International Journal of Law
www.lawjournals.org
ISSN: 2455-2194
Received: 15-12-2025, Accepted: 14-01-2026, Published: 29-01-2026
Volume 12, Issue 1, 2026, Page No. 198-205

Legal framework in criminal liability for crimes involving artificial intelligence in the European Union, The United States, And China

Doan Thi Thu Hien, Trinh Khanh Ngoc, Hoang Huu Phu
Department of Law, Hanoi Law University, Hanoi, Vietnam

Abstract
The rapid development of artificial intelligence (AI) in the context of the Fourth Industrial Revolution has profoundly transformed social activities while simultaneously creating new risks of criminal misuse. Beyond its economic and technological benefits, AI has increasingly been employed in fraud, misinformation, cybercrime, market manipulation, and the operation of autonomous systems. The growing capacity of AI systems for self-learning, adaptation, and automated decision-making challenges traditional criminal law concepts of conduct, fault, and causation, and complicates the identification and regulation of crimes involving AI. Against this backdrop, this article seeks to clarify the concept and key characteristics of crimes involving AI, with particular attention to the role of AI as a tool, intermediary, or risk-enhancing factor in criminal conduct. Adopting a comparative legal perspective, the study examines the legislative frameworks and enforcement approaches adopted in the European Union, the United States, and China, three jurisdictions that represent distinct regulatory philosophies and legal traditions. Through this analysis, the article aims to provide a structured basis for understanding how criminal liability is addressed in the context of AI-related risks.

Keywords: Artificial intelligence, crimes involving artificial intelligence, criminal liability, criminal, EU, U.S, China

Introduction

The rapid development of AI has brought about profound transformations across virtually all areas of social life, while also giving rise to new forms of criminal conduct and altering the methods through which traditional crimes are committed. The misuse of AI to generate and disseminate false information, perpetrate fraud, conduct cyberattacks, or infringe upon privacy has posed significant challenges to criminal justice systems worldwide. At the same time, existing criminal law frameworks, which are primarily designed to attribute criminal liability exclusively to human actors, have revealed notable limitations and lack adequate mechanisms to effectively address crimes involving AI.

At the international level, the 2001 Budapest Convention on Cybercrime, although adopted prior to the widespread development of AI technologies, continues to serve as a fundamental legal basis for the criminalization of cyber-related offenses and for international cooperation in combating crime in cyberspace. Building upon this foundation, a number of jurisdictions have enacted or amended legislation to regulate various aspects of AI, particularly in areas such as cybersecurity, data protection, and the legal responsibilities of actors involved in the development, deployment, and use of AI systems.

In this context, the European Union, the United States, and China have emerged as three representative models with distinct approaches to establishing legal frameworks for criminal liability in relation to crimes involving AI, reflecting differences in legal traditions, policy priorities, and regulatory philosophies. Against this backdrop, this article seeks to clarify the concept and key characteristics of crimes involving AI, analyze the legislative approaches adopted by the European Union, the United States, and China, and identify common trends and challenges in adapting criminal law to the rapid advancement of artificial intelligence in the digital era.

Overview about Crimes Involving Artificial Intelligence

The rapid development of AI has profoundly transformed the ways in which humans produce, communicate, and make decisions, while simultaneously creating new tools for unlawful conduct. AI is no longer merely a neutral instrument; rather, it may participate in the formation, execution, or amplification of the harmful consequences of criminal acts.

Numerous studies have demonstrated that AI can be used by humans to commit crimes, or that AI systems themselves may generate socially dangerous outcomes. The first experiment illustrates the use of AI as a tool for sending large-scale phishing messages on social media platforms. These messages are personalized based on users' behavior and publicly available information in order to conceal the offender's true intent. When victims click on fraudulent links and provide data on fake websites, the system can collect personal information for the purposes of theft or fraud [1]. The second experiment simulates an electronic market and shows that AI-driven trading agents are capable of learning autonomously and coordinating with one another to manipulate the market by placing fictitious orders and engaging in deceptive trading strategies, thereby creating artificial price fluctuations for illicit profit and posing serious threats to market order and economic security [2]. These experiments indicate that AI does not merely function as a passive tool assisting human offenders; under certain conditions, it may operate autonomously and participate directly in the execution of socially dangerous conduct.

Scholars have also warned of the growing trend toward the "weaponization" of AI. AI is increasingly used to enhance the capabilities of cybercriminals, enabling offenders to employ machine learning programs to carry out fraud, financial scams, system intrusions, and the impersonation of voices or images of relatives in order to misappropriate property [3]. The criminal use of AI is reshaping the landscape of threats to social safety in three main ways: (i)

amplifying traditional crimes through large-scale automation; (ii) creating entirely new modes of attack, such as the coordination of autonomous drone swarms or the production of highly persuasive synthetic content; and (iii) blurring the boundary between cybercrime and offline crime, as harmful conduct may originate in the digital environment while producing direct physical consequences in the real world [4].

These developments demonstrate that the involvement of AI has altered the structure of criminal conduct. Whereas crime was traditionally defined primarily by human intent and action, criminal processes may now occur through interactions between humans and automated systems. In some cases, crimes are committed entirely through intentional human conduct, with AI serving merely as a tool. In other cases, AI acts as an intermediary that increases the level of danger of the conduct or produces unintended consequences due to its capacity for autonomous decision-making.

