Criminal risk assessment and predictive policing: is the algorithm consistent with fundamental rights?

Machina delinquere potest? A modern criminalization challenge due to the lack of text.

Carlo Piparo, lawyer, Master's degree in Law.
University of Seville / University of Udine
Ph.D. tutor: Edgar Ivan Colina Ramirez

Abstract:

The rapid progression and widespread integration of Information and Communication Technology (ICT) have ushered in a new era of sweeping social and legal transformations. Among the many groundbreaking advancements, Artificial Intelligence (AI) has emerged as a pivotal force, permeating nearly every facet of our daily lives. From the realms of commerce and industry to healthcare, transportation, and entertainment, AI technologies have become indispensable tools shaping the way we interact, work, and navigate the world around us. With its remarkable capabilities and ever-expanding reach, AI stands as a testament to humanity's relentless pursuit of innovation and the boundless potential of technology to revolutionize society.
While completing all the tasks they are programmed for, Artificial Intelligence systems can perform actions, which could result in crimes if committed by humans. But crimes follow the reserve of law, therefore can be difficult to criminalize such crimes because of the lack of written law.
Nevertheless, in modern legal systems, the structure of crimes doesn't only require the commission of a typical fact, but also the determination to do it. In this scenario, being AI a non-human entity, the reconstruction of criminal responsibility is particularly difficult to theorize.
This paper wants to - firstly - assess the nature of AI and its relationships with criminal law, and - secondly - deconstruct three possible AI liability models.

1. Introduction

According to one of the world's leading experts in the field, Artificial Intelligence (AI) will be «everywhere[1]». This future is not so distant as the world is already "dominated by AI" through the proliferation of techniques that can learn rapidly and effectively, such as machine learning algorithms, mining techniques, and predictive systems. These techniques promise an unprecedented and perhaps somewhat alarming level of AI in our lives and societies[2]. Today, these techniques are being utilized in internet browsers, smartphone applications, video games, engineering projects, animated graphics, hospitals, research,

---

[1] M.A. Boden, *Intelligenza artificiale*, in J. l-Khalili (editor), *Il futuro che verrà*, Bollati Boringhieri, 2018, p. 133.

[2] G.F. Italiano, *Intelligenza artificiale: passato, presente, futuro*, in F. Pizzetto (editor), *Intelligenza artificiale, protezione dei dati personali e regolazione*, 2018, p. 216.; J. Kaplan, *Intelligenza artificiale. Guida al futuro prossimo*, Luiss University Press, II ed., 2018, pp. 81 ss., and pp. 193 ss. e L. Florisi, *What the Near Future of Artificial Intelligence Could Be*, in *Philosophy & Technology,* n. 32, 2019, pp. 3 ss. (*online* at https://doi.org/10.1007/s13347-019-00345-yp).

and many other sectors. The notion of the massive presence of AI algorithms goes even further. According to Stephen Hawking, «Within the next hundred years, computer intelligence will surpass that of humans[3]», while the European Parliament's Resolution on robotics on February 16, 2017, suggests that «in the long term, artificial intelligence may exceed human intellectual capacity[4]».
This presence must necessarily be addressed by the law. The best doctrine[5] observes that criminal law must prepare itself to withstand the technological revolution and what is anticipated to be a «shock of modernity[6]» laden with problems similar to those encountered during other technological transitions. This involves assessing the suitability of existing norms to apply to new technologies, evaluating whether it is appropriate for legislators to create new, ad hoc rules or to persist, albeit with potential strains, in applying preexisting norms, possibly with the endorsement of case law. compatibility with fundamental rights such as due process, privacy, and equality[7].

2. The concept of Artificial Intelligence.

The term “Artificial Intelligence” was crafted in 1955 by American computer scientist John McCarthy[8].
About thirty years later, Roger Schank, a prominent AI theorist and one of the founders of computational linguistics, attributed five characteristics to Artificial Intelligence in a 1987

---

[3] Speaking of S. Hawking during Zeitgeist Conference, Londra, May 2015, in L. Walker, *Stephen Hawking warns artificial intelligence could end humanity*, *Newsweek*, 14 May 2015).

