

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence

هل يُمكن معاقبة الروبوت؟ المسؤولية الجنائية في مواجهة الذكاء الاصطناعي

Mokkedes Amina*, Laboratory of Fundamental Private Law, Faculty of Law
University of Ain Temouchent, Algeria

amina.mokkedes@univ-temouchent.edu.dz

Bentaallah Zahira, Laboratory of Fundamental Private Law, Faculty of Law
University of Ain Temouchent, Algeria

zahira.bentaallah@univ-temouchent.edu.dz

Received: 14/07 /2025

Accepted: 31/10 /2025

Published: 29/11 /2025

Abstract:

As artificial intelligence advances, robots can commit acts that may cause harm or even amount to crimes, raising the question of whether robots themselves can be punished or if humans remain solely responsible. This paper explores AI-related crimes, such as fraud and killings by autonomous robots, and points out Algeria's lack of legal provisions for these issues. It argues that existing criminal laws need adaptation and suggests assigning civil liability to manufacturers and considering special legal status for AI systems. The study concludes that laws must evolve to balance justice and technological progress. Understanding these challenges is crucial for developing legal systems ready for the age of intelligent machines.

Key word : criminal responsibility; artificial intelligence; artificial intelligence crimes.

Introduction:

The era of artificial intelligence (AI) and robotics reflects a new criminal phenomenon, whereby criminals exploit AI systems and robots to commit offences. The use of these AI-driven devices is no longer limited to serving humanity; rather, they have increasingly been employed negatively, even to the extent of perpetrating crimes.

AI-related crimes can be defined as unlawful acts committed by means of AI systems or those exploiting the capabilities of such systems to carry out criminal activities, such as cyber intrusions, hacking, forgery, fraud, account breaches, and violations of privacy. Even more alarming, crimes committed by these entities can extend to endangering or taking human life, as is the case with medical robots, autonomous lethal robots, and weapons relying on AI technologies.

The subject of this study concerns criminal liability for AI-related crimes. Given that AI itself lacks consciousness and free will, it is necessary to investigate who should bear liability when AI systems commit acts that constitute crimes under the law.

With the rapid development of AI, there arises an urgent need for States—including the Algerian legislator—to develop legal frameworks and regulations that establish criminal liability and the corresponding sanctions. It is noteworthy that, to date, no specific international conventions exist to address AI-related crimes¹. Instead, there are merely some international initiatives dealing with related issues, pending greater

* Corresponding Author

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

cooperation among States to develop legal frameworks regarding AI, including the adoption of conventions to combat such crimes, which are unlikely to remain limited in scope given the ever-evolving nature of AI itself.

Based on the foregoing, our article raises several questions and seeks to answer them, most notably: Who bears criminal liability for acts committed by AI? Is it possible for the programmer or manufacturer of AI systems to be held criminally responsible? And if so, what types of sanctions would be appropriate for such crimes?

In this article, we shall adopt a descriptive, analytical, and comparative methodology, through which we will examine the concept of AI and the crimes emerging therefrom, and provide an analysis of the extent of criminal liability associated with the use of AI systems, as well as the appropriate sanctions for such offences.

1-Tilte I: The Concept of Artificial Intelligence Crimes

The widespread use of artificial intelligence (AI) raises numerous negative implications and challenges that significantly affect society. Economically, the adoption of AI in various tasks may lead to job losses, as software systems can replace human workers at lower costs and with higher efficiency, resulting in increased unemployment rates. This, in turn, may contribute to a rise in crimes associated with unemployment, such as theft, drug trafficking, sexual offenses, and even suicide.

Moreover, the proliferation of AI poses a threat to individuals' privacy and personal rights, particularly in the absence of adequate legal regulation and oversight of this technology. Most technological services currently available require users to consent to the use of AI systems for collecting personal data, whether from their mobile phones or other devices used to access such technologies. These systems analyze the collected data and employ it for various purposes, including commercial objectives, often without regard for individuals' privacy and personal rights.

Accordingly, in this section, we will address the definition of AI crimes, present examples of such crimes, and examine the legal elements that constitute these offenses.

1.1- First Subtitle: Definition of Artificial Intelligence Crimes and Some Examples of These Crimes

To define artificial intelligence (AI) crimes, it is necessary first to address the definitions of the terms "crime" and "artificial intelligence," and then to derive a specific meaning for AI crimes. To provide greater clarity, we will also discuss examples of certain AI-related crimes.

1.1.1 Definition of AI Crimes

We shall begin by defining each of the two terms: "crime" and "artificial intelligence."

Crime: From a legal perspective, crime is any act or omission committed by an individual for which the law prescribes a criminal penalty².

As for the definition of artificial intelligence, the term itself is composed of two words: "intelligence" and "artificial." Intelligence refers to the capacity for understanding, reasoning, analysis, and differentiation. The word artificial, on the other hand, denotes actions resulting from the creation or fabrication of things³.

Artificial intelligence is defined as a branch of computer science concerned with how machines can simulate human behavior. This field focuses on developing devices and software capable of thinking in ways that resemble the functioning of the human brain.

Accordingly, and based on the definitions of these two terms individually, we may deduce a definition of AI crimes as follows: they are crimes in which artificial intelligence is involved or that result from its use in unlawful ways.

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

1.1.2 Examples of Artificial Intelligence Crime

Crimes resulting from the actions of artificial intelligence (AI) entities are diverse and vary in the degree of human involvement depending on the level of technological advancement. These crimes include: electronic fraud, data manipulation for unlawful purposes, violations of privacy, incitement to hatred or violence, and even homicide. Below is a brief overview of selected AI-related crimes:

1.1.1.1- Fraud Crimes

- In 2019, a British company was defrauded of \$243,000 after criminals used AI to mimic the voice of its CEO.
- Similarly, drug cartels in Mexico have employed AI-controlled drones, granting them flexibility and coordination in targeting human threats or supply chains⁴.

1.1.1.2-Homicide by AI Robots

- In one of his books, Professor **Gabriel Hallevy** recounts a true story that reads like a science fiction plot. In 1981, a Japanese worker was killed by a machine equipped with AI after it mistakenly identified him as a threat. The machine pushed the employee toward another device nearby, resulting in his immediate death. This incident illustrates not only that AI systems are fallible, but also raises critical legal and ethical questions regarding the consequences of such errors⁵.
- In 2016, a police robot was involved in the killing of a suspect named Johnson following a confrontation in which five officers were killed and seven others injured. The police attached explosives to a bomb-disposal robot and detonated it when the robot approached the suspect⁶.
- Another tragic incident occurred in 2015 at a Volkswagen plant in Baunatal, Germany, where a robot killed an unidentified man. He was part of a team installing the machine when it grabbed and crushed him against metal panels, causing fatal injuries⁷.
- In 2009, Ana Maria Vital, a 40-year-old employee at Golden State Foods in Industry City, California, was killed by a robotic arm that was still operational. The robot mistook her for one of the boxes it was stacking and fatally crushed her torso, holding her tightly despite efforts from mechanics to free her⁸.

1.1.1.3-Crimes by Autonomous Vehicles

- One of the most notorious criminal cases involving autonomous vehicles occurred in Arizona in 2018, when a self-driving car operated by a company affiliated with **Uber** struck and killed a pedestrian⁹.
- Additionally, such vehicles have been involved in traffic violations, particularly in Canada.

1.1.1.4-AI Crimes Involving Facebook Algorithms

A legal question arises regarding Facebook's potential criminal liability in the event of a data breach involving its users. Facebook collects user data through complex AI-driven algorithms and cookies to enhance service functionality. While such data collection is not considered a crime—since users typically consent to it—the issue becomes problematic if this personal data is leaked. In such a case, legal accountability must be considered.

1.1.1.5-AI Weapons Crimes

Often referred to as "killer robots," these autonomous weapons systems have been defined by Human Rights Watch and the U.S. Department of Defense as machines capable of selecting and engaging targets without further human input once activated.

Recently, the development of new, AI-powered autonomous weapons has raised concerns about the possibility of future wars being fought by machines capable of making kill decisions independently. There is a growing fear that such robots might rebel against

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

their programmers. This concern was highlighted by Canadian journalist Jane Gerson, who authored an article titled *"Killer Robots: A New War on the Horizon?"*

One prominent example is the X-47B, an American unmanned combat aerial vehicle. Killer robots are among the most dangerous AI applications, as they may become uncontrollable, especially if compromised by viruses that disable command systems. This could result in civilian casualties or even friendly-fire incidents involving allied military forces.

1.2- Second Subtitle: The Elements of Artificial Intelligence Crimes

Every crime consists of three fundamental elements: the material element, the moral element, and the legal element. Accordingly, crimes involving artificial intelligence are subject to the same elements, detailed as follows:

1.2.1-The Material Element (Actus Reus)

This represents the conduct carried out by the perpetrator, which manifests externally and produces certain legal consequences. This element itself is composed of three components that constitute the physical structure of the crime:

The external conduct, which may be either an act (commission) or an omission.

The resulting consequence, whether it actually materializes or not.

The causal link connecting the conduct to the consequence¹⁰.

Therefore, this element applies equally to artificial intelligence crimes whenever the criminal conduct occurs, whether through an act or an omission.

1.2.-2The Moral Element (Mens Rea)

This element reflects the psychological or mental state accompanying the act. It is insufficient for a crime to exist based solely on the physical act; it must also be committed with free will. The moral element can take the form of intent, thereby rendering the crime intentional, as provided under Article 30 of the Algerian Penal Code, which stipulates that the offender's will must be directed toward committing the offense with knowledge of its legal elements. Alternatively, it may take the form of negligence or recklessness, in which case the crime is classified as unintentional.

1.2.-3The Legal Element (Legality)

This element comprises two aspects. The first is that the act must fall under a specific criminal provision that criminalizes the conduct and prescribes a corresponding penalty, embodying the principle of legality of crimes and punishments. The second aspect concerns the applicability of the criminal provision to the act committed in terms of territory and time, and the absence of any legal justification that might exempt the act from criminal liability¹¹.

2-Tilte I: Procedural Rules Governing Criminal Liability for Artificial Intelligence Crimes

To define criminal liability, it is necessary to examine the meaning of each term separately, as well as to explore its linguistic and technical (legal) definitions.

From a linguistic perspective, the expression "criminal liability" consists of two words: liability and criminal.

From a technical (legal) perspective, criminal liability is defined as the individual's responsibility for his or her criminal act, subjecting the person to the penalty prescribed in criminal law.

The Algerian legislator has not expressly defined criminal liability in either the Penal Code or the Code of Criminal Procedure, leaving its definition to legal scholarship and

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

judicial practice. The legislator has merely excluded criminal liability in cases where free will is absent¹².

Given the seriousness of artificial intelligence crimes—which are expected to increase with the development and proliferation of robots—it is necessary to establish a legislative solution to curb such violations and to clearly determine who bears criminal liability. This is crucial to prevent offenders from evading punishment under the pretext that the crime was committed through a device or robot devoid of human awareness and emotions.

Accordingly, we shall address the parties who may bear criminal liability for such crimes, and subsequently discuss the circumstances under which this liability may be excluded.

2.1- First Subtitle: Parties to Criminal Liability

The criminal legal system is founded upon several principles, foremost among them being the principle of criminal liability, whereby punishment is imposed on the actual perpetrator of a crime—that is, the individual who committed the criminal act. The significance of criminal liability lies in determining whether the accused should be punished, considering that, at least for now, the human being remains the only entity capable of bearing criminal responsibility. Consequently, the provisions of criminal law are directed exclusively toward natural persons.

However, with the emergence of artificial intelligence (AI) systems and their increasing integration into various fields—including those that may lead to criminal acts—a new legal challenge arises concerning the determination of the party who should bear criminal liability when a crime is committed. In this context, four potential parties to criminal liability can be identified.

.2.11. Criminal Liability of AI Itself

Under the current state of technology, it appears highly unlikely that AI systems could be considered liable entities under criminal law. Despite AI's capacity for self-learning and problem-solving, current AI systems are not sufficiently developed to operate entirely autonomously. The "decisions" made by AI systems are, in reality, determined by embedded programming, algorithms, and user inputs. Thus, AI cannot be attributed legal capacity, which includes the inability to be held criminally liable. Nevertheless, this conclusion might evolve with future technological advancements¹³.

2.1.2. Criminal Liability of the Programmer or Manufacturer

A distinction must be made between the programmer and the manufacturer. The programmer is the individual who writes the code guiding the system's operations, typically using machine languages, while the manufacturer is responsible for producing the physical components operated by the robot's AI. However, the roles of programmer and manufacturer may sometimes be embodied in a single entity.

2.1.2.1 Criminal Liability of the Programmer:

As previously noted, the establishment of criminal liability requires the existence of criminal intent (*mens rea*). Such intent may be present if a programmer deliberately designs AI systems to commit criminal acts—for instance, creating software or applications to manipulate markets or hack into bank accounts. In such cases, the programmer would be considered the moral author of the crime, with the robot merely serving as the instrument of its execution¹⁴.

Conversely, if crimes are committed due to programming errors attributable to negligence by the programmer, the latter could incur criminal liability for committing an unintentional offence arising from negligent acts (such as carelessness, recklessness, or lack of due diligence).

2.1.2.2 Criminal Liability of the Manufacturer:

The manufacture of AI technologies raises significant concerns regarding criminal liability. When

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

examining liability for conduct that constitutes a criminal offence, clarifying the role of the manufacturer is essential. Manufacturers may sometimes seek to shield themselves from criminal liability through specific clauses in user agreements, thereby transferring responsibility solely to the owner for offences committed via the AI entity.

However, manufacturers and designers may nevertheless bear criminal liability for defects in AI software or hardware that result in the commission of criminal acts. It is crucial to distinguish between liability arising from deliberate conduct and liability arising from negligence, as the applicable sanctions differ in each case¹⁵.

2.1.3. Criminal Liability of the Owner or User

The owner or user of an AI system is presumed liable upon acquiring and possessing it, given that they benefit from such technology and exercise control over it. Accordingly, determining the owner's or user's criminal liability requires distinguishing between intentional and unintentional crimes:

2.1.3.1 Intentional Crimes Committed by the Owner or User:

The owner or user is subject to punishment under the applicable criminal provisions. For instance¹⁶, if an AI system is employed to commit murder, the owner or user would be prosecuted under the legal provisions specific to homicide. The same applies if the AI system is used for theft or robbery, whereby the applicable penalties for those offences would be imposed.

2.1.3.2 Unintentional Crimes Committed by the Owner or User:

Criminal liability may arise even absent criminal intent—for example, due to negligence, recklessness, failure to exercise caution, or disregard of laws and regulations while using AI. In such instances, the owner or user would be subject to the prescribed penalties for the specific unintentional offence committed¹⁷.

In certain circumstances, an AI user may still be held liable for acts committed by the system. For example, one could imagine a person operating an autonomous vehicle remaining liable for an accident caused by the AI system—not because they intended to cause harm, but because they could and should have prevented the incident. This form of vicarious criminal liability appears possible only in cases of negligence (either criminal or quasi-criminal). By contrast, in cases involving specific or general intent crimes, the AI user may escape criminal liability due to the absence of personal intent¹⁸.

2.1.4 Criminal Liability of Third Parties

Third parties refer to individuals who are not directly linked to the AI system—i.e., parties other than the manufacturer, programmer, owner, or user. The criminal liability of third parties arises if they exploit vulnerabilities in AI devices or software—for example, through hacking or cyberattacks. Such acts are universally recognized as unlawful under both national and international law.

An example would be a third party—whether an individual or a group—gaining unauthorized access to the AI system's data or networks without assistance from the programmer, owner, or user. Such unauthorized access may threaten the integrity and confidentiality of the AI system, damage or destroy it, or jeopardize national security and economic interests. In such cases, the third party incurs criminal liability. Moreover, joint criminal liability may arise if the third party commits such offences due to negligence by the manufacturer, programmer, or owner, thus rendering liability shared among the responsible parties¹⁹.

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

2.2- Second Subtitle: Divergent Approaches Regarding Criminal Liability for Artificial Intelligence Crimes (Robot Crimes as a Model)

Legal and scholarly opinions have diverged on the question of whether devices associated with artificial intelligence, particularly intelligent robots, can be held criminally liable for acts constituting crimes. These views have essentially split into two main schools of thought:

- The first school supports the application of criminal liability provisions to crimes committed by such devices, provided that the material and moral elements of the crime are fulfilled.
 - The second school rejects the notion of subjecting such entities to criminal liability, based on their non-human nature and the legal and philosophical implications arising therefrom.
- Below is a brief exposition of each of these two approaches:

2.2.1- The Approach Supporting Criminal Liability of Robots

Proponents of this approach believe that intelligent robots are capable of fulfilling the requirements of criminal liability, both materially and morally.

From the material perspective, this school argues that acts committed by robots operating autonomously from human intervention can be regarded as criminal conduct, whether such conduct is by positive action or by omission.

Advocates for recognizing direct criminal liability of robots further contend that these machines can bear responsibility for criminal conduct constituting the material element of a crime, including omissions.

Moreover, this approach maintains that the mental element of crimes may also be satisfied in the case of robots, given the advanced cognitive capabilities developed through modern artificial intelligence technologies. These capabilities enable robots to make complex decisions and learn from prior experiences, thereby raising questions about the possibility of attributing intent or criminal purpose to them²⁰.

2.2.2- The Approach Rejecting Criminal Liability of Robots

By contrast, the second school of thought denies the possibility of holding artificial intelligence devices, such as robots, criminally liable. This position rests on several key arguments, most notably:

- The difficulty of attributing a criminal act personally and directly to a robot.
- The conflict between imposing criminal liability on robots and the philosophy underpinning criminal punishment, which is founded on deterrence²¹, retribution, and rehabilitation — objectives that cannot be realized in relation to a non-human entity incapable of experiencing pain, shame, or possessing legal free will in the traditional sense.

This approach further suggests that incidents caused by artificial intelligence systems should be treated as unavoidable events that fall outside the concern of criminal law. This view is exemplified in the opinion expressed by researcher Daphne Lima, who states:

“In cases of malfunction, it is proposed that humans learn to live with such unfortunate developments, in much the same way as they have learned to live with the consequences of a bridge collapsing due to a hurricane or a tyre bursting, leading to a car accident. Not everything can be anticipated, prevented, or controlled, and in daily life, there are many situations where no one bears blame — nor can anyone be held criminally liable — for an undesirable outcome.”

This path is, of course, the most straightforward to implement in practice and is sometimes considered the fairest approach, particularly in situations where an incident involving artificial intelligence does not result from human fault or criminal intent. For instance, in the context of self-driving vehicles, serious accidents may occur for which no specific party can be held responsible. Nonetheless, there remains significant concern in specific cases where a third party is injured or

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

killed by an artificial intelligence system, as the absence of any form of liability in such circumstances may appear shocking and legally unacceptable²².

It is noteworthy that the majority of legal systems, including Algerian legislation, condition criminal liability on the commission of criminal conduct by a natural person. As a general principle, criminal accusations are directed against natural persons, given that the provisions of criminal law were originally designed to regulate human behavior. However, it must also be acknowledged that legal persons have, in modern legal systems, become subject to criminal liability under specific conditions and legal frameworks established by contemporary legislation.

Conclusion:

With the rapid developments in artificial intelligence and its growing integration across various fields, it has become imperative for legal systems to adapt to these profound changes, especially in criminal law, which now faces unprecedented challenges posed by crimes committed through entities lacking human will, such as AI systems and intelligent robots.

This analysis has revealed a legislative gap in Algerian law regarding crimes resulting from the use of artificial intelligence. This gap places a significant burden on the judiciary, which must interpret existing legal provisions and attempt to apply them to emerging situations, despite the practical and legal difficulties this entails.

Based on the findings of this study, we recommend the following:

- Enacting specific Algerian legislation addressing artificial intelligence crimes, clearly defining criminal liability for both natural and legal persons involved in the development or operation of these systems.
- Considering the possibility of granting AI systems a limited legal personality, allowing for at least civil liability and ensuring compensation for victims of damages caused by errors or malfunctions.
- Increasing public awareness of the risks associated with the misuse of AI technologies and providing specialized training for legal and judicial professionals to address the unique challenges posed by such crimes.

The true challenge today lies in striking a balance between protecting society from the risks of artificial intelligence and harnessing its immense benefits, without hindering the path of innovation and technological progress.

Citations:

¹ Among the international initiatives related to artificial intelligence are: the Budapest Convention on Cybercrime (2001), the OECD Principles on Artificial Intelligence (2019), the UNESCO Recommendation on the Ethics of Artificial Intelligence (2021), and the United Nations Global Initiative on Artificial Intelligence.

² Boukhenouche, Amal. "The Term 'Crime' in Algerian Criminal Law Between Wording and Concept." *Al-Hikma Journal of Islamic Studies*, vol. 08, no. 01, 2021, pp. 24-43.

³ Ibrahim, Ali Ahmed. "Applications of Artificial Intelligence in Combating Cybercrimes." *Al-Qanuniyya Journal*, Cairo University, no. 08, 2021, p. 281.

⁴ Sigsworth, Romi. "Artificial Intelligence: What Implications for Organized Crime in Africa?" Published on 11 July 2023 in ISS Today. Available at: <https://issafrica.org/fr/iss-today/intelligence-artificielle-quelles-implications-pour-le-crime-organise-en-afrique>.

Accessed on 07 June 2024.

" Can the Robot Be Punished? Criminal Liability in the Face of Artificial Intelligence"

⁵Ellyson, Laura. "Criminal Liability and Artificial Intelligence: Some Avenues for Reflection." *Les Cahiers de propriété intellectuelle*, vol. 30, no. 03, p. 890.

⁶Al-Saddawi, Mostafa. "People Killed by Intelligent Robots." *Sayidaty*, 19 April 2018, <https://www.sayidaty.net/node/701311>. Accessed 09 May 2024 at 10:23.

⁷ Ibid.

⁸ Ibid.

⁹ <https://www.aljazeera.net/tech/2018/3/20>, Accessed on 07 June 2024.

¹⁰ Motaz Hamdallah Abu Sweilem, *Criminal Liability for Potential Crimes*, Master's thesis in Public Law, Faculty of Law, Middle East University, 2014, p. 26.

¹¹ Reasons for Justification (Causes of Justification)

The grounds for justification are defined in Article 39 of the amended and supplemented Algerian Penal Code as follows:

- If the act was ordered or authorized by law.
- The state of legitimate self-defence or defense of others, or property belonging to oneself or others (provided that the act is proportionate to the aggression).

¹² See Articles 47, 48, and 49 of Ordinance No. 66-156 of 18 Safar 1386, corresponding to 8 June 1966, containing the Penal Code, as amended and supplemented by Law No. 21-14 of 23 Jumada al-Awwal 1443, corresponding to 28 December 2021, Official Gazette dated 29 December 2021, No. 99, p. 5.

¹³ Laura Ellyson, op cit, p890.

¹⁴ Ahmed Keilan Abdallah, Mohamed Awni Alfet, "Criminal Liability for the Use of Robots – A Comparative Study –," *Al-Farabi Journal of Human Sciences*, vol. 02, no. 02, 2023, p. 11.

¹⁵ Laura Ellyson, ibid, p879.

¹⁶ Mahmoud Mohamed Soueif, *Artificial Intelligence Crimes (The New Criminals)*, Dar Al-Jameaa Al-Jadida, Alexandria, 2022, p. 144.

¹⁷ Ahmed Kilane Abdallah, Mohamed Aouni Alfet, op. cit., p. 16.

¹⁸ Laura Ellyson, ibid,p888.

¹⁹ Ahmed Al-Kailan Abdullah, Mohamed Aouni Alfet, op. cit., p. 16.

²⁰ Gabriel hallevy, The criminal liability of artificial intelligence entities, from science fiction to legal social control, the university of Akrom, *Akrom law journals*, 2016, p171.

²¹ Ahmed Al-Kailan Abdullah, Mohamed Aouni Alfet, op. cit., p. 09.

²² Laura Ellyson, ibid, p892.