Artificial Intelligence and Criminal Liability

Adriana-Iuliana STANCU[1]

Abstract: *Technology and internet users can gain from AI and cybersecurity working together. AI can be utilized to identify cyberattacks and develop more potent defences. Machine learning algorithms, for instance, can be trained to identify odd computer network behaviours or suspect traffic patterns. They can aid in the prompt detection of cyberattacks, resulting in a quicker and more efficient reaction when handling security-related issues. What happens when artificial intelligence goes rogue, buys drugs on the darknet, or commits other criminal acts? Can it be punished? Only humans are subject to criminal accountability; legal persons are also subject to criminal liability, for which the primary sanctions are less effective than the complementary ones. In Dutch law, the use of AI is permitted in this capacity by amending the criminal provisions of the legislation, but the concept of the victim is assumed only when the victim is a human being because only he is legally protected, from the use of rights when he is also the beneficiary of social values to the protection of criminal law. Similar to the incrimination of legal persons, it would be able to incriminate AI that engages in criminal activity.*

Keywords: *artificial intelligence; criminal liability; criminalization; cybersecurity; benefits*

1. Introduction

We can start from the fact that AI is a system that performs certain objectives that require the presence of human intelligence. This requires the use of an unadulterated language, facial recognition, data analysis, and more. Cybersecurity, on the other hand, refers to protecting computer systems and networks from cyberattacks.

There are several well-known cybersecurity tools that use Artificial Intelligence technology to improve information security and protect personal data, including:

[1] Associate Professor, PhD, Faculty of Law and Administrative Sciences, Head of the Legal Science Department, "Dunărea de Jos" University of Galati, Romania, Address: Domnească Street no. 111, Galati, Romania, Corresponding author: adriana.stancu@ugal.ro.

A. Darktrace – It can be a cybersecurity tool that uses Artificial Intelligence to detect cyberattacks and respond in real time, helping to prevent data loss and cyberattacks.

B. IBM Watson for Cybersecurity – This is a cybersecurity tool that uses IBM Watson technology to identify and prevent cyberattacks. Using advanced data analytics and machine learning, IBM Watson can detect potential security threats quickly and efficiently.

C. Cylance – It can be a cybersecurity tool which incorporates Artificial Intelligence technology to identify and prevent cyber threats. Cylance uses machine learning and behavioral analysis to identify and prevent cyber-attacks.

D. McAfee – It can be a cybersecurity tool which incorporates Artificial Intelligence technology to protect personal data and security information. McAfee uses behavioral analysis and machine learning to identify security threats and prevent data loss.

Cybersecurity officials say the biggest fear is a type of deepfake technology that relies on photos and videos to create unreal images or completely unknown avatars. This new technology can generate images or videos that cause people to believe that what they see belongs to reality, and this is precisely what IT experts fear. "If cybercriminals can find ways to take your identity or create a new person, a fake identity that is not found in reality, and then they can verify it online, then we are at enormous risk", he explains. He goes on to show that artificial intelligence is now being used to commit money laundering and fraud on online platforms. In the most recent instrumental case, deepfake technology was used by a person who posed as a director of a medium-sized corporation. The resulting fraud? Millions of dollars! But this method is now being used for other purposes as well. For example, many European politicians accused deepfakes when they were misled into meeting with a man who was assumed to be a personality from the Russian opposition, as people from Navalny's entourage were considered.

2. Unsolved Criminal Issues

An annual report known as the "Global Threat Assessment" is released by the US government, outlining its assessment of the biggest risks to national security. There was no reference to IT problems in the 2007 report (Negroponte, 2007). The problem of computer security first surfaced in 2011, ranking last in terms of the harm it poses, being seen as nearly irrelevant. (Clapper, 2011). These reports have ranked IT as the top priority since 2013 (Clapper, 2013).

In 2010, a full-scale cyber war took place between the US and Iran, in which the US planted a computer worm (a type of self-replicating virus) in a uranium enrichment plant in Iran. Once there, he replaced the images from the recording cameras, deactivated the main alarm system, gathered information about the activity of that plant and accelerated the rotation of the centrifuges well beyond the normal limit for 15 minutes, after which he suddenly reduced the number to just 2 rotations per minute for another 15 minutes. The center's control and surveillance systems, which were also the focus of the computer system attack, reported regular operation for this entire period. To avoid raising suspicions, this kind of activity was repeated at specific intervals. This attack is thought to have compromised one-fifth of the uranium supply, which had a significant impact on Iran's nuclear development (Farwell & Rohozinski, 2011, pp. 23-40).