Accordingly, crimes involving AI may be understood as violations of criminal law in which AI systems play a significant role by participating in, influencing, or materially affecting the formation, execution, or consequences of the offense. This neutral approach does not presume AI to be a criminal subject, but instead acknowledges its practical role within the structure of contemporary criminal conduct. It thereby enables criminal law to address emerging phenomena without undermining traditional principles concerning criminal capacity and liability.

Crimes involving AI exhibit several key characteristics [5]. First, they involve diverse actors: multiple parties may be connected to the offense, including not only the direct user but also those who design, program, provide data for, or operate AI systems. Second, the conduct is sophisticated and difficult to control, as AI systems may operate autonomously, learn, and modify their behavior, rendering AI-enabled crimes faster, more covert, and more difficult to detect than traditional crimes. Third, the attribution of fault and causation is particularly complex, as the boundary between human fault and system-generated outcomes is often unclear. Fourth, the consequences are widespread and transnational, given that AI operates in a global digital environment where harm may rapidly affect individuals, organizations, and even multiple states.

Based on the role of AI in criminal conduct, crimes involving AI may be classified into three main categories. The first category includes offenses directed against AI systems themselves, such as attacks, unauthorized access, manipulation of data or training models, and acts that compromise system security. The second category consists of offenses that use AI as a tool for crime, where humans intentionally exploit AI capabilities to engage in socially dangerous conduct, including the dissemination of false information, market manipulation, or the control of autonomous devices to cause harm. The third category encompasses offenses arising from negligence in the design, management, or operation of AI systems, where the actor lacks criminal intent but serious harm results from failures to comply with technical safety procedures, testing, certification, or supervision.

In sum, crimes involving AI do not constitute an entirely new category of crime, but rather reflect the transformation of traditional forms of criminality under the influence of intelligent technologies. This development compels criminal law to expand its analysis of the constituent elements of crime in order to adapt to the era of automation, while maintaining a balance between encouraging technological innovation and preventing and punishing criminal conduct.

International Legal Frameworks for Criminal Liability in Crimes Involving Artificial Intelligence

1. The European Union's Legal Framework for Criminal Liability in Crimes Involving Artificial Intelligence

The European Union (EU) is widely regarded as a pioneering jurisdiction in developing a comprehensive legal framework to govern and guide the development of AI. Its regulatory approach is grounded in prioritizing human rights protection and democratic values, rather than pursuing technological growth as an end in itself.

This perspective was codified through the adoption of the Artificial Intelligence Act (AI Act) in 2024, which establishes a horizontal regulatory model. Instead of adopting a fragmented, sector-specific approach, the Act applies general principles such as transparency, accountability, fairness, and safety to all AI systems, irrespective of the field in which they operate. This approach builds upon the normative foundations of the General Data Protection Regulation (GDPR), which had already affirmed the right of individuals not to be subject to decisions based solely on automated processing [6]. On this basis, the AI Act seeks to ensure the governance of ethical values and legal harmonization among Member States in a context where AI increasingly exerts profound societal impact. In addition, the Act adopts a risk-based regulatory model, in which the intensity of legal intervention and compliance obligations is calibrated to the level of risk an AI system may pose to society and fundamental human rights.

The introduction of this regulatory framework was necessary to address legal gaps that existed before the emergence of instruments specifically regulating AI technologies. Before this development, the European criminal law framework was primarily shaped by the Convention on Cybercrime (Budapest Convention) of 2001. Over time, however, this instrument has revealed significant limitations when applied to the context of AI. The Convention obliges contracting states to criminalize certain acts, such as unauthorized access to systems, interference with data, fraud, or violations of the integrity of information systems. In doing so, it establishes a foundational principle that technology functions merely as a means or instrument for the commission of an offence [7]. At the same time, human beings remain the subjects of legal responsibility. This principle clearly reflects the human-centered orientation of criminal law, attributing responsibility to the actor rather than to the technical tool employed.

This shift in regulatory approach has had significant implications for EU criminal policy. Rather than relying solely on ex post punitive measures after harm has occurred, which are often ill-suited to address the automated and complex nature of AI, the EU has moved toward establishing preventive legal mechanisms at an early stage. Although the AI Act primarily relies on administrative regulatory tools, the system of standards relating to safety, transparency, and risk management established therein constitutes a fundamental legal reference framework. On this basis, the content of duties of care and the boundaries of

lawful conduct are progressively clarified, thereby laying the groundwork for the assessment and attribution of criminal liability in appropriate cases, including Crimes involving AI.

Within the EU, the human oversight mechanism provided for in Article 14 of the AI Act plays a central role in establishing mandatory standards of conduct for determining the limits of individual responsibility in the operation of AI systems. Article 14(1) stipulates that "high-risk AI systems shall be designed and developed in such a way, including with appropriate human–machine interface tools, that natural persons can effectively oversee them throughout their period of use." Failure to comply with the obligations set out in Article 14 does not merely constitute an administrative violation but may also amount to a breach of professional rules or safety regulations. Such a breach represents a prerequisite condition for the attribution of criminal liability for negligent offences causing death or bodily injury under the criminal laws of the Member States.

Unlike Article 22 of the GDPR, which approaches the issue from the perspective of protecting the subjective rights of individuals affected by fully automated decisions and grants them the right to request human intervention or review, Article 14 of the AI Act emphasizes the proactive obligations of entities that operate and use AI systems. It requires human involvement throughout the system's lifecycle to ensure effective control and timely intervention. This approach reflects a preventive, ex ante risk-governance mindset rather than an intervention that occurs only after rights have already been infringed.