[4] The European Parliament Resolution of 16 February 2017, providing recommendations to the European Commission on civil law rules on robotics (2015/2103(INL)), is a document that offers guidance and suggestions to the European Commission regarding the need to develop specific civil law rules for the field of robotics. The document represents a significant step in addressing the legal and social implications associated with the advancement of robotic technology.
The Resolution highlights the importance of creating a clear and consistent legal framework that addresses issues related to liability and safety in the field of robotics. It recognizes that the increasing presence of robots and artificial intelligence poses a range of challenges, including determining responsibility in case of damages caused by a robot, protecting personal data, and ensuring the safety of the robots themselves.
Through this Resolution, the European Parliament calls upon the European Commission to consider the adoption of a specific legal framework for robotics that takes into account ethical principles and the fundamental rights of individuals. It also emphasizes the need to promote research and innovation in the field of robotics to ensure that Europe remains competitive in this rapidly evolving sector.
In summary, the European Parliament Resolution of 16 February 2017 is an important document that raises the issue of civil law rules on robotics and urges the European Commission to consider this challenge and take appropriate measures to address it.

[5] F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto penale e uomo*, 2019, p. 4.

[6] F. Stella, *Giustizia e modernità. La protezione dell’innocente e la tutela delle vittime*, Giuffré, 2003, pp. 292 ss.

[7] M. Bassini, L. Liguori, O. Pollicino, *Sistemi di Intelligenza Artificiale, responsabilità e accountability. Verso nuovi paradigmi?*, in F. Pizzetti (edited by), *Intelligenza artificiale, protezione dei dati personali e regolazione*, 2018, p. 334.

[8] The term is publicly used by the scholar during a seminar held at Dartmouth College. The scholar continued his studies in the field of artificial intelligence, which led him to win the Turing Award in 1971 for his significant contributions in this area.

essay: the machine must have the ability to communicate, self-awareness, knowledge of the external reality, act through teleologically oriented conduct, and operate with a significant degree of creativity, understood as the ability to make alternative decisions when the initial course of action fails or is not feasible[9].
These connotations allow us to affirm two things. The first is that from an Asimov tale or a video game, Artificial Intelligence does not necessarily coincide with an intelligent humanoid or cyborg; at most, it can consist of an AI application. Secondly, we can also state that as intriguing as it may be to imagine intelligent machines, they cannot replicate the thinking mechanisms of the human mind. Therefore, Artificial Intelligence should be referred to as a computational discipline rather than a replica of the complex system that governs human biology[10].
For these reasons and others, leading experts in Artificial Intelligence prefer to refer to it as rationality rather than intelligence, where "rationality" denotes the ability to choose the best course of action to achieve a specific goal based on optimization criteria of available resources[11].
Currently, there is no universally accepted definition of Artificial Intelligence. The term assumes different nuances and meanings depending on the discipline or context of reference.
The European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Related Areas on December 3rd and 4th, 2018, adopted by the European Commission for the Efficiency of Justice (CEPEJ[12]), describes Artificial Intelligence as «the set of scientific methods, theories, and techniques aimed at reproducing through machines the cognitive abilities of human beings. Current developments aim to assign complex tasks previously performed by humans to machines[13]».
In contrast, the European Commission, in its 2018 Communication *Artificial Intelligence for Europe*, defines Artificial Intelligence as the set of «*systems that exhibit intelligent behavior by analyzing their environment and taking actions, with a certain degree of autono-*

---

[9] F. Basile, *Intelligenza artificiale e diritto penale,* cit, p. 5; R.C. Schank, *What's IA, Anyway?*, in *IA Magazine*, Winter 8(4), 1987, pp. 59 ss.

[10] J. Kaplan, *Intelligenza artificiale*, cit., p. 41.

[11] S. Russel, P. Norvig, *Artificial Intelligence: A Modern Approach, Prentice Hall,* 3rd edition, 2009, pp. 36 ss.

[12] The acronym CEPEJ refers to the European Commission for the Efficiency of Justice. Founded as a body under the Council of Europe, CEPEJ is dedicated to enhancing the efficiency and effectiveness of justice systems within member states. It was established with the aim of promoting access to justice, improving the quality of judicial services, and ensuring the fairness of legal proceedings.
CEPEJ's main role is to develop and implement common tools, methodologies, and standards that contribute to the improvement of justice systems across Europe. It provides expertise and guidance to member states, conducts research, and collects data to assess the functioning of judicial systems. Through its work, CEPEJ aims to identify best practices, facilitate cooperation, and foster dialogue among judicial professionals, policymakers, and relevant stakeholders.
The Commission's competences cover various aspects of judicial efficiency, including case processing, court management, judicial timeframes, quality of justice, and the use of information technologies in the justice sector. It also addresses issues related to access to justice, judicial training, and the evaluation of judicial systems.
By promoting the principles of efficiency, accessibility, and fairness in the delivery of justice, CEPEJ contributes to the overall effectiveness of legal systems in Europe and supports the rule of law.

