

# Difficulties and Countermeasures in Defining Criminal Subjects in Criminal Law in the Era of Artificial Intelligence

Zhuoying Ding

City University of Hong Kong, Hong Kong, China

## Abstract

**With the development of artificial intelligence constantly intervening in various aspects of social life, numerous legal issues arising from it urgently need to be resolved. Issues such as whether artificial intelligence has the qualification as a criminal subject in the field of criminal law have become increasingly controversial. The article reflects on the dilemma of artificial intelligence as a criminal subject, such as the lack of corresponding legal personality, uncertainty of subjective elements, and the lag of criminal law; Through typical examples such as autonomous driving, smart healthcare, and finance, demonstrate the shortcomings of current criminal law in terms of artificial intelligence accountability; Proposed the establishment of a responsibility system centered on "people", combining the absorption of legal agency ideas, exploration of innovative responsibility methods, and enrichment of pre-existing norms to form a diversified responsibility system. At the same time, it is necessary to establish legislative innovation and institutional improvement under artificial intelligence technology while adhering to the basic principles of criminal law.**

## Keywords

**Artificial Intelligence; Criminal Subject; Adaptability of criminal Law; Attribution System; Legal Innovation.**

## 1. Introduction

We are currently in the midst of a development boom in artificial intelligence technology. Autonomous vehicles, automated diagnostic healthcare, financial algorithm trading, and smart homes are all mapped to various aspects of people's lives. It is estimated that by 2023, the size of the artificial intelligence market will reach approximately \$200 billion and expand at a rate of over 30% per year. This wave of technological development has not only brought more and more convenience to people, but also brought unprecedented impact to modern laws built on the foundation of industrial civilization[1].

The issue of "criminal subjects" is facing unprecedented challenges in the criminal law, which is the last line of defense for regulating social order and maintaining social security. Traditional criminal law defines the perpetrator of a crime as a natural person who has the ability to bear responsibility for committing harmful acts, and later expands to a legal entity[2]. However, artificial intelligence systems with autonomous learning and decision-making capabilities are different from natural persons and traditional legal entities, and their unique legal status is hindered by the current criminal law legislative system in the face of AI crimes.

In January 2022, a doctor in the United States believed too much in artificial intelligence diagnosis, resulting in the death of a patient and facing a medical malpractice lawsuit as a result; In April 2022, a autonomous vehicle in Germany made a wrong judgment when encountering a complex situation, and thus a traffic accident occurred. All of these pose real challenges to the application of criminal law in the era of artificial intelligence. Therefore, it is necessary to

discuss from the perspective of criminal law whether artificial intelligence can qualify as a criminal subject, which has important practical significance.

Next, the author will take whether artificial intelligence can become a criminal subject in criminal law as the starting point, and based on summarizing the issue of attribution in current relevant scenarios, provide some inspiration for the improvement of the current criminal law system[3].

## **2. Specific legal challenges faced by artificial intelligence as a criminal subject**

### **2.1. Theoretical dilemma of legal personality recognition**

In China's criminal legislation, the establishment of criminal subject qualification has always been considered based on the concept of "legal personality". The possession of consciousness freedom and moral autonomy is the basic prerequisite for the establishment of legal personality, and it is also the main problem for artificial intelligence to obtain legal personality in legislation.

On the one hand, the establishment of the legal status of individuals is based on their rationality. Kant believed in the philosophy of law that any rational being has the ability to legislate for itself, which is the philosophical foundation of the legal person. The mainstream theory of criminal law accepts this idea and regards freedom as the basis for attribution. However, the latest development of artificial intelligence has always relied on algorithms and data as the basis for intelligent assistance or decision-making, which is detached from true conscious behavior and value judgment. Taking intelligent decision-making in automatic driving as an example, whether it is a rule-based auto drive system or a auto drive system based on deep learning, no matter how complex the program design is, it only completes the algorithm program with specific rules, rather than value judgment[4].

Firstly, from the perspective of legal technology, artificial intelligence is difficult to meet the requirements of criminal subjects. The Criminal Law of our country clearly requires that the criminal subject must be a natural person or a legally established unit with corresponding criminal responsibility capacity and corresponding criminal responsibility age. However, artificial intelligence has no biological age or independent property capacity, and cannot bear traditional punishments such as freedom punishment. If we consider giving it a pseudo legal personality similar to that of a legal person, there are also theoretical difficulties: the establishment of the legal personality of a legal person depends on the collective will behind it, while the "will" of artificial intelligence comes from algorithmic calculations, and the essential difference hinders the application of simple analogies[5].

### **2.2. Practical difficulties in determining subjective elements**

The subjective aspect is an important component of confirming a crime. Intention and negligence, as the two mainstream states of subjective composition, cannot be separated from the recognition of the perpetrator's inner activities. Due to the emergence of artificial intelligence, the identification of this subjective state has become even more difficult to discern.

In addition, from the perspective of intentional criminal determination, artificial intelligence programs do not have true "criminal intent". For example, for intelligent investment advisory programs, once their system causes heavy losses to investors due to algorithm errors, this is an objective result of criminal behavior. However, the intelligent investment advisory program itself does not have subjective intent to illegally possess or harm the interests of others. The "program decision-making" of robo advisors is not "intentional" in the sense of criminal law[6].

In addition, artificial intelligence negligence crimes also face the issue of determining the duty of care due to their specificity. Traditional negligence is defined as the duty of care based on the concepts of "should have foreseen but not encountered" or "anticipated but easily believed and

avoided". However, artificial intelligence does not rely on human cognitive level when reasoning and cognition. Therefore, "should have foreseen but not encountered" and "anticipated but easily believed and avoided" seem to be at a loss. For example, medical AI may recognize pathogenic situations that human doctors cannot see after deep learning, but may overlook the importance of the symptoms only due to the missing characteristics of pathogenic factors in some training data. In this case, how to define the scope of AI's duty of care is worth exploring when applying the law[7].

### **2.3. The contradiction between legal lag and technological development**

The written law criminal law, with its certainty and conservatism, often has a significant lag in adaptability to the impact of new technological iterations and upgrades on criminal law. The contradiction between the exponential growth rate revolution presented by artificial intelligence technology and the systematic evolution and improvement of traditional criminal law has gradually become prominent.

Secondly, lag. Firstly, legislative techniques. The current criminal law provisions are designed for human behavior patterns and cannot address the behavioral characteristics of artificial intelligence. As stipulated in the Criminal Law, "behavior" allows individuals to have direct control, but strong artificial intelligence has strong decision-making abilities that may exceed the direct control of the creator. For example, the causal relationship in criminal law is determined based on the experiential connection between humans and outcomes, while the causal relationship in the "black box" of artificial intelligence is difficult to judge[8].

Thirdly, the expansion of criminal law interpretation has become increasingly evident. When dealing with emerging artificial intelligence crimes, judicial authorities often use analogical interpretations of criminal law, based on their understanding of the nature of new criminal acts, to give new or incomplete meanings to certain provisions of criminal law. Although it can quell criminal issues in individual cases, it may theoretically undermine the standardization and clarity of criminal law. If problems with artificial intelligence products are attributed to the negligence of AI designers, it will confuse the boundary between product liability and criminal liability.

## **3. Realistic dilemmas: attribution challenges in different contexts**

### **3.1. Responsibility dilemma in the field of transportation**

With the rapid advancement of autonomous driving technology, the original way of defining liability for traffic accidents has been greatly tested. According to the automatic driving classification standards released by SAE, when a car with L3 level or above functions truly achieves a state of no human operation in a certain application scenario, it can be considered that the driver's subject has undergone a significant change[9].

For example, in a typical accident in 2022, the L3 level autonomous driving of a certain intelligent driving car was driving on a highway and failed to recognize a certain type of obstacle on the road, resulting in a rear end collision with the following car. After investigation, it was found that there was an algorithm error in the decision-making process of the system, but it does not entirely belong to the car owner (the car was in a regulated state rather than a driving state at the time of use), nor can it be fully attributed to the manufacturer (this scenario has not occurred in existing autonomous driving tests). This example reveals the uncertainty of the responsible person in the context of autonomous driving.

How should the responsibility for accidents caused by human failure to respond correctly within a limited time when the decision-making power of human-machine interaction (driving mode switching power) is transferred to humans by autonomous vehicles? Under the existing system of traffic accident liability allocation, it is difficult to adjust the rights and obligations

between the new human-machine interactive liability subjects through established norms and clauses.

### **3.2. The problem of attribution in the field of healthcare**

The development of AI in medical diagnosis, surgical assistance, and other fields is becoming increasingly common, and the issue of medical responsibility for medical artificial intelligence is gradually receiving attention. In 2023, a tertiary hospital in China used an AI lung nodule diagnostic tool to misdiagnose lung nodules in approximately 310 patients. Some patients missed the best window for timely treatment due to the misdiagnosis by the "AI doctor". This has also raised questions about whether "AI doctors" should be held responsible for diagnosis and treatment errors.

The traditional subject of doctor-patient disputes is clear, namely hospitals/doctors and patients, but with artificial intelligence robot systems, the responsibility chain is not clear. For example, even though the patient has a confirmed diagnosis, the doctor still insists on listening to the incorrect treatment advice provided by the intelligent robot. Who is the one who should bear the burden of proof at this point? Are they doctors who rely solely on artificial intelligence robots, workers who develop artificial intelligence robots, or medical institutions that purchase artificial intelligence systems? In reality, it is a situation of multiple parties shirking responsibility.

More noteworthy is the difficulty in applying the principle of fault liability brought about by the "black box" transformation of AI when it has undergone sufficient data learning to completely exceed the intended use of the developer, resulting in a medical liability that exceeds the original design and current regulations are insufficient to predict its completely new use.

### **3.3. Systemic risk challenges in the financial sector**

The deep application of AI in the financial field has brought both new changes in financial service models and new systemic risks. The changes brought by AI applications such as algorithmic trading, intelligent investment advisory, and risk control systems have not only improved efficiency but also created new risk transmission chains.

In 2022, a securities firm's intelligent trading system experienced algorithm vulnerabilities that led to a "flash crash" phenomenon, causing severe fluctuations in securities trading prices. The intelligent system of the securities firm is based on an algorithm of an untested machine learning model, which generates effect transmission in specific market environments. The characteristic of intelligent trading systems is that they are difficult to terminate once they are in operation, and when there are flaws in the intelligent and automated system, it is difficult for regulatory authorities to intervene.

These two incidents have revealed two new risks. Firstly, the reaction time of financial AI itself is faster than that of humans, making traditional financial regulation and control based on human risk control logic somewhat inadequate; Secondly, complex interactions between different artificial intelligence can also occur, resulting in certain systemic risks that exceed the financial institutions' ability to bear their own risks[10].

### **3.4. Responsibility gaps in personal information processing**

The implementation of the Personal Information Protection Law has strengthened the requirements for data compliance. With the increasingly widespread application of artificial intelligence technology in the field of personal information, it will further bring new problems to the determination of relevant responsibilities.

For example, in the process of using algorithms to collect and use user information on a certain e-commerce platform, due to design flaws in the algorithm, some user information is collected and used beyond the scope. At this time, the responsibility for the algorithm should be

attributed to the algorithm designer, the data controller, or the platform provider. There is currently no corresponding legal definition.

More seriously, deep learning algorithms may cause the collection and use of personal information to exceed the authorized scope at the time of collection, and this "drift" phenomenon has become a dilemma faced by traditional informed consent rules in the field of data security. Due to the difficulty for users to understand the complex content of the algorithm model, it is also difficult to give consent in this situation.

## **4. Exploring solutions: building a multi-faceted responsibility system**

### **4.1. Adhere to the core responsibility of 'people'**

Based on this, the author believes that "upholding the responsibility of 'people'", although just a slogan, has the following theoretical and practical values in the face of the "rapidly developing AI technology" itself: firstly, it does not conflict with the basic principles of criminal law and is in line with the basic functions of criminal law; Secondly, the stability of criminal law can promote its healthy development.

Firstly, artificial intelligence should be regarded as a tool. The tool attribute is an indispensable and important attribute of artificial intelligence. It cannot be denied that artificial intelligence can exercise its own artificial intelligence behavior. However, no matter how advanced the development of artificial intelligence is, it is still essentially a tool manufactured and used by humans. As Germany's Roxin said, "Even the most complex technology is not an object with responsible behavior ability. Classifying artificial intelligence as a tool can effectively avoid unlimited liability expansion.

The second is to implement the principle of graded responsibility allocation. There should be different principles of responsibility allocation for different types of artificial intelligence. Artificial intelligence for human assisted decision-making should strengthen human review responsibility; Artificial intelligence for autonomous decision-making should strengthen the responsibility of designers and regulators. For example, in the field of medical artificial intelligence, the professional review responsibility of physicians for the diagnostic recommendations provided by artificial intelligence should be commensurate with the credibility and interpretability of the AI.

Finally, refine the rules for the responsible parties behind it. On the basis of establishing the duty of care for designers, producers, users, and regulators, clarify the responsibility chain, and separately legislate for high-risk autonomous driving to clarify the duty of care for each node.

### **4.2. Introducing the concept of "legal representation" or "guardianship"**

For highly autonomous artificial intelligence devices, a responsibility system for artificial intelligence can be constructed based on the principles of statutory agency and guardianship in civil law. This not only does not deny the uniqueness of artificial intelligence devices, but also facilitates the inheritance of the responsibility system.

In terms of specific institutional design, we can draw on the guardianship system for minors and establish a systematic "guardian" system within the artificial intelligence system. The owner or controller of the system bears a responsibility similar to that of a "guardian", responsible for the AI behavior of the system under their own control. The size of its guardianship responsibility should be based on the principle of benefit, and the more people benefit from the use of artificial intelligence, the more risks they should bear in the use of AI systems.

One is to recognize the "technology agent" system and view highly autonomous AI systems as an extension of their ownership technology. The legal effect of AI systems causing harm through autonomous operations within their authorized scope should belong to the owner. The

establishment of the "technical agent" system is not only in line with the technical attributes but also facilitates the relief of victims.

### **4.3. Exploring new forms of legal liability**

To establish a new form of responsibility, it is necessary to innovate new technological forms of responsibility that adapt to the development of artificial intelligence, while fully considering the compatibility with modern criminal legal concepts and criminal technology characteristics, that is, to innovate the development of artificial intelligence technology for legal governance.

Firstly, it is possible to consider establishing an "algorithm liability" system, where the designer of the algorithm should bear special and strict liability for damages caused by algorithm defects. The criteria for establishing such liability should be based on the algorithm security requirements and industry standards mentioned above, and should take into account the phased characteristics of technological progress.

Secondly, refine technical sanctions. For security AI products with potential risks, technical sanctions can be implemented, such as mandatory product upgrades, reduction of product functions, interruption of system usage, etc. This is not only conducive to achieving technical risk prevention and control, but also in line with the technical characteristics of artificial intelligence. For example, for autonomous driving products, a "technology ban" system can be established, and when there are significant defects, they can be forcibly withdrawn from the market.

The third is to establish an artificial intelligence liability insurance mechanism. A multi-level sharing mechanism can be constructed by combining compulsory insurance with commercial insurance, especially in high-risk areas such as autonomous driving and medical AI, where industry wide liability insurance funds should be established to ensure timely compensation for victims.

### **4.4. Promoting Pre-emptive Legal Norms**

Before good governance, it is the source and prerequisite legal norms of risk prevention and control. The establishment of good laws for artificial intelligence should balance risk prevention and technology neutrality.

One is to establish a grading management system for artificial intelligence products. Establish a system for implementing differentiated supervision based on the level of system autonomy and the risk level of application areas, implement an admission system for high-risk AI systems, and mandate safety certification and ethical review.

Secondly, establish requirements for algorithm registration and transparency. Establish a filing and review mechanism for algorithms in key areas, requiring algorithm developers to explain the basic principles of algorithms, risk control measures, etc. Propose interpretability requirements of different levels based on the risk level of the scenario.

Thirdly, comprehensive supervision. Supervise the entire lifecycle of AI systems from design and development, deployment and operation to retirement and termination. Especially in system upgrades or algorithm updates, a strict change management system should be established.

Fifth, accelerate the formulation and improvement of industry standards and technical specifications. Establish and improve a multi-level standard system for the healthy development of artificial intelligence, including industry standards, group standards, etc., and timely formulate and improve relevant standards that reflect technological progress.

## 5. Conclusion

Artificial intelligence poses a transformative challenge to criminal law. We must acknowledge, rather than resist, the change and conduct systematic construction and research based on adherence to the principles and doctrines of criminal law.

In summary, this article believes that there are two major path dependence dilemmas in the qualification determination of artificial intelligence criminal subjects: substantive law (how to attribute responsibility in specific scenarios and different types) and procedural law (attribution of legal personality and subjective constitutive elements). The theoretical dilemma lies in the lack of legal personality, difficulty in proving subjective elements, and the inability to achieve attribution of subjective elements; The actual dilemma arises from the differences in specific scenarios and types of behavior, and highlights the shortcomings of the legal framework itself. The solution to difficulties cannot be achieved without the joint research of law, judiciary, and theoretical studies.

Another way to solve the above problems is to establish a comprehensive multi-level and collaborative responsibility system based on the concept of legal agency, new types of responsibility forms, and pre-existing norms, with the "responsible subject as the axis". It should be noted that the legal system should be innovated based on relevant technical rationality and legal analysis, that is, it cannot produce "technical blindness" or "legal blindness".

Of course, in the future, with the further development of technologies such as general artificial intelligence, there will be some new contradictions in our criminal technology, which requires a more open and forward-looking legislative spirit. On the basis of maintaining existing legal stability, there should be room for its appropriate and sufficient development, so as to truly realize the legal order value of artificial intelligence in the new era - "good law".

## References

- [1]Feng W . The Dilemmas and Solutions in the Application of Criminal Law to Property Crimes in the Age of Artificial Intelligence [J]. Economics, Law and Policy, 2024, 7 (2): DOI: 10.22158/ELP.V7N2P125.
- [2]Shen Y . Research on Criminal Risks in the Age of Artificial Intelligence [J]. Journal of Artificial Intelligence Practice, 2024, 7 (2): DOI:10.23977/JAIP.2024.070208.
- [3]I. A S . Problems of qualification of murder by perfect group methods (item "G" part 2 of article 105 of the Criminal Code) [J]. Yugra State University Bulletin, 2023, 19 (1): 23-28. DOI:10. 18822 /BYUSU20230123-28.
- [4]Zhang D ,Li P . The Challenges and Response of Criminal Proceedings in the Era of Artificial Intelligence [J]. Management, 2022, 5 (3): DOI:HTTPS://DOI.ORG/10.31058/J.MANA.2022.53004.
- [5]Tarkhanov I ,Gayfutdinov R . The Subjective Scope of Crime and Its Signs in Criminal Law [J]. Uchenye Zapiski Kazanskogo Universiteta Seriya Gumanitarnye Nauki, 2020, 162 (2): 161-176. DOI:10.26907/2541-7738.2020.2.161-176.
- [6]Елена Анатольевна Пашенко. MORALITY AND CRIMINAL LAW CRIMINALIZATION OF THE ACTIONS OF THE SUBJECT CRIMINAL ASSISTANCE TO SUICIDE [J]. NORTH CAUCASUS LEGAL VESTNIK, 2019, 1 (4): 143-147. DOI:10.22394/2074-7306-2019-1-4-143-147.
- [7]V. S E . SUBJECT OF BRIBERY UNDER THE MODERN CRIMINAL LAW OF RUSSIA [J]. Juridical Journal of Samara University, 2018, 4 (4): 159-159. DOI:10.18287/2542-047x-2018-4-4-159-165.
- [8]CERNOMOREȚ S ,HORTOPAN M D . The analysis of the subjective side of the crime provided by Article 264 of the Republic of Moldova's Criminal Code [J]. Relații Internaționale Plus, 2017, 1 (11): 341-351.
- [9]Malcher S D F ,Deluchey Y F J . Crime and Subjectivity: Reflections on the Discourses and Practices of Legitimizing Discriminatory Exercise of Criminal Law [J]. Revista de Criminologias e Politicas Criminas, 2016, 2 (2): 99-116. DOI:10.26668/INDEXLAWJOURNALS/2526-0065/2016.V2I2.1458.

- [10]Kuznetsova I N . The Subject of the Crime Provided by Article 2581 of the Criminal Code of the Russian Federation [J]. Vestnik Omskoj Ūridičeskoj Akademii, 2016, (32): 54-57. DOI:10.19073/2306-1340-2016-3-54-57.