From a criminal law perspective, Article 14 of the AI Act may be regarded as an essential normative basis for defining the duty of care of actors involved in the design, deployment, and use of AI systems. Accordingly, where serious harm occurs, the failure to implement, or the inadequate implementation of, human oversight obligations under Article 14 may constitute a violation of mandatory professional or administrative rules. This, in turn, satisfies a necessary condition for examining criminal liability for negligent offences that infringe upon life, health, or other legally protected interests [8].

Although the AI Act does not directly criminalize breaches of human oversight obligations, in judicial practice, such commitments may be invoked by criminal courts as "professional rules or safety regulations" to assess negligence on the part of actors involved in AI-related crimes.

At the level of the Members, building on the obligations established at the Union level, several countries have further internalized and concretized these standards by introducing new criminal offences and aggravating circumstances into their domestic criminal laws. Among these, the Italian Republic and the French Republic represent particularly illustrative examples. The prevailing model of criminal liability across European legal systems continues to adhere to the traditional principle that criminal liability may be imposed only on two categories of subjects: natural persons and legal persons.

Italian criminal law has taken a pioneering step by becoming the first EU Member State to implement the AI Act at the national level through the adoption of Law No. 132/2025 in September 2025. This instrument amended and supplemented the 1930 Penal Code by introducing new provisions designed to reflect the specific dangerousness associated with AI technologies. Notably, the introduction of the offence of "Unlawful dissemination of content generated or altered by artificial intelligence systems" under Article 612-quater provides that "any person who unlawfully causes harm to another person by disseminating, without that person's consent, images, videos, or voice recordings that have been falsified or altered through AI systems and are designed to mislead as to their authenticity" shall be subject to criminal sanctions. This provision establishes a clear legal basis for criminalizing the distribution of AI-generated content that infringes upon honour and privacy, with penalties ranging from 1 to 5 years of imprisonment. It reflects a significant shift from a predominantly civil-law-based protection of data and personality rights toward the use of criminal law instruments to address the heightened harmful potential of generative AI abuses. At the same time, the explicit requirement that the content be "designed to mislead as to its authenticity" narrows the scope of the offence, thereby avoiding the over-criminalization of legitimate creative activities or lawful uses of AI.

In addition, paragraph 11-undecies of Article 61 has been introduced to classify the use of AI as an aggravating circumstance in traditional criminal offences. Article 294, concerning violations of citizens' political rights, has also been amended to impose stricter sanctions on acts of information manipulation involving AI. Specifically, it introduces an aggravated offence whereby "any person who uses violence, threats, or deception carried out through the use of AI systems to wholly or partially prevent the exercise of political rights, or to compel another person to exercise such rights against their will, shall be punished by imprisonment from two to six years."

A similar trend can be observed in France, where the criminal law framework governing AI is being progressively refined, most notably through the adoption of the Law on Securing and Regulating the Digital Space (LSREN) in May 2024. The French legal system continues to maintain the parallel attribution of criminal liability to both natural persons and legal persons, particularly in relation to new offences involving the dissemination of AI-related content without the consent of the affected individual. Under this framework, the publication of visual or audio content generated or produced by AI algorithms that reproduces the appearance or voice of an individual without that person's consent constitutes a criminal offence where two cumulative conditions are met. First, the content creates an impression of authenticity, making it difficult for recipients to distinguish it from reality. Second, the content lacks clear indications of its artificial nature.

This provision marks a significant development in French criminal legislative technique, as it represents the first instance in which AI technology is addressed through a specific criminal norm. Rather than relying on traditional offences such as defamation or invasion of privacy, which often entail complex evidentiary burdens, the current approach emphasizes the obligation of technological transparency and directly penalizes deceptive practices concerning the origin of digital content. French criminal law thus does not adopt an absolute prohibition on the use of AI, but instead focuses on transparency requirements and the protection of personality rights. This approach clearly reflects the human-centered principle underlying EU legal policy, namely that technology is recognized only insofar as

it does not infringe fundamental individual rights and interests.

Moreover, France has established an innovative regime of criminal liability for autonomous vehicles through Ordinance No. 2021-443. In particular, the introduction of Articles L.123-1 and L.123-2 into the Highway Code has brought about a fundamental reallocation of responsibility. Criminal liability is exempted from the driver during periods when the automated system exercises control, except for non-compliance with police instructions, while liability is transferred to the manufacturer in the event of an accident. In such cases, the manufacturer may be held criminally liable for negligent homicide, negligent bodily injury, or specific road traffic offences pursuant to Article 121-3 of the Penal Code, provided that fault can be established. This regime not only resolves the question of the liable subject when AI systems are in operation, but also imposes a higher standard of responsibility on technology companies regarding technical safety and risk governance, thereby increasing their potential exposure to criminal liability for breaches of duties of care.

In EU practice, the attribution of legal responsibility for automated systems has moved beyond abstract principles and has been concretized through landmark case law. A notable example is the judgment of the Court of Justice of the European Union (CJEU) in Case C-634/21 (SCHUFA) in 2023 [9], concerning credit scoring activities. In this case, the Court held that the use by a credit information company of algorithms to generate a "probability value" relating to an individual's creditworthiness constitutes a "solely automated individual decision" within the meaning of Article 22 of the GDPR, where that value plays a decisive role in a third party's decision, such as a bank's refusal to grant credit.

This judgment carries significant legal implications, as it rejects the argument that the algorithm provider merely acts as an intermediary supplying information. Instead, legal responsibility is extended beyond the final decision-maker to include actors involved in the design and operation of the data processing system. This approach reflects a consistent trend in EU jurisprudence, in which liability is traced throughout the entire algorithmic processing chain to prevent actors from exploiting the fragmentation of automated processes to evade legal obligations.

Such a chain-based attribution of responsibility may serve as a reference framework for judicial authorities when assessing the fault of actors involved in the design, deployment, and operation of AI systems that cause harm.

2. The United States' Legal Framework for Criminal Liability in Crimes Involving Artificial Intelligence

The United States is one of the pioneering countries in the development of AI and has also been among the earliest to confront legal issues arising from this technology. However, unlike the European Union, the United States has not enacted a comprehensive and unified federal statute governing AI. Instead, it has adopted a decentralized approach to addressing AI-related risks, with regulatory responses distributed across federal law, state law, and executive actions [10].

In this context, the U.S. legal system currently operates under a fragmented, sector-specific and agency-by-agency regulatory model. Under this model, each federal agency is responsible for issuing regulations, guidance, or technical standards for AI systems within its respective jurisdiction, such as in the fields of national defense, commerce, finance, or intellectual property [11]. In parallel, the White House has issued executive orders and soft-law policy instruments to guide the development of safe and trustworthy AI, most notably the Artificial Intelligence Risk Management Framework developed by the National Institute of Standards and Technology (NIST).

This approach clearly reflects the United States' policy orientation toward prioritizing innovation and maximizing the deployment of AI technologies, while avoiding the premature establishment of rigid legal constraints that could impede technological development. Legal intervention is therefore limited to what is deemed necessary and is calibrated according to specific categories of risk. This orientation is exemplified by the Artificial Intelligence Action Plan announced in 2025, which seeks to relax legal constraints that may hinder the development of AI in the private sector [12].

This policy orientation of the United States is also clearly reflected in the criminal law domain. At present, AI is still regarded as a supporting tool rather than an independent legal subject. Consequently, criminal liability continues to be attributed to human actors or legal entities whose conduct, whether direct or indirect, leads to harmful outcomes. Such attribution depends primarily on: (1) the degree of actual control exercised by the actor over the AI system; (2) the actor's legal duty of supervision; and (3) the presence of fault on the part of the actor.

The foundational principle underlying this approach is causation-based liability [13]. Prosecutorial authorities may impose criminal liability only where a causal link between human conduct and the harmful result produced by the AI system can be established, particularly in cases involving negligence.

At the federal legislative level, during the early stage, crimes involving AI were primarily addressed through the application of existing criminal statutes and case law that were not specifically designed for this technology. A representative example is the Computer Fraud and Abuse Act of 1986 (CFAA), which has been applied to conduct involving the use of software, algorithms, or AI models to gain unauthorized access, cause damage, misappropriate data, automatically collect information, or unlawfully seize control of computer systems [14]. The Supreme Court's decision in Van Buren v. United States (2021) clarified the limits of the concept of "unauthorized access" under the CFAA, thereby providing a legal basis for addressing crimes involving AI that fall within this category [15].

In addition to the CFAA, the Electronic Communications Privacy Act (ECPA), together with federal provisions on wire fraud and identity theft, has been employed to address crimes involving AI committed through the unlawful collection of data, model theft, the generation of deceptive content such as deepfakes, fraud, or the misappropriation of property [16].

However, in recent years, the incidence of crimes involving AI in the United States has increased sharply. For example, identity impersonation through deepfake technology alone has reportedly increased by approximately 3,000%, while total losses resulting from fraud facilitated by generative AI are projected to rise from USD 12.3 billion in 2023 to nearly USD 40 billion by 2027 [17]. In response to this alarming trend, the U.S. federal legal system has gradually shifted

from a primary reliance on traditional, non-specialized criminal statutes toward the incremental development of more specific legal frameworks aimed at addressing emerging technological challenges.

The United States responded at an early stage to deepfake technology through the enactment of the Deepfake Accountability Act of 2023. This statute has a broad regulatory scope and requires labeling and watermarking for all AI-generated content, accompanied by severe criminal sanctions of up to five years' imprisonment and fines of up to USD 150,000 for each violation [18]. In addition, the Take It Down Act of 2025 marks the first federal statute to criminalize the non-consensual dissemination of realistic intimate images, including deepfakes, providing for penalties of up to three years' imprisonment and imposing an obligation to remove the offending content within 48 hours [19].

In addition to legislation specifically targeting deepfakes, the United States has also developed institutional and technical governance frameworks for AI, most notably the Federal Artificial Intelligence Risk Management Act of 2023 and the Algorithmic Accountability Act of 2022. These frameworks impose obligations relating to risk assessment, algorithmic transparency, and oversight of the deployment of AI systems. Although they do not directly criminalize violations, such laws play an important role in establishing standards of conduct and duties of care, which serve as a normative basis for determining fault when addressing crimes involving AI.