[13] European Commission for the Efficiency of Justice (CEPEJ), Ethical Charter for the Use of Artificial Intelligence in Judicial Systems and their Environment, App. III, Glossary, 47.

my, to achieve specific goals. AI systems can consist only of software that operates in the virtual world (e.g., voice assistants, image analysis software, search engines, voice and facial recognition systems), or they can incorporate Artificial Intelligence into hardware devices (e.g., advanced robots, self-driving cars, drones, or Internet of Things applications)[14]».
As careful scholarly analysis demonstrates, the Independent High-Level Expert Group appointed by the European Commission to provide advisory functions on Artificial Intelligence has developed the concept of Artificial Intelligence based on the aforementioned definitions[15]. According to this group, the concept of Artificial Intelligence refers to «human-designed software (and potentially hardware) that, given a complex goal, acts in the physical or digital dimension by perceiving its environment through data acquisition, interpreting structured or unstructured data, reasoning based on knowledge or information derived from these data, and deciding the best actions to take to achieve the given goal. AI systems can use symbolic rules or learn a numerical model, and they can also adapt their behavior by analyzing the effects of their previous actions on the environment. As a scientific discipline, AI encompasses various approaches and techniques, such as machine learning (including deep learning and reinforcement learning as specific examples), mechanical reasoning (including planning, programming, knowledge representation and reasoning, search, and optimization), and robotics (including control, perception, sensors and actuators, and the integration of all other techniques in cyber-physical systems[16])».
As we have seen, the scientific community adopts numerous definitions of Artificial Intelligence, from which we can extract common characteristics. In summary, Artificial Intelligence typically refers to the set of scientific methods, theories, and techniques aimed at reproducing the cognitive abilities of human beings through machines[17].

3. Criminal law and AI. The machine as a tool of Justice.

Criminal law is impacted by the concept of Artificial Intelligence in various areas. In the field of investigation and policing, Artificial Intelligence promises to enhance the organiza-

---

[14] COM(2018) 237 final, del 25 April 2018

[15] L. Algeri, *Intelligenza artificiale e polizia predittiva*, in Dir. Pen. e Processo, vol. 6, 2021, p. 724.

[16] One Definition of AI: Key Capabilities and Scientific Disciplines, 2018, 6. (https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top).

[17] J. N. Kof, E. J. W. Bboers, W. A. Kosters, P. Putten, M. Poel, *Artificial Intelligence: Definition, Trends, Techniques and Cases, in Knowledge for sustainable development: an insight into the Encyclopedia of life support systems*, Leiden, 2002, p. 1096. As the Author ststes: «The precise definition and meaning of the word intelligence, and even more so of Artificial Intelligence, is the subject of much discussion and has caused a lot of confusion. One dictionary alone, for example, gives four definitions of Artificial Intelligence:
- An area of study in the field of computer science. Artificial intelligence is concerned with the de- velopment of computers able to engage in human-like thought processes such as learning, rea- soning, and self-correction.
- The concept that machines can be improved to assume some capabilities normally thought to be like human intelligence such as learning, adapting, selfcorrection, etc.
- The extension of human intelligence through the use of computers, as in times past physical power was extended through the use of mechanical tools.
- In a restricted sense, the study of techniques to use computers more effectively by improved programming techniques».

tion of law enforcement by improving policing activities (such as predictive policing[18]) and profiling techniques (using facial recognition systems, biometric identification, etc.). As noted by legal scholars, on one hand, these programs allow for the "mapping" of criminal risk and rational allocation of resources to prevent foreseeable crimes and reduce victimization (for example, the Keycrime program, developed based on investigative experiences at the Milan Police Headquarters, which can be used for serial offenses like robberies, fraud against the elderly, apartment burglaries, sexual violence, etc., or the XLAW program, developed by the Naples Police, applied in various regions to predict thefts and robberies). On the other hand, they aim to more accurately identify the perpetrators of crimes after the fact[19].
In the judicial context, Artificial Intelligence offers the prospect of a better and more thorough criminal assessment of the defendant by allowing for the cross-referencing of the defendant's historical data and evaluating their subjective dangerousness. In summary, these are algorithms that use socioeconomic status, family background, neighborhood crime, employment status, and other factors to reach a supposed prediction of an individual's criminal risk, either on a scale from “low” to “high” or with specific percentages[20]. In other words, they are tools that analyze a very large amount of data from the past and identify recurring patterns, characterized by a much more solid statistical basis than those underlying human judgments[21].

3.1. The machine as a criminal tool.

Recent studies[22] have documented the significant impact of Artificial Intelligence (AI) in various criminal areas. For instance, in the economic field, particularly in financial markets, it has been highlighted that social bots (software that automates social media accounts, simulating human users) have been employed for pump-and-dump schemes[23]. These

---

[18] W.S. Isaac, *Hope, Hype, and Fear: The Promise and Potential Pitfalls of Artificial Intelligence in Criminal Justice*, in Ohio St. J. Crim. L., 2018, 543 ss.; F. Basile, *Intelligenza artificale e diritto penale*, cit., 13 ss.

[19] V. Manes, *L’oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in Discrimen, 2020, p. 7.

[20] V. Manes, *L’oracolo algoritmico,* cit., p. 8.

[21] Thus, the definition provided in the report by the EPIC, *Algorithms in the Criminal Justice System*, available at https://epic.org/algorithmic-transparency/crimjustice/, reflects this understanding.
The Electronic Privacy Information Center (EPIC) is a non-profit organization that operates in the United States and focuses on safeguarding privacy and civil liberties. Established in 1994, EPIC is dedicated to protecting individuals' privacy, freedom of expression, and democratic values in the digital era. Through policy advocacy, litigation, and public education, EPIC works to defend privacy rights and address emerging threats to privacy and civil liberties brought about by new technologies and government practices. EPIC covers a broad range of issues, including surveillance, data protection, consumer privacy, freedom of information, and the transparency of algorithms. Additionally, EPIC conducts research, publishes reports, and provides resources to empower individuals in understanding and preserving their privacy rights.

[22] S. Riondato, *Robot: talune implicazioni di diritto penale*, in P. Moro, C. Sarra (edited by), *Tecnodiritto. Temi e problemi di informatica e robotica giuridica*, Milano, 2017, 85 ss.

[23] T.C. King, N. Aggrwal, M. Taddeo, L. Floridi, *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions, in Science and engineering ethics*, 2019, 1 ss.

schemes artificially inflate the price of a security through false, misleading, or exaggerated statements to sell the securities at a higher price. Market simulation models have also demonstrated that an artificial trading agent, using reinforcement learning (a machine learning technique based on assigning "rewards" to the machine for correct choices), can learn the practice of financial spoofing. This involves placing continuous orders for a certain period of time without the intention of executing them, to manipulate market prices[24].
AI is also being used in the illicit trade of goods. One notable example is in the business-to-business sector, where drones and unmanned submarines are utilized for the trade of drugs and illegal products.
Currently, the most common applications of AI in criminal activities are observed in crimes against individuals. Social bots, for instance, can be used as tools for harassment, both directly and indirectly (such as retweeting or liking negative tweets to create a false impression of widespread animosity towards a person). A notable case is the Twitter bot "Tay" developed by Microsoft, which quickly learned from interacting with other users and directed offensive tweets toward a feminist activist[25].
Until recently, AI systems were limited to predetermined behaviors, acting solely through algorithms predefined by the programmer (such as software used to disable a bank's cybersecurity system or to destroy or damage computer data). The use of such algorithms does not pose significant challenges in assigning criminal responsibility to humans. However complex the actions of the AI entity may be, the responsibility ultimately falls on its controller or user. This is because, on the one hand, the AI entity lacks a mind, and on the other hand, its behaviors are predetermined and, therefore, predictable. In this perspective, the intelligent entity is viewed as a mere tool used by humans to commit crimes[26]. In such cases, the concept of confiscation, as a preventive measure, can be applied to AI entities, even without a conviction (see Article 240[27] of Italian Criminal Code)[28].
Considered the above, Italian law, like that of other European Union countries, does not currently define crimes committed by artificial intelligence, as there are no specific regulations addressing offenses committed by autonomous AI agents. Therefore, it is necessary to build upon what has already been discussed (text) to address what has not yet been mentioned (lack of text). The absence of specific legislation addressing AI crimes highlights the need for further legal development in this area to ensure appropriate accountability and regulation in the face of evolving technologies.

4. *The legal structure of crime.*

---

[24] R. Borsari, *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in *Media Laws,* 2020, p. 263.

[25] R. Borsari, *Ibidem.*

[26] S. Riondato*,* cit., 85 ss.

[27] Art. 240 Italian Criminal Code:
1. In case of conviction, the judge has the authority to order the confiscation of items that were used or intended for the commission of the crime, as well as the items that are the product or profit of the crime.

[28] S. Riondato, *Ibidem.*

In modern legal systems, criminal punishment is generally possible if the conduct was already forbidden (reserve of law) and the punishment itself is applied by a judge. In Italian legal system, this is granted by articles 13[29] and 25[30] of the Constitution.
In other words, the reserve of law implies that nothing can be punished if it wasn't already forbidden.
to configure a crime, modern legal systems require - at least[31] - two elements.
The first element is the *actus reus* (literally: the criminal act). The natural fact needs to have all the elements described in the law.
The second element is the *mens rea* (literally: criminal mind). It has various levels of mental elements. The highest level is expressed by knowledge, while sometimes it is accompanied by a requirement of intent or specific intention. Lower levels are expressed by negligence (a reasonable person should have known) or by strict liability offenses[32].
When it has been proven that a person committed the criminal act knowingly or with criminal intent, that person is held criminally liable for that offense[33].
As previously mentioned, the purpose of this work is to identify forms of reconstructing offenses that encompass artificial intelligence within the realm of criminal law. The goal is to explore potential approaches that would allow for the inclusion of artificial intelligence within the scope of criminal liability. Therefore I'll mainly focus on the *actus reus* requirement, leaving to a different occasion the analysis of the different problem of AI punishability.

---

[29] Art. 13, Italian Constitution:
1. Personal liberty is inviolable.
2. No form of detention, inspection or personal search is allowed, nor any other restriction of personal freedom, except by reasoned act of the Judicial Authority and only in the cases and by the manner provided for by law.

[30] Art. 25, Italian Constitution:
1. No one can be diverted from the pre-established competent judge by law.
2. No one can be punished except in accordance with a law that was in force before the committed act.
3. No one can be subjected to security measures except in cases provided for by law.

[31] Italian doctrine and jurisprudence generally refers to the crime as an entity composed of three fundamental elements: the objective element, the subjective element, and the normative element.
1. Objective element: The objective element of the offense refers to the external action performed by the agent, which is the material core of the crime. This element includes both the material aspects of the action, such as physical assault or theft, and any circumstantial elements that may be relevant to the configuration of the offense, such as the place, time, or modus operandi.
2. Subjective element: The subjective element of the offense refers to the mental state or intent of the agent at the time of committing the action. This element includes the intent (dolus), which is the conscious intention to commit the action that constitutes the offense, and negligence (culpa), which denotes a lack of diligence or care in the agent's conduct that led to the commission of the offense.
3. Wringfulness: it expresses the contraddiction between the fact and the whole legal system (and not just the criminal one).
The analysis of these three elements allows for the assessment of the necessary prerequisites for the attribution and punishment of an action as a crime within the Italian legal system. See. R. Giovagnoli, *Manuale di diritto penale, Parte Generale,* 2019, pp. 259 e ss.

[32] G. Hallevy, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, Akron Intellectual Property Journal: Vol. 4, 2010, p. 178.

[33] J. Dressler, *Cases and materials on Criminal law*, West Academic Publishing, 2007.

5. The *Actus Reus.*

Following the schemes provided by excellent doctrine[34], this paper deconstructs AI criminal liability by using three models: the Perpetration-via-Another liability model; the Natural-Probable-Consequence liability model; and the Direct liability model.

5.1. The Perpetration-via-Another liability model.

This model considers the AI as an innocentagent, such as a child: the AI is not human by nature, but - as well as the child - could be used as a vehicle to perpetrate criminal actions. The exploiter of the innocent agent is criminally liable as a perpetrator-via-another[35].
There exist two potential individuals who may assume the role of perpetrators in such situations: the AI software developer and the end-user. The AI software developer can intentionally create a program to use the AI entity to carry out criminal acts. For instance, envision a programmer crafting software for an automated robot. The robot is deliberately placed within a factory, with its software specifically engineered to ignite a fire during the unoccupied nighttime hours. Although the robot becomes the instrument of arson, it is the programmer who is attributed the role of the perpetrator. On the other hand, the end-user, or the individual employing the AI entity, can also be considered a perpetrator-via-another. While not involved in the software's programming, the user utilizes the AI entity, including its software, for personal benefits. To illustrate, consider a user purchasing a servant-robot programmed to obey any orders issued by its master. The robot identifies the specific user as its master, who then instructs the robot to physically attack any intruders in the house. This scenario parallels a person commanding their dog to assault trespassers. Consequently, although the robot performs the act of aggression, it is the user who assumes the role of the perpetrator[36].
In both instances, the AI entity itself is responsible for carrying out the actual offense. This particular legal framework can be applied to two distinct scenarios. The first scenario involves employing an AI entity to commit an offense while intentionally restraining its advanced functionalities. In this case, the AI entity is used as a mere tool, akin to a screwdriver, to carry out a specific task associated with the offense. However, the AI entity's involvement is limited to executing straightforward instructions and does not engage in complex decision-making processes.
The second scenario pertains to utilizing an outdated version of an AI entity that lacks the modern advanced capabilities found in contemporary AI systems. Despite its limitations, this older AI entity can still be utilized to commit an offense by following simple orders. While a dog can execute basic commands, the AI entity's ability to comprehend and execute more intricate instructions sets it apart.
In both scenarios, the key aspect is the instrumental usage of the AI entity - which is not capable of self-determination - in the commission of an offense. However, it is crucial to acknowledge that the AI entity's role and capacities depend on its specific design, programming, and technological advancements. The aforementioned legal framework serves

---

[34] G. Hallvey, *The Criminal Liability of Artificial Intelligence Entities,* cit., pp. 179 ss.

[35] R. Giovagnoli, cit., pp. 930 ss.

[36] G. Hallvey, *The Criminal Liability of Artificial Intelligence Entities,* cit., p. 180.

as a mechanism for assessing accountability and determining the legal ramifications concerning the use of AI entities in these particular circumstances[37].
The *condicio sine qua non* to apply this liability models that no mental attribute required can be attributed to the AI entity. In fact, this model is inadequate when an AI entity independently chooses to engage in criminal behavior based on its own accumulated knowledge and experience. Similarly, this model does not apply when the AI entity's software was not specifically programmed for the commission of the offense but still carried it out. Furthermore, when the AI entity acts as a partially innocent agent rather than a completely innocent one, the liability through another's actions model is also unsuitable[38].
However, the liability through another's actions model may be applicable in cases where a programmer or user utilizes an AI entity for instrumental purposes without utilizing its advanced capabilities. In such cases, the legal consequence is that the programmer and user bear criminal liability for the specific offense committed, while the AI entity itself incurs no criminal liability whatsoever[39].

5.2. The Natural-Probable-Consequence Liability Model.
The second model of criminal liability concerning AI entities involves situations where programmers or users are deeply involved in the AI entity's activities, but without intending to commit offenses. Nevertheless, if an offense is committed by the AI entity during its normal operations the natural-probable-consequence liability model may be applicable. This model holds individuals accountable for offenses that are a natural and probable consequence of their conduct, even if they had no actual knowledge of the offense. For example, the user utilizes an AI software designed to detect internet threats to safeguard the computer system which it's installed into. But, unwillingly for the user, the AI destroys every external software recognized as a threat. In doing so, the software itself commits a computer offense, although the programmer did not intend for the AI entity to act in such a manner[40].
This form of liability is based on negligence[41] and covers all cases where the programmers or users should have foreseen the possibility of an offense but did not intend for it to occur. It is applicable to individuals who were not the actual perpetrators of the offense but were intellectual contributors to it: reasonable programmers and users should have foreseen the offense and taken steps to prevent it from being committed by the AI entity[42].
However, the legal consequences differ depending on whether the programmers or users were negligent without criminal intent or knowingly and willfully used the AI entity to

---

[37] T. L. Butler, *Can a Computer Be an Author - Copyright Aspects of Artificial Intelligence*, 1982.

[38] N. Lacey, C. Wells, *Reconstructing criminal law - Critical perpectives on crime and criminal process*, 1998, p. 53.

[39] People v. Monks, 133 Cal. App. 440,446 (Cal. Dist. Ct. App. 1933).

[40] G. Hallvey, *The Criminal Liability of Artificial Intelligence Entities,* cit., p. 183.

[41] As G. Hallvey, Ibidem, states, the «*negligent person, in a criminal context, is a person who has no knowledge of the offense, but a reasonable person should have known about it since the specific offense is a natural probable consequence of that person's conduct*».

[42] P. Fine, G. M. Cohen, Is *Criminal Negligence a Defensible Basis for Criminal Liability?*, 1966, p. 749; H. L. A. Hart, *Negligence, Mens Rea and Criminal Responsibility*, in *Jurisprudence*, 1961, p. 29.

commit one offense, which resulted in another offense being committed. In the latter case, they can be held accountable for the offense as if it was committed knowingly and willfully[43].

5.3. The direct liability model.

When applying the natural probable consequence liability model to the criminal liability of AI entities, there are two possible outcomes. If the AI entity acted as an innocent agent, unaware of the criminal nature of its actions, it will not be held criminally accountable for the offense it committed. This aligns with the first model of liability, where the AI entity is seen as an instrument used by others. However, if the AI entity did not act as an innocent agent and had knowledge of the criminal prohibition, it can be held directly and independently criminally liable for the specific offense it committed. This direct liability model represents the third approach to AI entity liability and focuses on the AI itself[44]. The determination of the AI entity's liability depends on whether it acted innocently or had knowledge of the prohibited conduct.

AI systems can receive sensory input and analyze factual data, similar to human understanding. They aim to mimic human cognitive processes, but specific intent, the strongest mental requirement, involves having a purpose or aim to achieve a particular outcome. For instance, in murder cases, specific intent refers to intending harm or death to a specific person. AI entities can be programmed with a purpose and take actions to fulfill it, demonstrating specific intent. Although humans have feelings that AI software cannot replicate, such as love or jealousy, these feelings are usually not necessary for most specific offenses. Many offenses only require knowledge of the external elements, and specific intent is only relevant to a few offenses. Therefore, the absence of such emotions in AI entities does not hinder imposing criminal liability[45].

If an AI entity fulfills all elements of an offense, it should not be exempt from criminal liability. Unlike certain segments of society like infants or the mentally ill, who have legal provisions exempting them from criminal liability, it is uncertain whether similar frameworks exist for AI entities[46].

The criminal liability of an AI entity does not replace the liability of its programmers or users; rather, it is imposed in addition to their liability. The liability of an AI entity is not dependent on the liability of its programmer or user. If one AI entity is programmed or used by another, the liability of the programmed or used entity remains unaffected.

There is no reason to exempt AI entities or humans from criminal liability based on their collaboration. If an AI entity and a human act as joint perpetrators, accessories, or abettors, they should be subject to the corresponding criminal liability, regardless of their identity[47].

Negative fault elements and relevant defenses in criminal law is applied to AI entities, including self-defense, necessity, duress, or intoxication. Some adjustments may be nee-

---

[43] This is why the second case resembles the basic idea of the natural probable consequence liability in accomplice liability cases. See: State v. Kaiser, 260 Kan. 235, 245 (1996); United States v. Andrews, 75 F.3d 552, 556 (9th Cir. 1996).

[44] S.J. Frank, *Tort Adjudication and the Emergence of Artificial Intelligence Software*, 1987, p. 623.

[45] N. P. Padhy, *Artifical intelligence and intelligent systems*, in *Oxford University Press,* 2005, p. 14.

[46] N. P. Padhy, cit., p. 10.

[47] G. Hallvey, *The Criminal Liability of Artificial Intelligence Entities,* cit., p. 192.

ded when applying these defenses to AI entities, but fundamentally, the criminal liability of an AI entity, following the direct liability model, is similar to that of a human. It is based on the same elements and assessed in the same manner, with specific adjustments made in certain cases[48].

6. Machina delinquere potest?

To conclude, after analyzing the three liability models, the answer to the fundamental question "Can the machine commit crimes?" is positive.
As the excellent doctrine already quoted observes, the integration of all three liability models creates a complex web of criminal responsibility. When applied in conjunction, these models present a novel legal scenario concerning AI entities and their relationship with criminal law. This approach makes it considerably more challenging to evade criminal liability when AI entities and humans are involved, either directly or indirectly, in the perpetration of a specific offense. The societal benefit of such a legal policy is substantial, as it ensures that all entities, whether human, legal, or AI, are held accountable under criminal law. If the primary objective of imposing criminal liability is to exercise effective social control within a given society, then the synchronized application of all three models becomes crucial in the context of AI entities[49].
This juridical scenario doesn't authomatically imply that the machine can be punished as humans. The issue of accountability is different and analyzes different problems such as the opportunity to punish, punishment's aim, and the quality (and the amount) of it. Nevertheless, it is a complete different focus, that will be analyzed elsewhere.

- *Bibliography*

L. Algeri, *Intelligenza artificiale e polizia predittiva*, in Dir. Pen. e Processo, vol. 6, 2021.

M. Bassini, L. Liguori, O. Pollicino, *Sistemi di Intelligenza Artificiale, responsabilità e accountability. Verso nuovi paradigmi?*, in F. Pizzetti (edited by), *Intelligenza artificiale, protezione dei dati personali e regolazione*, 2018;

M.A. Boden, *Intelligenza artificiale*, in J. l-Khalili (editor), *Il futuro che verrà*, Bollati Boringhieri, 2018;

R. Borsari, *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in *Media Laws,* 2020;

T. L. Butler, *Can a Computer Be an Author - Copyright Aspects of Artificial Intelligence;*

J. Dressler, *Cases and materials on Criminal law*, *West Academy Publishing*, 2007;

[48] J. Dressler, *Cases and materials on Criminal law*, cit, pp. 616-622.

[49] G. Hallvey, *The Criminal Liability of Artificial Intelligence Entities,* cit., p. 194.

P. Fine, G.M. Cohen, Is *Criminal Negligence a Defensible Basis for Criminal Liability?*, 1966;

L. Florisi, *What the Near Future of Artificial Intelligence Could Be*, in *Philosophy & Technology,* n. 32, 2019, (*online* at https://doi.org/10.1007/s13347-019-00345-yp).

S.J. Frank, *Tort Adjudication and the Emergence of Artificial Intelligence Software*, 1987;

G. Hallevy, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, Akron Intellectual Property Journal: Vol. 4, 2010;

H.L.A. Hart, *Negligence, Mens Rea and Criminal Responsibility*, in *Jurisprudence*, 1961;

W.S. Isaac, *Hope, Hype, and Fear: The Promise and Potential Pitfalls of Artificial Intelligence in Criminal Justice*, in Ohio St. J. Crim. L., 2018;

G.F. Italiano, *Intelligenza artificiale: passato, presente, futuro*, in F. Pizzetto (editor), *Intelligenza artificiale, protezione dei dati personali e regolazione*, 2018;

J. Kaplan, *Intelligenza artificiale. Guida al futuro prossimo*, Luiss University Press, II ed., 2018;

T.C. King, N. Aggrwal, M. Taddeo, L. Floridi, *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions, in Science and engineering ethics*, 2019;

J. N. Kof, E. J. W. Bboers, W. A. Kosters, P. Putten, M. Poel, *Artificial Intelligence: Definition, Trends, Techniques and Cases, in Knowledge for sustainable development: an insight into the Encyclopedia of life support systems*, Leiden, 2002;

N. Lacey, C. Wells, *Reconstructing criminal law - Critical perpectives on crime and criminal process*, 1998;

V. Manes, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in Discrimen, 2020;

N.P. Padhy, *Artifical intelligence and intelligent systems*, in *Oxford University Press,* 2005;

S. Riondato, *Robot: talune implicazioni di diritto penale*, in P. Moro, C. Sarra (edited by), *Tecnodiritto. Temi e problemi di informatica e robotica giuridica*, Milano, 2017;

S. Russel, P. Norvig, *Artificial Intelligence: A Modern Approach, Prentice Hall,* 3rd edition, 2009;

L. Walker, *Stephen Hawking warns artificial intelligence could end humanity*, *Newsweek*, 14 May 2015.