEGAL TERMS?

The term 'AI' has been used to refer to a machine that can succeed in a task that needs human cognition (Abbot and Sarch 329). The prevalent existence of these intelligent machines has been increasing each day. As a result of this increase, people began to recognize the existence of AI and give them rights. For example, robot Sophia became the first robot with more rights regarding what it can do without a male companion than Saudi women after getting citizenship from Saudi Arabia in 2017. With the beginning of granting rights to AIs, some people began to question whether AIs should be held responsible for their mistakes and claimed that people could treat AIs like humans on legal terms. However, people should not treat AIs like humans on legal terms considering the following points: most AI will not reach the level of intelligence required to justify their crimes at court; unlike the crimes committed by a person, criminal liability for AI crimes should be distributed to the user, developer and AI itself; since the penalties imposed on humans cannot restrict the freedom of AI, legal authorities should give distinct penalties to individuals and AI.

One reason why AI should not be equal in legality with humans is that most AIs will not have the intelligence needed to defend their actions. Keeping AI legally equal to humans may not only require holding AI accountable for their actions, but also giving them the legal rights of people, such as the right to self-defense and the right to object to the punishment imposed. However, most AIs may not fulfill some prerequisite of these rights, such as explaining the reasons for their actions. According to Abbott and Sarch, three possible characteristics of domain-based AI are acting uncertainly, mysteriously, and independently (330-331). Therefore, people often see AI as a 'black box'. In this respect, the unpredictability and inexplicability of the actions of AI may make it challenging to give AI the legal rights of people. For example, due to the complexity of the causality of AI's behavior, it would be meaningless for the judge to demand an explanation from AI to justify its actions; thus, the judge may give the punishment without considering the rights of self-defense of AI, which would be against human rights when applied to people. In the hypothetical situation, AI can have people's legal rights, such as self-defense, if it can justify its past actions and explain its future actions. Turing proposes a test for determining whether particular AI indistinguishable from people in terms of thinking ability. The test contains three participants: AI, human, and the interrogator, and the interrogator striving to distinguish who is human by examining the written responses of the other two participants to the questions (Turing 433-434). The significance of this test is that it reveals the equal level of thinking skills of humans and AI, without prejudices such as physical appearance. Therefore, some people can argue that the Turing test may be used to determine whether AI can explain itself and, thus, be an indicator for giving AIs human legal rights such as self-defense. Their claim seems reasonable as thinking ability is the core requirement for legal personhood (Van den Hoven van Genderen 36-37). However, this argument fails to consider that such human-level intelligent AI may not exist in practical terms. In 1950, Turing claimed that AI with '125 MB' of memory could pass this test in fifty years (442). Passing seventy-one years, up to now, no AI has passed the test. Considering that today's AI has much more memory than he predicted and nevertheless still not passed the test, the human brain may not be simple to imitate by humans. Therefore, it seems unlikely that all AIs today and most future AIs have human-level intelligence. Just as genius people have the same legal rights as ordinary people, the legal rights of AI should be set by assuming that they have no intellectual skills to justify their actions since today and in the future, the majority of AIs lack human-level intelligence. Thus, AIs should not be treated the same as humans on legal terms since they could not obtain the legal rights of people.

Another significant reason why AI should not be treated the same with people legally is that unlike the crimes committed by individuals, there may be many criminals in the background of AI crimes, and criminal responsibility should be distributed according to their level of guilt. In general, most people who commit crimes assume responsibility for their actions, whereas, in the actions of AI, all responsibility may not always belong to AI. As Hallevy criticizes the criminal liability of AI, he defines two necessary elements of the criminal liability as criminal conduct (actus reus), which means the action itself, and intent (mens rea), which means did someone do the action willingly while aware of the consequences (177). In most human crimes, people fulfill both the actus reus and mens rea requirements for criminal liability because they may intentionally commit a crime by considering what they might gain or lose if their crime is revealed and make their decisions accordingly. On the other hand, animals do not have criminal liability as Hallevy illustrates, a parrot can repeat words but is unable to arrange words to libel; therefore, not fulfilling the mens rea, intention, requirement for the crime (178). In AI crimes, AI may always fulfill the actus reus by committing the crime; however, the mens rea, intention, requirement may be fulfilled by the user, developer, or AI agent. Since AI fulfills the actus reus but does not always fulfill mens rea and the developer or user does not fulfill actus reus but sometimes fulfills the mens rea requirement, legally, none of them may be held responsible for the crime if the AI crimes considered the same with human crimes. Therefore, AI crimes should be examined differently from ordinary crimes. For that, there are three models for criminal liability of AI crimes. The first model, The Perpetration-via-Another Liability model, is the case that AI does not intend to commit a crime and the developer or user has the intention to crime (Hallevy 179). Giving a concrete example, to provide a better service, AI that stores the user's searches and the developer who sells this information to advertisers is an example of such crimes. In this model, AI is only a tool for the crime, so it should not be held guilty. The second model, The Natural-Probable-Consequence Liability Model, is the case that AI has no criminal intention and the developer or the user unintentionally commit a crime (Hallevy 181). An autonomous vehicle that causes the driver's death due to an unexpected situation is an example of such a crime. In that case, AI is not fulfilling the mens rea, intention, requirement; however, the developer or user may partially fulfill the mens rea requirement because they may know in advance that some malfunction can happen. Van den Hoven van Genderen shows a particular point of a Dutch Civil Code as quoting "A party who brings a product to the market of which, at the state of scientific and technical knowledge at the time he put the product into circulation, it was not possible to discover the existence of the defect of the product, will not be responsible for the defect." (48). This quotation emphasizes that whether the crime is caused by negligence or incompetence should be determined. To understand the reason for the crime, Doshi-Velez et al. propose a two-stage technique. The first stage, Local Explanation, is tracking the reason for the specific decision of AI (Doshi-Velez et al. 6-7). Believing that a decrease in the vision capacity of the autonomous vehicle at night is the reason for a particular accident, seeking to prove this claim by conducting tests on the vision capacity of the identical model vehicles is an example of a local explanation. The second stage, Counterfactual Faithfulness, consists of examining the result of the previous stage and determining whether the reason for the crime is found (Doshi-Velez et al. 7). In the example above, if it turns out that autonomous vehicles vision is decreasing at night, the developer may be responsible for the crime, regardless of the system's internal structure since, according to the Dutch civil code, the developer can notice this malfunction during the testing process and decide to fix it or terminate the project. The third model, The Direct Liability Model, is when AI commits a crime intentionally and both the developer and the user are innocent (Hallevy 186). When trading on the internet, self-deciding AI started to drug trade when it noticed that drug trading is profitable is an example of this model. This self-deciding AI should be held responsible because it fulfills actus reus, committing a crime, and mens era, intention, requirements. However, some people claim that people should always be guilty of AI crimes by assigning responsible people for all AI agents in advance (Abbott and Sarch 378-379). What this argument fails to consider is that the companies can make the people they hire accountable legally for AI products by giving these employees big sums of money. If this practice allows companies to get rid of their responsibility, their care when testing AI products may be reduced. Hence, the criminal responsibility for AI crimes should be distributed to the user, developer, and AI according to their level of guilt.

Another main reason why people should not treat AIs like humans on legality is that a new criminal law should be prepared for AI, which is better than the criminal law for humans in terms of restricting the freedom of AI. Some people assert that punishing AI is meaningless since deterring AI is impossible; therefore, punishing AI does not produce any positive outcomes. Abbot and Sarch have stated that "[Punishing AI] will not produce any affirmative harm-reduction benefits because AI is not deterrable." (344). Advocates of this idea argue that the benefits of prison on humans, like causing them to lose their hate or make them more compatible with society, cannot be obtained by punishing AI (Abbot and Sarch 344). Although the core of this claim is valid in terms of AIs do not have feelings that punishments can develop, it suffers from underestimating the scope of the penalty for AI. Penalizing AI can be justified for three reasons. The first reason is because of its societal benefits. On this issue, Abbott and Sarch state, "Expressing condemnation of the harms suffered by the victims of an AI could provide these victims with a sense of satisfaction and vindication." (346). Thus, not penalizing the mistakes made by AI can decrease people's sense of trust in justice. Also, the presence of non-punishable AI causes people to be scared when using them since no one will be responsible for the damage AI may cause. Hence, integrating AI into social life may be hindered. The second reason for the necessity of giving penalties to AI is that corporations are also punished with various penalties, even though they do not have a real personality. Corporation punishments are significant in showing that the legal identity is different from human identity, and legal identity can be applied to different objects. However, deterrence is not applied to every legal identity in the same way. For people, it can be imprisonment, whereas, for corporations, it can be imposing significant fines that eventually reduce the power of the company. Thus, although AI cannot be deterred in the way people perceive, it may be possible to deter AI like corporations by giving them customized penalties. The last reason that penalties for AI are justifiable is the general acceptance of giving punishments. On this issue, Abbot and Sarch indicate that even though AI does not perceive the punishment as offensive, giving them penalties is required, like there is no obstacle to give an imprisonment sentence to a criminal who wants to live in prison (364). Therefore, it is not an unnecessary practice to punish AI. Even though punishing AI can be justifiable, punishments should be revised. Hallevy proposes possible alternatives to punishments imposed on people for AIs. His first suggestion is that instead of a death sentence, a deletion sentence can be given to AI (195-196). This sentence is meaningful since both punishments restrict the right to life of entities. His second suggestion is that the alternative of community service sentence and fine sentence may be made the service offered by AI free for a certain time (197-199). This penalty can be imposed on AI services used worldwide, as it causes loss of money to the owner of AI and causes satisfaction in the public. His last suggestion is that instead of incarceration of AI physically, AI can be shut down for a limited period (197). Considering that incarceration and shutting down punishments restrict freedoms, one may expect that this punishment is reasonable. However, since AIs have no consciousness, putting them out of use may not be different from killing them temporarily. Therefore, a better alternative could be to restrict their freedom by disconnecting them from the internet. Without the internet, their communication with the outside world may be highly minimized. Additionally, one new punishment may be proposed. This new punishment may be changing the place of duty of AI. For example, a judge AI who has been proven to be racist can be deployed where there are no people with whom AI can be racist. This way, effective usage of the AIs may be provided. Thus, the punishment of AI is beneficial in many ways; however, they should be punished by new laws that further restrict their freedom than human penalties.

In conclusion, people should not treat AIs like humans on legal terms because the majority of the AI's will not satisfy the requirements to have a right to defend their crimes at court; the criminal responsibility of AI crimes does not belong only to AI, instead, it should be distributed between the user, developer and AI itself according to their level of intent in crime; AI cannot be equal to humans in terms of punishment, because punishments imposed on humans may not restrict AI's freedom. For AI to be legally equal with humans, it should have rights such as self-defense, full responsibility for its crimes, and receive the same punishments as humans for its crimes. However, AIs cannot be given the right to defend themselves because the one with a human-level causality does not exist now, and it does not seem possible to be in the future due to the difficulty of imitating the human brain. Although AI commits a crime in AI crimes, it may not be the AI that intends to commit the crime, so it may not be suitable to punish only AI in such crimes, and the developer and the user should also be punished according to their intention levels. If AI is found to be guilty, it should be given some punishments limiting its freedom. For example, rather than giving AI physical imprisonment, disconnecting it from the internet would be equivalent to leaving it alone with itself because it would not get new information from the internet. The world is moving into the age of AI, and people may need to remember them as a new species someday. In order not to be caught unprepared when that day comes, people should start drawing the limits of AI's legal responsibilities today.

WORKS CITED

Abbott, Ryan Benjamin, and Alex F. Sarch. “Punishing Artificial Intelligence: Legal Fiction or Science Fiction.” *SSRN Electronic Journal*, 2019, pp. 323–384., doi:10.2139/ssrn.3327485.

Doshi-Velez, Finale, et al. “Accountability of AI Under the Law: The Role of Explanation.” *SSRN Electronic Journal*, 2017, pp. 1–15., doi:10.2139/ssrn.3064761.

Hallevy, Gabriel. ”The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control.” *Akron Intellectual Property Journal*, vol. 4, Issue 2, Article 1, 2010, pp. 171-202. <https://ideaexchange.uakron.edu/akronintellectualproperty/vol4/iss2/1>

Turing, Alan. “I.—Computing Machinery and Intelligence,” *Mind*, vol. 59, Issue 236, October 1950, pp. 433–460. <https://doi.org/10.1093/mind/LIX.236.433>.

Van den Hoven van Genderen, Robert. “Do We Need New Legal Personhood in the Age of Robots and AI?” *Robotics, AI and the Future of Law*, 2018, pp. 15–55., doi:10.1007/978-981-13-2874-9\_2.