This attack was undoubtedly deliberate, carefully planned and coordinated. But another threat to AI comes from its own autonomy. As technology advances, human actors become increasingly redundant. In some situations, the human actor can even represent a vulnerability for trouble-free operation systems. These conditions create the possibility of a lack of legal liability.

To close this gap, there is growing debate on artificial intelligence's potential for criminal activity, for which it would bear responsibility. The purpose of this article is to outline these ideas, define legal capacity, examine the implications of such an approach, and offer a conclusion regarding the applicability of legal capacity in bridging the possible liability gap. According to certain writers, criminal aptitude is required to stay up with new and developing technology (Laukyte, 2019, pp. 209–213). These authors argue that the classic notions of liability, product liability and the like are no longer capable of guaranteeing the administration of justice and protecting the legitimate interests related to this technology. However, these writers contend that the way artificial intelligence is perceived is starting to shift from being seen as merely tools in a user's hands to something that goes beyond this idea. According to these authors, the shift in perception happens as technology exhibits ever-increasing levels of autonomy and intellect, as well as social skills, perception, and empathy. Other writers contend that economic considerations, rather than technological advancement, are the things that drive change (Karnow, 2018, p. XIX and onwards). For him, major changes in legislation occur when technology creates greater economic risks. These authors then explain how artificial intelligence has evolved over the decades and conclude that significant amounts of money are currently being invested in this sector. According to the cited article, some experts on the subject (Tractica) estimate that artificial intelligence (AI) revenues will increase by about $60 billion by 2025. A 2021 State of Technology assessment basically shows that artificial intelligence is still a lucrative and expanding area (McKinsey Analytics, 2021). The study concluded, among other things, that at least

56% of the study participants had purchased at least one of these systems. Most of these purchases were aimed at optimizing the services offered, while improving products based on the implementation of AI. in second place, and the automation of communication with product recipients (communication services) ranked 3rd.

3. Criminal Liability of AI, Desire and Possibility

According to the same study, in 2019 and 2020, the costs of adopting new technologies decreased simultaneously with the increase in revenues from exploring this market. It is true that artificial intelligence has important practical effects that have been observed by investors.

In terms of criminal law, the main questions concern, in our opinion, the subjective side of the crime and the causal link. It differs from other technologies because it presents a certain degree of autonomy. Some systems of this type denote such autonomy that their vision of the program becomes problematic. In some cases, the AI even deals with the programming of the AI, so that the human actor is further removed. Under these conditions, the process of establishing the subjective position, especially when it comes to unintentional crimes, becomes an extremely difficult process. At the same time, in terms of causality, the extent to which the intervention of artificial intelligence does not break the causal chain (deviant risk) is problematic. If AI deviates from its initial programming, can that action still be attributed to a person? If so, to whom? The developer, the manufacturer or the user?

Other questions in legal liability does not refer to artificial intelligence so much the theoretical framework as the practical one. Suppose a self-driving car fatally injures a person, and the evidence reveals a fault in the car's programming, although it is not excluded that this fault belongs to a single person, in practice it will most likely be a simultaneous fault. As for criminal law, for the existence of the crime, in principle, the degree of guilt does not matter, this aspect will be considered in individualization. However, in terms of criminal proceedings, prosecuting an entire department of programmers and, possibly, the legal entity as well, is a difficult and disproportionate approach. In addition, such a phenomenon would significantly discourage work in this field, given the general lack of predictability of how criminal liability would operate. A significant part of the workforce would be encouraged to choose a framework that is more protected from legal risks.

Finally, the fact that multiple states have previously changed the statutory framework or thoroughly examined this matter provides additional support for the discussion's practical significance. The Ministry of Internal Affairs is now proposing a draft regulation in Romania that would allow autonomous vehicles to be driven on public roads. It is noteworthy in these circumstances that the proposed

amendment to point 4 of the draft law states that "if the provisions of art. 131 para. (1) and (2) are fulfilled, then the verification on public roads of a vehicle that has a fully automated steering system, but only on condition that it is driven by a human, is considered fulfilled". Actually, we have already underlined the necessity of amending the Highway Code from the standpoint of causality and acceptable risk, within the context of earlier studies on the subject (Husti, 2019, pp. 84-86).

Even leaving aside all the other aspects mentioned, the mere entry into force of this rule – if it ever enters into force – would raise specific problems in the case of a road event with criminal significance.

Furthermore, although these questions are the most important, this does not mean that other elements of the general theory of crime are not challenged by new technologies.

For all these reasons, we recognize that the discussion is not strictly theoretical, even if it is at a higher degree of abstraction, but it ultimately has serious practical ramifications.

The deliberate aberrant use of AI to carry out sophisticated attacks, such as AI-based malware, through complex engineering at the societal level, to use fake accounts on social networks, which can be considered DdoS attacks via AI, through sourcing from the source to obtain unreal data, as a model, through AI assistance, or through illegal access to passwords, etc. The above also includes the use of AI systems to reinvent themselves and undermine other AI systems, that is, precisely those that use the most advanced AI techniques specifically invented to increase the efficiency of all ordinary systems against attacks of this kind. Cybersecurity is one of the underlying capabilities of AI-resilient solutions. For the safe state-wide implementation of AI throughout the European Union, this will act as an exponential starting point. But only until a stakeholder awareness of the many important risks and the difficulties they present is continuously developed will this be possible. As AI advances and numerous other technologies continue to be integrated, the threat landscape for AI is broad and constantly expanding. AI has the potential to significantly influence criminal activity (King, Aggarwal, Taddeo, & al., 2020, pp. 89–120). For this reason, the doctrine refers to a brand-new category of crime called

"AI Crime" (AIC) (King, Aggarwal, Taddeo, & al., 2020). With the potential for crimes of any kind perpetrated by AI to eventually be considered crimes under the law, AIC places a strong emphasis on using AI as a means or technique of committing crimes (Stănilă, 2019, pp. 130-157) Researchers have shown the presence of AIC through studies where they persuaded social network users to click on

phishing sites[1]. This kind of research makes it abundantly evident that artificial intelligence poses a significant and radically novel threat (King, Aggarwal, Taddeo, & al., 2020, pp. 90-91)

Once this phenomenon is recognized (Stănilă, 2019, pp. 67-68), analysing the threats that AIC poses to the growth of social relationships and ensuring sufficient and efficient criminal protection of the social values that AIC threatens are essential. Verifying if existing incriminations are adequate to safeguard social relationships and values against criminal activities committed by AI is important in the domain of solutions. "The transnational concept of order, justice, and solidarity must be reflected in domestic regulations." (Cotterel, 2017, p. 22)

Thus, a very careful verification of the effects produced by t is necessary to specifically designate agents that do not function properly or those that lead to results that contravene human rights by not respecting them. Not being human persons, the way in which AI systems behave cannot be established according to human moral standards. These days, several writers (Stănilă, 2019) stress that human particularities like empathy, cooperation, and choice are necessary for moral judgment. This rational explanation holds that artificial agents are incapable of possessing a particular moral code; instead, their behaviour is solely influenced by human characteristics.

In this sense, the applications of AI when making government decisions, especially in applied forensics, can turn into a reason for concern. Current data shows that algorithms take over and can even exacerbate preexisting disparities (Acemoglu, 2021). These are then regarded as the adverse effects of artificial intelligence, which is frequently used in politics and democratic presentations. This fact stems from algorithmic fakes on social media as well as the expanding ability of governments or big businesses to monitor people's personal lives worldwide, going beyond the democracy they claim to support (Acemoglu, 2021).

The use of AI in criminal matters: A draft report on the use of AI in law enforcement by criminal authorities was presented to a committee that deals with civil liberties, the state of the justice system, and the police environment at the end of June 2021. The report was adopted by a majority vote. The study emphasizes the dangers of using AI, which could have disastrous consequences, even as its advantages are acknowledged (European Parliament's Libe Committee Adopted a New Draft Report On The Use Of AI By The Police And Judicial Authorities, 2021).

[1] Phishing is an online fraud method that attempts to obtain personal or confidential data from the clients of various organizations. These can then be used illegally by criminals to carry out transactions on the client's account.

Psychology describes the threats AI poses to a user's mental health, and AI can lead them to commit crimes. This was demonstrated by Joseph Weizembaum (Bierstedt, 1976) after doing exercises on human-bot interaction, in which people revealed their most personal details while under the influence of AI.

Like numerous other technologies, artificial intelligence has the potential to be employed for both beneficial and detrimental reasons. Many jobs that are often completed by humans can be completed by AI, and in most cases, it outperforms them in terms of efficiency, speed at which results are established, and, most importantly, objectivity (Brundage, Avin, & Clark, 2018). Thus, the conclusion that AI can be used to perform crimes that were previously based on human thought processes far more effectively. According to the legal literature (Dupont, Stevens, Westermann, & Joyce, 2018), one of the primary features of AI as a tool for criminal activity is its ability to elevate the offender's status in relation to the victims, which makes the investigation and proof of the crime more challenging. According to this viewpoint, artificial intelligence is a real "vector" of crime (Dupont, Stevens, Westermann, & Joyce, 2018).

Crimes that are almost or entirely untraceable could be committed by humans using AI-powered technologies (Stănilă, 2020, pp. 123, 124). As a result, most people cannot mimic the voices of others or produce audio files that sound like recordings of real-world conversations. Beyond this, remarkable advancements in AI speech synthesis systems that aim to mimic human voices have been made recently. Without some specifically created safeguards, there is no real chance that AI outputs won't be identical to real recordings (Brundage, Avin, & Clark, 2018, p. 20).

The SOCTA 2017 (Europol, The EU Serious and Organised Crime Threat Assessment (SOCTA) 2017, 06 Dec 2021) research demonstrates how different kinds of criminal organizations are beginning to correlate with technological crime. Furthermore, legal doctrine demonstrates that the study of organized crime may find a new field of focus in the research of AI and technological crime (King, Aggarwal, Taddeo, & al., 2020, p. 30). As a result, AI can be crucial to criminal groups like drug cartels, for instance (Stănilă, 2020, pp. 123, 124). Understanding these phenomena will irreversibly lead to the taking of preventive measures. Thus, in Romanian legislation we find crimes that can be committed only with the help of AI. For instance, computer fraud, illegal access to computer systems, and unauthorized transmission of computer data are all offenses that are incriminated by the Criminal Code. A whole chapter on the computer field, specifically in Romanian criminal law, is also included in the same code. The treaty that governs the European Union's operations, which includes laws passed by the European Parliament and the Council pertaining to serious and especially serious crimes with cross-border ramifications, contains the first mentions of cybercrime. In this sense, Directive 2018/1673/EU on money

laundering through criminal measures (Zlati, 2020, p. 10) also includes cybercrime which was expressly provided for among the crimes affecting the economy and society, similar to the provisions of European legislation from 2013, with the specification that the latter Directive also leaves room for other crimes of the same kind (Zlati, 2020, p. 10), not limiting itself like the previous Directive.

Cybercrime has also been the subject of numerous court decisions issued by national courts, decisions that have also examined problematic usage pattern of AI tools in the criminal proceedings. As an example, we cite the Decision of 25 July 2019, of the Supreme Court in Glasgow (United Kingdom), in which a person was convicted for the first time, for sexual abuse of children in a live broadcast, which took place in the Philippines, and the Decision of December 14, 2018, of the Court of Appeal in Amsterdam in the Netherlands, in which the defendant was sentenced to the maximum penalty in the context of committing multiple crimes, obtaining, using, selling or renting a child, pornography, access and possession of software for this purpose, blackmail and fraud (Eurojust, December 2019, pp. 10-12). In the world of technology, less happens in two years than we anticipate, to paraphrase Bill Gates. It is the same with AI. The relationship between AI and crime already exists and will continue to exist, whether we like it or not. What matters is how we manage to protect ourselves, and this protection can only be ensured through a well-designed and comprehensive legislative framework that provide guarantees for the proper application of fundamental human rights and freedoms, which is particularly important harmonious coexistence with AI. We believe that discussions on this topic should be treated ethically. As a result, we must examine our beliefs about human dignity and conscience, paying particular attention to the organic and synthetic aspects. According to Erik Hoel, whose viewpoint we also share (Hoel, 2021): a machine should never be made in the likeness of the human mind, because the Artificial Intelligence that we should fear is already here.

4. Criminal Liability of Artificial Intelligence *de lege ferenda*

G. Hallevy believes that AI/robot systems can, by law, have criminal capacity. In his reasoning, he analyzes each characteristic of the crime and concludes that it is applicable to the matters at hand. Given the autonomy of the law of these social sciences, the author argues that the crime is the only basis for criminal liability and that an entity will be held criminally liable regardless of other philosophical or ethical considerations as long as it can confirm all the components of the crime's structure (Hallevy, 2015, pp. 68-70, 102). However, this perspective is predicated on the idea that not all human cognitive abilities are required to perpetrate a crime, hence it is inconsequential that artificial intelligence does not yet fully replicate the human brain and has not attained a higher degree of autonomy and independence

(Hallevy, 2015, pp. 68-70, 102): "human capacities, which are irrelevant for the commission of a particular crime, to the extent that they are not expressly required by law, are not taken into account when determining whether or not there is criminal liability".

Moreover, the fact that the perpetrator of the robbery crime possesses extraordinary culinary skills is generally irrelevant in establishing a person's criminal guilt. However, this conclusion in no way legitimizes the fact that AI/robots can be assimilated to humans to such an extent that we can speak about their possible criminal capacity, in our opinion. Based on the same conclusion presented above, one could easily consider that animals or objects can be sent to trial. Homicide, in accordance with Romanian criminal provisions, is defined as "the murder of a man or a women". Well, a bear that kills a man in the forest commits an act like the objective side of the crime of homicide. In parallel, only sometimes, it happens cognitive capabilities like those of humans. Why is the author's conclusion about the lack of need for all human capabilities not valid in his case? In fact, killing people by animals is a relatively common phenomenon, but the question of their criminal liability is certainly not raised in contemporary criminal law.

The opinion's author makes the case that AI is more like people than animals, particularly since robots can communicate more easily and employ formal logic, whereas animals are mostly instinctive and emotional beings (Hallevy, 2015, pp. 68-70, 102).

In addition to this dubious conclusion, we recognize that the real problem in establishing the criminal capacity of robots/AI is not so much the fact that they must be like humans in all respects, but the choice of minimum criteria to confer this capacity. Professor Halevy, although he addresses the issue of animals and objects in his work, does not identify and defend the existence of these minimum criteria and, consequently, does not demonstrate their compliance by technology. Moreover, using this author's system, any robot is capable of criminal action, be it a coffee maker, a self-driving car, a smart light bulb or an armed drone.

Going beyond these observations, the author I referred to analyzes in detail how the current general theory of crime applies to the criminal liability of robots. As for the objective side, he completely separates it from any elements related to will and concludes that any robot can fulfill the objective side of a crime.

5. Conclusions

Referring to the criminal liability of AI, we can observe that there is a tendency to sanction its criminal behavior, by analogy with what exists today as punishable in various legal systems. In Romania, not only natural persons but also legal entities

are sanctioned if they participate in the commission of the acts. Their sanctions are stronger and more effective in the field of complementary penalties that prohibit them from the various rights used in the commission of crimes. An analysis of the way in which AI crosses the line between licit and illicit would be required, and an amendment regarding those presented in the Romanian Criminal Code of sanctioning field would be required, to include AI in criminal liability.

6. **Bibliography**

Acemoglu, D. (2021, July 21). *Opinion: The AI we should fear is already here*. Retrieved 01 30, 2025, from The Washington Post: https://www.washingtonpost.com/opinions/2021/07/21/ai-we-should-fear-is-already-here/.

Bierstedt, R. (1976). JOSEPH WEIZENBAUM. Computer Power and Human Reason: From Judgment to Calculation. *The ANNALS of the American Academy of Political and Social Science, 426(1)*, 266-267. Retrieved 01 30, 2025, from https://journals.sagepub.com/doi/10.1177/000271627642600162.

Brundage, M., Avin, S., & Clark, J. e. (2018). The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation. *Report*. Retrieved 01 30, 2025, from https://maliciousaireport.com/.

Clapper, J. R. (2011, March 10). Statement for the Record on the Worldwide Threat Assessment of the U.S. Intelligence Community for the Senate Committee on Armed Services. Retrieved 01 30, 2025, from https://www.dni.gov/files/documents/Newsroom/Testimonies/20110310_testimony_clapper.pdf.

Clapper, J. R. (2013, March 12). Worldwide Threat Assessment of the US Intelligence Community Senate Select Committee on Intelligence. Retrieved 01 29, 2025, from https://www.dni.gov/files/documents/Intelligence%20Reports/2013%20ATA%20SFR%20for%20SSCI%2012%20Mar%202013.pdf.

Cotterel, R. (2017). The concept of "crime" and transnational networks of community. In V. Mitsilegas, P. Alldrige, & L. Cheliotis, *Globalization, Criminal Law and Criminal Justice* (p. 22). New York: Bloomsbury.

Dupont, B., Stevens, Y., Westermann, H., & Joyce, M. (2018). *Artificial Intelligence in the Context of Crime and Criminal Justice, in Korean Institute of Criminology*. Montréal: ICCC, Université de Montréal. Retrieved 01 30, 2025, from https://www.researchgate.net/publication/337402608_Artificial_Intelligence_in_the_Context_of_Crime_and_Criminal_Justice.

Eurojust. (December 2019). *Judicial Cybercrime Monitor, no. 5*. Retrieved 01 30, 2025, from https://www.eurojust.europa.eu/publication/cybercrime-judicial-monitor-issue-5.

European Parliament's Libe Committee Adopted a New Draft Report On The Use Of AI By The Police And Judicial Authorities. (2021). Retrieved 01 30, 2025, from Team AI Regulation: https://ai-regulation.com/libe-new-draft-report-on-the-use-of-ai-by-the-police-and-judicial-authorities/.

(06 Dec 2021). *Europol, The EU Serious and Organised Crime Threat Assessment (SOCTA) 2017*. Europol. Retrieved 01 30, 2025, from https://www.europol.europa.eu/publications-events/main-reports/european-union-serious-and-organised-crime-threat-assessment-2017.

Farwell, J. P., & Rohozinski, R. (2011). Stuxnet and the Future of Cyber War. *Survival, 53:1*, 23-40. doi:10.1080/00396338.2011.555586.

Hallevy, G. (2015). *Liability for Crimes Involving Artificial Intelligence Systems*. New York: Springer Edition. Retrieved from 10.1007/978-3-319-10124-8.

Hoel, E. (2021). We need a Butlerian Jihad against AI. A proposal to ban AI research by treating it like human-animal hybrids. *The Intrisic Perspective*. Retrieved 01 15, 2025, from https://www.theintrinsicperspective.com/p/we-need-a-butlerian-jihad-against.

Husti, G. (2019). Guilt and other elements regarding criminal liability in the case of autonomous cars. *Caiete de Drept Penal/Criminal Law Notebooks, no. 3*, 84-86.

Karnow, C. E. (2018). *Research Handbook on the Law of Artificial Intelligence.* Cheltenham: Edward Elgar Pub.

King, T., Aggarwal, N., Taddeo, M., & al., e. (2020). Artificial Intelligence Affordances: Deep-fakes as Exemplars of AI Challenges to Criminal Justice Systems. *Sci Eng Ethics 26*, 89–120. doi:https://doi.org/10.1007/s11948-018-00081-0.

Laukyte, M. (2019). AI as a Legal Person. *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law (ICAIL '19)* (pp. 209–213). New York: Association for Computing Machinery. doi:https://doi.org/10.1145/3322640.3326701

McKinsey Analytics. (2021, December). The state of AI in 2021. Retrieved 01 30, 2025, from https://mck.co/3xKiXhr.

Negroponte, J. D. (2007, January 11). Annual Threat Assessment of the Director of National Intelligence. Retrieved 01 30, 2025, from https://www.dni.gov/files/documents/Newsroom/Testimonies/20070111_testimony.pdf.

Stănilă, L. (2019). Artificial Intelligence and the Criminal Justice System – Criminal Risk Assessment Instruments. *Revista Română de Drept al Afacerilor/Romanian Business Law Journal, no 3*, 130-157.

Stănilă, L. (2020). *Artificial Intelligence, Criminal Law and the Criminal Justice System. Memories of the Future.* Bucharest: Universul Juridic.

Zlati, G. (2020). *Cybercrime Treatise, Vol. 1.* Bucharest: Ed. Solomon.