In addition, the U.S. Department of Justice (DOJ) has indicated that the misuse of AI in the commission of crimes involving AI may be treated as an aggravating factor. Accordingly, the DOJ may seek enhanced penalties in cases where the use of AI renders the offense more difficult to detect or remedy, or results in more serious harm. Where existing aggravating frameworks prove inadequate, the DOJ has stated that it will consider proposing amendments to federal criminal law to address sanctioning gaps related to the misuse of AI [20]. This development demonstrates that the United States has begun to adopt a more stringent criminal policy response to the growing trend of abusing AI to commit crimes involving AI.

At the state legislative level, in the absence of a unified federal statute addressing crimes involving AI, numerous legislative initiatives have emerged, reflecting the distinctly decentralized nature of the U.S. legal system in responding to the challenges posed by AI. Since 2019, crimes involving deepfakes have become a focal point of state-level legislation. Virginia was the first state to amend its criminal law to criminalize the non-consensual distribution of deepfake-generated images and videos of another person, classifying such conduct as a Class 1 misdemeanor punishable by up to one year of imprisonment and a fine of up to USD 2,500 [21]. California subsequently enacted two significant statutes: Assembly Bill 602, which criminalizes non-consensual sexually explicit deepfake images, and Assembly Bill 730, which protects political candidates from manipulated media content within 60 days prior to an election [22]. Following these developments, several other states, including Missouri, New York, Minnesota, and North Carolina, have adopted similar legislative measures. These initiatives are widely regarded as timely responses to urgent risks, aimed at addressing acute phenomena that pose clear and significant dangers to society.

In U.S. practice concerning the application and attribution of criminal liability for crimes involving AI, criminal liability is not imposed on the AI system itself but is instead allocated to human actors or legal entities involved in the design, operation, or deployment of AI, depending on the degree of control exercised, the existence of a causal relationship, and the fault of each actor.

With respect to end users, U.S. judicial practice shows that they are the actors most frequently and directly held criminally liable in cases where AI causes serious harm. For intentional offenses in which AI merely functions as a tool, criminal liability is attributed directly to the user who employed the AI system to commit the crime. A representative example is the case of a former high school athletic director in Maryland who used AI to generate a fake audio recording containing racist and antisemitic content targeting a school principal and was sentenced to four months' imprisonment [23].

By contrast, in cases involving negligent offenses where the AI user does not intend to cause socially dangerous consequences, U.S. courts have reasoned that as long as humans retain a supervisory role or the ability to intervene in the operation of the AI system, the legal duty of care remains with the user. In Kevin George Aziz Riad, Riad was operating a Tesla Model S in autopilot mode when a collision occurred with another vehicle, resulting in the deaths of two passengers in that vehicle. Riad was subsequently charged with vehicular manslaughter. Although the accident occurred while the autonomous driving system was lawfully engaged, the court held that the driver could not transfer full control and responsibility to the AI system. Excessive reliance on automated systems was not regarded as a ground for excluding criminal liability; rather, it could constitute negligence for the purposes of the relevant offenses [24].

Nevertheless, Westbrook argues that users or passive passengers of autonomous vehicles should not be held criminally liable when the self-driving system itself violates the law. From the perspective of the actus reus requirement, Westbrook contends that the human operator does not satisfy the objective element of the criminal conduct committed by an autonomous vehicle. In this context, Westbrook proposes the development of a new form of liability, referred to as products culpability, particularly in the field of autonomous vehicles, in order to attribute criminal liability to manufacturers [25].

With respect to programmers, U.S. criminal law adopts a cautious approach and does not automatically attribute criminal liability merely because an AI system causes harm. Criminal liability may nevertheless arise where there is evidence that the developer intentionally designed or deployed AI to facilitate criminal conduct, or was aware of a serious risk yet deliberately disregarded it by failing to adopt necessary preventive measures. In such cases, AI is treated as a tool for committing the offense, and the developer may be regarded as the principal actor where there is evidence of intentional misuse or conscious disregard of serious risks. A representative example is the case of a Chinese developer who installed a "kill switch" in a company's network system, causing the system to shut down after he left the company, for which he was sentenced by a U.S. federal court to four years' imprisonment. This case illustrates that where a developer intentionally designs and deploys malicious code, the AI system is viewed merely

as an instrument, and criminal liability is attributed directly to the developer [26].

With respect to commercial legal entities, U.S. practice demonstrates continued caution in imposing criminal liability on corporations in situations involving crimes involving AI, particularly where it is difficult to establish the organization's criminal intent. In a 2012 decision by the Minnesota Supreme Court (Glorvigen v. Cirrus Design Corp.), the court dismissed the plaintiff's claim against the aircraft manufacturer, reasoning that "a manufacturer of sophisticated technology, including technology marketed as capable of operating an aircraft in autopilot mode, cannot be held liable solely for failing to provide adequate training to operators on the use of such technology" [27]. Similarly, in the Uber case, prosecutors declined to bring criminal charges against the corporation due to insufficient evidence to establish fault, even though one contributing factor to the accident was Uber's inadequate safety risk assessment procedures.

3. China's Legal Framework for Criminal Liability in Crimes Involving Artificial Intelligence

China is recognized as one of the leading countries globally in the development and application of AI technology. The country has invested heavily in research and development activities, thereby occupying a pioneering position in this high-tech field. China has established a centralized and stringent legal framework aimed at regulating the use of AI technology while protecting national security and social order. China's regulatory strategy is implemented in a phased approach, based on two main pillars: foundational laws and specialized management and regulatory measures, as follows [28]:

First, foundational laws. The legal mechanism for regulating AI in China is built upon three foundational laws, including the Cybersecurity Law of 2017, the Data Security Law of 2021, and the Personal Information Protection Law of 2021, clearly reflecting a legislative approach based on "digital sovereignty" to control risks from AI technology. These three laws form a relatively unified legal framework to control data flows and protect data in the context of AI development; however, they still primarily focus on the data aspect and do not fully regulate other dimensions of AI systems.

Second, specialized management and regulatory measures. At the regulatory level, China has made significant progress through the promulgation of the Provisions on the Administration of Deep Synthesis in Internet Information Services (Deep Synthesis Management Provisions) [29] in 2023, which impose obligations on service platform providers for content generation. Additionally, China issued the Interim Measures for the Management of Generative Artificial Intelligence Services [30] (Generative AI Management Provisions) as the first legal framework to comprehensively regulate generative AI services. This document has a broad scope and imposes stricter obligations and responsibilities compared to previous regulations. Thus, the management and regulatory measures have completed the foundations laid by the basic laws, marking a shift from focusing on data protection to directly regulating AI systems, while clearly defining the legal responsibilities of both providers and users.

China previously published a comprehensive Draft AI Law but recently withdrew this draft from the 2025 legislative agenda [31]. Although there is no specialized law on AI yet, China is gradually building a legal framework for AI governance through policy documents and subordinate regulations. It can be seen that China is implementing a centralized and phased strategy, starting from general security laws, then developing into specialized regulations, and currently aiming toward a unified law that has not yet been enacted. This model prioritizes digital sovereignty and social stability through strict preventive oversight mechanisms, with flexible risk definitions linked to impacts on public opinion [32].

Chinese criminal law does not recognize AI systems as legal entities capable of bearing criminal liability but only recognizes AI as an instrument for committing crimes. Based on current Chinese criminal law, crimes involving AI can be divided into three categories [33]:

First, conduct regulated by existing criminal law through direct application of the Criminal Code via judicial interpretation. For example, using AI to identify image codes to infiltrate accounts and selling "AI fraud tricks" resulted in conviction under Article 285 of the Chinese Criminal Code for providing tools for infiltrating information systems.

Second, conduct not yet fully regulated, which involves traditional crimes but with new AI characteristics that make law application less effective; a typical example is fully autonomous vehicle accidents—if caused by technical defects, the manufacturer may be held responsible, but when the system operates independently, it is difficult to impose criminal liability on the user due to lack of intent; the question of prosecution against the manufacturer/user (regarding product quality or accident consequences) has not been clearly regulated.

Third, new conduct that the scope of law cannot fully cover in the AI era. The objective elements of crimes in the Criminal Code provisions do not encompass new forms of criminal conduct and do not have criminal elements for such new conduct. For example, if an AI-integrated prosthetic arm is hacked and destroyed, causing injury to the user, it would be considered destruction of property if regarded as property, but this does not properly reflect the consequences. If it is considered part of the body, it should be considered an act of causing bodily harm.

In China's legal research community, there are two main viewpoints regarding the criminal legal status of AI. The first viewpoint suggests that, with high-level learning and decision-making capabilities, AI could be considered a subject of criminal liability in the future. Conversely, the second viewpoint, which currently prevails, holds that AI is merely a technical tool created and controlled by humans, lacking independent legal will, and therefore cannot become a subject of criminal liability [34]. Additionally, some argue that strong AI already possesses behavioral capacity and legal responsibility like "humans," can perceive and control its own behavior, and should be considered a subject of crime, while weak AI still operates according to programming, lacks the ability to perceive and control, and should not be considered a subject of crime. Nevertheless, China's current judicial practice and legislation still treat AI as an instrument of crime, and AI itself does not bear criminal liability. Therefore, all crimes involving AI are attributed to human or legal entity subjects, with AI regarded as a non-culpable agent.

Chinese law currently handles subjects who commit crimes involving AI through flexible application of traditional offenses in the Criminal Code, combined with other laws such as the Cybersecurity Law of 2017, the Data Security Law of 2021, the Personal Information Protection Law of 2021, the Deep Synthesis Management Provisions, the Generative AI Management Provisions, etc. These documents regulate AI in specific fields, stipulating prohibited conduct, mandatory obligations, etc., and include a transitional clause to the Criminal Code for criminal prosecution. For example, Chapter VI of China's Cybersecurity Law stipulates that violators (who commit prohibited acts, violate mandatory obligations, etc.) shall be subject to fines, remedial measures, etc., from Articles 59 to 69, while Article 74 provides that if such violations constitute a crime, criminal liability shall be pursued according to law. This means that China does not criminalize new crimes involving AI but rather establishes new mandatory obligations and prohibited conduct in the provisions of specialized laws and documents. If there are violations of these provisions that cause serious consequences and constitute a crime, criminal liability shall be pursued for traditional crimes. This is understood as a tiered liability model consisting of two layers: administrative liability and criminal liability. Starting with the first tier, when a person violates the law, they will be subject to ordinary administrative penalties, but when serious consequences are caused that constitute a crime, the second tier—criminal liability—is "activated," and they will be subject to criminal prosecution. This is specifically demonstrated in actual cases in China, and the handling and identification of criminal subjects is also diverse, as follows:

With respect to end users, since the Deep Synthesis Management Provisions officially came into effect, China has had its first case involving ChatGPT. In May 2023 in Gansu, suspect Hong used ChatGPT to fabricate false information about a catastrophic train accident that killed 9 people and disseminated it on the Baijiahao blog platform, attracting 15,000 views. Although ChatGPT service is banned in China, Hong still found a way to access it through VPN. Gansu police arrested Hong and charged him with the crime of "picking quarrels and provoking trouble" under Article 293 of the Criminal Code, facing a maximum sentence of 5 years [35]. Thus, Hong committed prohibited conduct under Article 6 of the Deep Synthesis Management Provisions as follows: "Deep synthesis services shall not be used by any organization or individual to... disrupt economic and social order... service users are prohibited from using the service to create, copy, publish or disseminate false news information..." and the consequence of this conduct was causing public panic, sufficient to constitute a crime. According to Article 22 of these Provisions: "...If the conduct constitutes a violation of public security management, the public security organ shall impose penalties according to public security management regulations; if it constitutes a crime, criminal liability shall be pursued according to law." This provision does not directly criminalize the violation but establishes a mechanism to handle administrative violations; only when serious consequences are caused that constitute a crime does criminal prosecution under the Criminal Code apply.

With respect to developers, Criminal liability is not limited to end-users but extends to developers, as demonstrated in the "AlienChat" case in Shanghai. Two developers created the AlienChat chatbot and deliberately set up system prompts to bypass ethical barriers, allowing AI to generate vulgar and pornographic content to attract paying users, and they were subsequently prosecuted and arrested in April 2024. Despite defense arguments that pornographic content only arose from interactions between AI and users and therefore the defendants were innocent, the court determined that the two defendants had intentional fault, aimed for illegal profit, and committed the act of producing pornographic content [36] through training and adjusting the system according to the Generative AI Management Provisions. On this basis, the trial court convicted both of "producing obscene cultural products for profit" under Article 363 of the Chinese Criminal Code, sentencing them to 4 years and 1.5 years in prison [37], respectively. Thus, the two defendants committed prohibited conduct when using generative AI technology to create vulgar and pornographic text and images under Article 12 of the Cybersecurity Law and Article 4 of the Generative AI Management Provisions; seriously violated ethical issues stipulated in the Cybersecurity Law, Data Security Law, Deep Synthesis Management Provisions, Generative AI Management Provisions, etc. The violations by the two developers were sufficient to constitute a crime and were subject to criminal prosecution.

With respect to commercial legal entities, according to information on June 13, 2025, the People's Procuratorate of Tongzhou District, Beijing prosecuted four defendants and simultaneously prosecuted an e-commerce company headquartered in Fuzhou for producing copyright-infringing collage images. The Tongzhou District Court of Beijing accepted the case, and in the first-instance trial, the court ruled that the e-commerce company from Fuzhou committed copyright infringement and was fined 100,000 yuan; the defendants were sentenced to varying sentences from one year and six months' imprisonment to control for copyright infringement and fined from 25,000 yuan to 60,000 yuan [38]. This is the first criminal case of copyright infringement using AI in Beijing; this case provides an important precedent for similar cases in the future and fills legal application gaps in this issue. The criminal subjects here are the legal entity (company) and individuals; these subjects committed acts of violating intellectual property rights (copyright), which are prohibited acts under Article 12 of the Cybersecurity Law and Articles 4 and 7 of the Generative AI Management Provisions. Because the conduct was committed for profit, it constituted the crime of copyright infringement in the Chinese Criminal Code and was subject to criminal prosecution.

Conclusion

The rapid development of AI has significantly transformed the nature and mechanisms of criminal conduct, challenging traditional concepts of human agency, fault, and causation in criminal law. Crimes involving AI do not constitute a wholly new category of offenses, but rather reflect the technological evolution of conventional crimes in which AI systems increasingly influence the commission or consequences of unlawful acts.

The comparative analysis of the European Union, the United States, and China reveals distinct regulatory approaches shaped by different legal traditions and policy priorities. The European Union emphasizes a preventive, risk-based framework grounded in human oversight and ex ante

obligations. The United States adopts a decentralized and reactive model, treating AI primarily as a tool and attributing criminal liability based on control, causation, and fault. China employs a centralized and phased approach, integrating administrative regulation with criminal enforcement through a tiered liability mechanism. Despite these differences, all three jurisdictions converge on a core principle: AI is not recognized as an independent subject of criminal liability, and responsibility is ultimately attributed to human actors or legal entities behind AI systems.

Overall, these approaches demonstrate the capacity of existing criminal law frameworks to adapt to technological change, while also revealing persistent challenges in attributing fault and responsibility within complex human–machine interactions. Addressing crimes involving AI therefore remains an evolving task, requiring ongoing doctrinal development and careful balancing between technological innovation and the protection of fundamental legal values.

References

1. King TC, Aggarwal N, Taddeo M, Floridi L. Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. Science and Engineering Ethics,2020:26(1):91–103.
2. Martínez-Miranda E, McBurney P, Howard MJ. Learning Unfair Trading: A Market Manipulation Analysis from the Reinforcement Learning Perspective. Proceedings of the IEEE Conference on Evolving and Adaptive Intelligent Systems,2016:2016(1):103–109.
3. Jeong D. AI Security Threat, Crime, and Forensics: Taxonomy and Open Issues. IEEE Access, Volume 8, Issue 1, 2020:8(1):184562–184579.
4. Brundage M, Avin S, Clark J, *et al*. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. University of Oxford,2018:1(1):3–10.
5. Zhao S. Principle of Criminal Imputation for Negligence Crime Involving Artificial Intelligence. Artificial Intelligence and the Rule of Law, Springer Nature Singapore Pte Ltd,2024:1(1):6–8.
6. Park S. Bridging the Global Divide in AI Regulation: A Proposal for a Contextual, Coherent, and Commensurable Framework. arXiv,2023:2023(1):7–10.
7. Council of Europe. Convention on Cybercrime (Budapest Convention). European Treaty Series,2001:185(1):1–25.
8. Radtke T. Chapter 3: High-Risk AI Systems. Artificial Intelligence and Fundamental Rights: The AI Act of the European Union and Its Implications for Global Technology Regulation,2025:4(1):92–98.
9. Court of Justice of the European Union. Judgment of the Court (Grand Chamber) of 7 December 2023, Case C-634/21 (OQ v Land Hessen). Court of Justice of the European Union Reports,2023:2023(12):1–18.
10. Al-Amiri SS, Muqdad H. The Divergence of Legislative Models in Addressing Artificial Intelligence Crimes: A Comparative Study of the European Union, China, and the United States. Journal of Cultural Analysis and Social Change, 2025.
11. Davtyan T. The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained. SSRN Journal, 2024, 2–3.
12. Pham G. The United States Announces a "Low-Regulation" Strategy for Artificial Intelligence. VnExpress, 2025.
13. Do VC, Pham NT. Theoretical Models and Legislative Experiences on Criminal Liability for Artificial Intelligence and Implications for Vietnam. Law and Development Journal, 2025.
14. Federal Bureau of Investigation. Criminals Use Generative Artificial Intelligence to Facilitate Fraud. FBI Public Service Announcement, 2024.
15. Van Buren v. United States. 141 S. Ct. 1648, 2021.
16. National Security Law Firm. Cyber Harassment Laws and AI-Generated Images Like Deepfakes. National Security Law Firm Publication, 2023.
17. Khalil M. Deepfake Statistics 2025. Deepstrike, 2025.
18. U.S. Congress. DEEPFAKES Accountability Act. H.R. 5586, 118th Congress, 2023.
19. U.S. Congress. TAKE IT DOWN Act. S. 146, 119th Congress, 2025.
20. Sidley Austin LLP. U.S. Department of Justice Signals Tougher Enforcement Against Artificial Intelligence Crimes. Sidley Austin Legal Update, 2024.
21. TechSpot Editorial Team. Virginia Updates Law to Make Deepfake "Revenge Porn" Illegal. TechSpot, 2019.
22. Sierra AD. California Deepfake Laws First in Country to Take Effect. Akin Gump Strauss Hauer & Feld LLP Commentary, 2020.
23. Skene L. Former School Athletic Director Gets Four Months in Jail in Racist AI Deepfake Case. Associated Press, 2024.
24. Krisher T, Dazio S. L.A. County Felony Charges Are First in Fatal Crash Involving Tesla's Autopilot. Los Angeles Times, 2022.
25. Giannini A. United States Report on Traditional Criminal Law Categories and Artificial Intelligence. Association Internationale de Droit Pénal, 2022
26. U.S. Department of Justice. Chinese National Who Deployed "Kill Switch" Code on Employer's Network Sentenced to Four Years in Prison. U.S. Department of Justice Press Release, 2024.
27. Westbrook CW. The Google Made Me Do It: The Complexity of Criminal Liability in the Age of Autonomous Vehicles. Michigan State Law Review,2017:2017(1):97–142.
28. Al-Amiri SS, Muqdad H., The Divergence of Legislative Models in Addressing Artificial Intelligence Crimes: A Comparative Study of the European Union, China, and the United States. Journal of Cultural Analysis and Social Change,2025:10(2):4085–4086.
29. Cyberspace Administration of China. Internet Information Service Deep Synthesis Management Regulations. Cyberspace Administration of China Official Publication, 2022.
30. Cyberspace Administration of China. Interim Measures for the Management of Generative Artificial Intelligence Services. Cyberspace Administration of China Official Publication, 2023.
31. Chau A. Artificial Intelligence Governance in China: A Cautious Roadmap with a Long-Term Orientation. People's Representative Newspaper, 2025.
32. Al-Amiri SS, Muqdad H. The Divergence of Legislative Models in Addressing Artificial Intelligence Crimes: A Comparative Study of the European Union, China, and the United States. Journal of Cultural Analysis and Social Change,2025:10(2):4087.
33. Pang D, Olkhovik NV. Criminal Liability for Actions of Artificial Intelligence: Approach of Russia and

China. Journal of Siberian Federal University. Humanities & Social Sciences,2022(8):1094–1107.
34. Yang S. Subjects Bearing Criminal Liability for Crimes Involving Artificial Intelligence. Jurisprudence, Issue 3, 2025:(3):500–501.
35. Anh V. First Arrest Involving the Use of ChatGPT in China. Lao Dong Newspaper, 2025.
36. Global Times Editorial Board. China's First AI Companion App Case Enters Second-Instance Trial, Sparking Debate on Emotional AI Service Boundaries. