THE INFLUENCE OF ARTIFICIAL INTELLIGENCE ON CRIMINAL LIABILITY

Maxim DOBRINOIU[*]

Abstract

The nowadays deployment of Artificial Intelligence (AI) and its expected relatively rapid integration into various instances of the socio-economic or governmental life (e.g. household, health, industry, trade and so on) represent a great development opportunity for every nation, as well as a key element for the evolution of the mankind. The elements of AI have already started to take over certain human-type workouts or tasks, while it will take not so long until they will almost completely replace individuals in performing their jobs, and thus evolve from the status of simple tools to the status of "electronic persons" or even subjects of law. During their interaction with the human-dominated world, the AI-driven entities may either be in compliance or a conflict relationship with the law and the society protected by the law, especially when a loss, a damage or a casualty occurs. The article aims at studying the electronic persons' behavior and pointing out whether would be possible or not to further treat the elements of AI as liable against the law, in general, and criminal law, in particular.

Keywords: *Artificial Intelligence, criminal law, criminal liability, electronic person*.

1. The concept of artificial intelligence and its impact on social life

At the European level, the term "artificial intelligence" (AI) was officially referred to as "systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals".[1]

From its already far deployment in areas like: medicine, transportation, industry, agriculture, military, public order, Cybersecurity, client-interaction, technology research and improvement, Internet of Things – IoT and so on, the AI proved to be "real", to be "live", and to be a significant part of our socio-economic life.

It is worth understanding, in a first phase, what really means both "artificial" and "intelligence". While "artificial" may be regarded as a good "made by people, often as a copy of something natural"[2], "intelligence" has at least the following meanings: "the ability to learn and understand or to deal with new or trying situations", "the skilled use of reason", and "the ability to apply knowledge to manipulate one's environment or to think

[*] Associate Professor, Faculty of Law, "Nicolae Titulescu" University, Bucharest (email: office@e-crime.ro).

[1] See the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, *Artificial Intelligence for Europe*, COM (2018) 237 final, Brussels, 25.04.2018.

[2] The Cambridge Dictionary, available at https://dictionary.cambridge.org/dictionary/english/artificial, accessed 25.04.2019.

abstractly as measured by objective criteria"[3].

Other authors[4] define AI as artificially developed intelligence, which is, to some extent, correct and logic.

It is pretty much obvious that AI was created as an alternative to humans, a crafted machine with embedded learning and analysis capabilities, mastered to comply with real-life situations and to perform, as much as accurately possible, the tasks and works once done by men. Thus, the combined above definitions may conclude that an element of AI could be perceived as a unnatural product designed with human-like form of intelligence.

However, as written in the preamble of the Montreal Declaration for a Responsible Development of Artificial Intelligence (2018), AI poses a major ethical challenge and social risks, with intelligent machines that can restrict the choices of individuals and groups, lower living standards, disrupt the organization of labor and the job market, influence politics, clash with fundamental rights, exacerbate social and economic inequalities, and affect ecosystems, the climate and the environment.[5]

The evolution of AI-type entities (such as robots) conducted in time to the development of autonomous and even cognitive features – such as the ability to learn from experiences and take independent decisions, thus evolving them more and more to agents that interact with their environment and are able to alter it significantly. That's why the European experts came to the conclusion that "*the legal responsibility arising from a robot's harmful action becomes a crucial issue*".[6]

In terms of liability, the same EU legal document (mentioned above) states that "the most autonomous robots are, the less they can be considered simple tools in the hands of other actors (such as the manufacturer, the owner, the user, etc.)" and this, in turn, "makes the ordinary rules of liability insufficient and calls for new rules which focus on how the machine can be held – partly or entirely – responsible for its acts or omissions", while "as a consequence, it becomes more and more urgent to address the fundamental question of whether robots should poses a legal status".[7]

Another interesting point driven to the attention of the EU Parliament's Committee on Legal Affairs is that "robot's autonomy raises the question of their nature in the light of the existing legal categories – of whether they should be regarded as natural persons, animals or objects – or whether a new category should be created, with its own specific features and implications as regards the attributions of rights and duties, including liability".[8]

It seems to be commonly agreed at the European level that "the existing rules of liability cover cases where the cause of the robot's act or omission can be traced back to a specific human agent such as the manufacturer, the owner or the user and

[3] The Merriam-Webster Dictionary, available at https://www.merriam-webster.com/dictionary/intelligence, accessed on 25.04.2019.

[4] Stuart Russell, Peter Norving, *Artificial Intelligence: A Modern Approach*(3rd edn, NJ Prentice Hall 2009) p. 4-5.

[5] https://docs.wixstatic.com/ugd/ebc3a3_c5c1c196fc164756afb92466c081d7ae.pdf.

[6] See the Committee on Legal Affairs Draft Report with recommendations to the Commission on Civil Law Rules on Robotics 2015/2103 (INL) available at http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML%2BCOMPARL%2BPE-582.443%2B01%2BDOC%2BPDF%2BV0//EN.

[7] *Ibidem*, pct S on Liability.

[8] *Ibidem*, pct T on Liability.

where that the agent could have foreseen and avoided the robot's harmful behavior".[9]

Among other significant aspects, the experts calls on the European Commission, when carrying out an impact assessment of its future legislative instrument, to explore the implications of all legal solutions related to the AI entities (robots), by far the most important one being the "creation of a specific legal status for robots, so that at least the most sophisticated autonomous robots could be established as having the **status of electronic persons with specific rights and obligations**, …, and applying electronic personality to cases where robots make smart autonomous decisions or otherwise interact with third parties independently".[10]

Some authors[11] developed a scale of AI, based on different forms of intelligence they poses and the implication of humans, such as: level 1 – AI with human supervision, level 2 – AI with deterministic autonomy, level 3 – machine learning-type AI, and level 4 – multi agents systems AI.

2. Doctrine views on Criminal Liability

A crime is the only legal ground for the criminal liability. For a crime to be indicted to a specific person (individual or legal), certain elements must exist, such as: a legal provision (depicting the offence), the commission of one or several material acts (*actus reus*), the mental state (*mens rea*) of the person charged with that offence, the unjustifiable ground for the person's criminal behavior, and the attribution (one's moral involvement in committing a crime).

In the large majority of the national criminal systems, one of the most important elements of a crime is *mens rea* – the mental element[12] which drives a person to commit a crime or to trespass a legal provision.

As all the legal practitioners know, that guilty mind of a culprit consists of three different forms: the intent (with its sub-categories: direct intent – when the person foresees the result of his actions and pursue that result, and oblique intent – when the person foresees the result of his actions, and, while not pursuing that result, only accepts the occurrence of that result), the guilt (with its sub-categories: recklessness – when the person foresees that a particular result may occur and further acts without taking care whether that result happens or not, and criminal negligence – when the person does not foresee the result of his actions while he could or should have foresee it), and the overt intent.

From the Romanian legislation perspective, the guilt or the moral responsibility (involvement) of the person who commit a crime is a subjective process consisting of two factors: the consciousness and the will.[13]

In what regards the consciousness, the culprit has the representation of his actions, of the conditions he acts in, and of the causal relation between the culprit's action/inaction and the result. In his mind there comes the idea of committing the crime and, furthermore, the deliberation of the reasons why he, however, should commit the crime.

[9] *Ibidem*, pct U on Liability.
[10] *Ibidem*, pct 31, p. 11.
[11] Ikenga Oraegbunam, Uguru Uguru, *Artificial Intelligence Entities and Criminal Liability: A Nigerian Jurisprudential Diagnosis*, in African Journal of Criminal Law and Jurisprudence, no. 2, 2018.
[12] Malice aforethought (US Criminal Code).
[13] V. Pasca, *Criminal Law. General Part.* 2nd edition, Universul Juridic Publishing House, Bucharest, 2014, p. 156.

At the end of this process, the culprit takes the decision to commit the crime.[14]

In what regards the will, the culprit moves from the mental state to the physical state of his actions, thus mobilizing his energies (at his disposal) towards realizing the external behavioral acts. This will comes to be very important, because the person, being in full control of its actions and without any (internal or external) constraints (physical or moral), has a free and unconditioned determination to act in the desired manner, thus to also commit a crime.

These above analyzed factors are entirely acknowledged and fully recognized as being human-related. They are specific to any individual, whose conscience and will are not affected in any way by various forces, and there is no clue that they may be associated with any form of machine, even world class high performance computers, run with the most advanced pieces of software and applications.

3. From "electronic person" to active subject of a crime

The human-level AI seems to be the next generation of AI, capable of performing almost all the intellectual tasks an individual can do, and also to have feelings (worries, angers, happiness or maybe love) and to control them through autonomous human-like behavior. Many[15] believes that it is a question of time until the AI will become a true forms of intelligence (or a human-based or human-type intelligence), replacing human judgement, also think independently and act for itself.

As we all know, nowadays, in law, a *person* is identified as *individual* (human) person and *legal* person, both having certain degrees of liability when involved in any way in the commission of a crime.

Different authors identified some particular aspects that shape the elements of AI, and play a significant role in explaining the difficulties of assessing the criminal liability share between the "synthetic person" and the "natural person". And these are: increasing autonomy[16] (that meaning a decreasing control from humans), unpredictability[17] (meaning AI lacks of cognition may lead to reactions totally different than human like), and unaccountability[18] (while not applied with legal personality, AI elements cannot be held responsible for their harmful actions).

In order to analyze the actual and real involvement of an AI entity in committing a crime, it is first needed to clarify the role of different other actors in the *doing* or *undoing* (action or inaction – meaning *actus reus*)[19]. And here the "user", the "supervisor" and the "producer" of the AI entity have an important role in a respective criminal investigation, as being the humans behind the machine, thus firstly questionable about the conditions the AI entity acted upon, the software they designed and implemented into the machine, and the computer

[14] V. Dobrinoiu and colab., *The New Criminal Code Commented. General Part.*, 3rd edition, Universul Juridic Publishing House, Bucharest, 2016, p. 134.

[15] Karel Nedbálek, *The Future Inclusion of Criminal Liability of the Robots and Artificial Intelligence in the Czech Republic*, Paradigm of Law and Public Administration, Interregional Academy for Personnel Management, Ukraine, 2018, available at http://maup.com.ua/assets/files/expert/1/the-future-inclusion-of-criminal.pdf.

[16] Ryan Calo, Robotics and the Lessons of Cyberlaw, California Law Review, 2015, p. 532.

[17] Sherer M, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies and Strategies (2016)*, Harvard Journal of Law and Technology, p. 353.

[18] Mireille Hildebrandt, *Criminal Liability and "Smart" Environments in R.A. Duff and Stuart P Green (eds) Philosophical Foundations of Criminal Law* (2011), p. 506.

[19] Weaver J F, *Robots Are People Too: How Siri, Google Car and Artificial Intelligence Will Force Us to Change Our Laws* (2014).

instructions they performed on it or even the omission to intervene when they are noticed about the AI element acting wrongfully, harming an individual or damaging goods.

When it comes to autonomous agents or machine learning, the real problem is the way they actually "learn" from the environment or from their own experiences. With little or even no human control of the learning process (in the future), we will have to deal with unpredictable entities, which may turn harmful or at least unlawful in performing their actions.

The doctrine is still reluctant to clearly attribute the responsibility of committing a crime entirely to the AI element, and rather prefers to identify a human being as the offender – the main actor liable (see the "user", the "supervisor" or the "producer" of the AI element).

According to some authors[20], "the harm from the actors' behavior does not occur immediately, but it may occur in the future when the AI acts", while "the launch or use of any AI somewhat presupposes a duty of control and supervision over the AI and its actions".

On the other hand, other authors[21] believe that AI criminal liability requires legal personhood for the AIs, and that would be similar to corporate criminal liability that some legal systems are recognizing. And, therefore, legal personhood for AI is consequently a question whether AIs should have rights and duties in accordance with the law.

Moreover, the general opinion is that, in contrast with corporations, the AI elements should be liable only for their own actions or inactions (behavior), and not for those initially attributed to certain individuals.

There is an idea that a possible solution would be a system enforcing AI criminal liability within a system that accepts only the *actus reus* condition when assessing a crime, but this seems to be unacceptable from the general principles of the criminal law. We agree with the opinion that such a case, when *mens rea* is excluded, would be similar to the involuntary acts that excludes criminal liability at all.

In one of his remarkable articles on this subject, an author[22] envisaged three models of liability concerning the AI entities, that can be considered separately of in conjunction (for better liability solutions): 1) Perpetration-via-Another Liability Model, 2) Natural-Probable-Consequence Liability Model, and 3) Direct Liability Model.

We agree with the author that in the first model, when a crime involves an AI entity, this AI entity should be regarded as "innocent agent" (like in the *longa manus* theory), thus mere an instrument in the commission of that crime, and not an active (principal or secondary) participant. In this case, due to the lack of *mens rea* of the actual perpetrator, the criminal charge will always pursue the producer, the programmer or the end-user of that particular AI entity.

The second model addresses the cases of the "foreseeable offences committed by AI entities", where, in the opinion of the author, the producer or the programmer do

[20] AP Simester, A von Hirsch, *Crimes, Harms, and Wrongs. On the principles of criminalisation*, Hart Publishing, 2011 at https://www.researchgate.net/publication/241643522_AP_Simester_and_Andreas_von_Hirsch_Crime_Harms_and_Wrongs_On_the_Principles_of_Criminalisation.

[21] Ashworth A, *Principles of Criminal Law* (4th edn, OUP 2003) and Mireille Hildebrandt, *Criminal Liability and "Smart" Environments* in the thesis of Matilda Claussen-Karlson, *Artificial Intelligence and the External Element of Crime*, Orebro University, Sweden, 2017.

[22] Gabriel Hallevy, *The Criminal Liability of Artificial Intelligence Entities – From Science Fiction to Legal Social Control*, Akron Intellectual Property Journal no. 4, University of Akron, 2016.

not have any involvement, nor they acknowledge of any offence until this is actually committed by the AI entity they designed, produced or programmed.

In this scenario, we agree that human activity is merely linked to the malfunction of the AI entity in the manner that the producer, the programmer or the user should have thought about (or should have considered the possible consequence of) a crime being committed (in certain circumstances) by that AI entity. Therefore, we support author Gabriel Hallevy that considers the criminal liability of the human factor rather negligence[23], than intention, although there may be situations when the (human) offender foresees the result of its actions (upon AI entity), does not pursue it, while accepting this result to occur one day.

The third model of Gabriel Hallevy focuses on the AI entity itself[24], while considering the direct liability as similar applicable to societal individuals (offenders). While there are argues that AI elements should be put aside of the criminal liability similar to children and mentally ill persons (*doli incapax*), the new technology developments prove that AI entities are able to interpret large amounts of data from its sensors, to make difference between "right" and "wrong", and even to analyze what is "permitted" or "forbidden".

It is still a question whether these capabilities (irrespective they are the result of a good programming or the result of its own learning feature) may be seen as signs of consciousness or internal elements (*mens rea*) needed for the existence of the criminal liability.

If so, we also need to consider the various forms of participation to the crime commission, depending on the relations between the AI entity and the other human perpetrators, and each other's involvement in pursuing the criminal activity. In these scenarios, the AI entity may find itself in the capacity of principal, accessory, accomplice or abettor.

Although some authors believe the contrary, we consider that is beyond reasonable acceptance to consider AI elements as qualifying to all the defenses against criminal liability (e.g. self-defense, necessity, consent, error, physical or mental constraint etc.), due to the fact that, in our opinion, there are more other internal elements to be taken into account when analyzing the possible fulfillment of all the requirements.

Conclusions

Trying to find the best solutions for AI-related legal problems, some authors[25] envisaged various approaches, from a "precautionary" one – in which the autonomous agents are precluded or prohibited due to their associated risks and uncertainties, to a "permissive" one – permitting the deployment and development of AI entities and autonomous agents, while accepting the risks and the social costs until properly regulating the domain.

As revealed by the above analysis, in the crimes committed with the involvement of AI elements, for the criminal liability to exists there is a strong need for both *actus reus* and *mens rea* to exist in the behavior of the respective artificial intelligence agents.

And we observed that at least *mens rea* is hard to be taken into consideration in what regards AI.

But, before "thinking" and "acting", there is a strong need for an AI element to

[23] *Ibidem.*

[24] *Ibidem.*

[25] See Peter M. Asaro, The Liability Problem for Autonomous Artificial Agents, presented to the Association for the Advancement of Artificial Intelligence (www.aaai.org), Spring Symposia 2016.

learn (or to be taught) about the law. Civil and criminal. And if is about a autonomous AI or an advanced machine learning, the producer, the programmer or the user must ensure that the most important routines of instructions comply with the existing laws and regulations, and the entity is (somehow) forced to "learn" the most prevalent principles of the living societies (not to kill, not to harm, not to steal, not to destroy etc.), to abide these laws and regulations and to keep away from any sort of autonomous actions that may be considered as unacceptable harmful behavior.

And this should be the main task of all the future projects involving the development of AI or legal bids to consider (and further treat) AI as "electronic person", with rights and obligations, similar to human beings.

Also, considering that the future will probably belong to the AI elements, the basics of the criminal law must be adjusted according to the principle *nullum crimen sine lege*, assuming that for the new society (electronic) members we may need to create special legal provisions and maybe a new legal system[26].

We share the same views with other authors[27] claiming that the AI entities should be considered as both objects and subjects of legal relations, "perhaps somewhere between legal entities and individuals, combining their individual characteristics with regards to relevant circumstances".

Another system that should be revised in the future is the penalty one, as it is hardly believable that actual criminal sanctions may apply to AI accordingly (such as: imprisonment, penal fine, safety measures or educative measures). There are multiple possibilities to be considered, such as: the destruction, the dismemberment, the decommissioning (partially or totally), the removal from duty or the reprogramming.

In all the cases, we think that there will be no effect on both re-education of the "convicted" AI entity, and the prevention of future crimes – as the principal aims of any penalty system in place, due to the fact that AI existence and behavior rely on computer programs and logic instructions and not on human-like emotions or feelings like shame, fear, care, love, guilt, outrage, regret, suffering, worry, rejection, social connection, need, sense of freedom etc.

For all that, the national criminal justice systems are required to adapt themselves and include clear and comprehensive provisions in order to ensure the public order, the safety of people and their goods and property.

References

- https://www.coe.int/en/web/human-rights-rule-of-law/-/thematic-session-on-artificial-intelligence-and-criminal-law-responsibility.
- https://www.torontopubliclibrary.ca/detail.jsp?Entt=RDM3061779&R=3061779.
- https://www.researchgate.net/publication/321321904_Criminal_liability_of_the_artificial_intelligence/download.
- https://arxiv.org/ftp/arxiv/papers/1802/1802.07782.pdf.
- https://core.ac.uk/download/pdf/55634657.pdf.
- http://www.diva-portal.org/smash/get/diva2:1115160/FULLTEXT01.pdf.
- http://plaw.nlu.edu.ua/article/view/105661.
- https://lawreview.law.ucdavis.edu/issues/51/5/Essays/51-5_Greely.pdf.
- http://maup.com.ua/assets/files/expert/1/the-future-inclusion-of-criminal.pdf.

[26] See also Karel Nedbálek.

[27] Radutniy Olecksandr Eduardovich, *Criminal Liability of the Artificial Intelligence*, in Problems of Legality, issue 138, 2017.

- https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&rid=1
- http://journals.ezenwaohaetorc.org/index.php/AFJCLJ/article/view/501.
- http://www.europarl.europa.eu/doceo/document/A-8-2019-0019_EN.html.
- https://www.inverse.com/article/30285-robot-crime-artificial-intelligence-punishment.
- http://brandonmarinoff.com/artificial-intelligence-and-criminal-law/.
- https://www.natlawreview.com/article/legal-ethics-ethical-dilemma-artificial-intelligence.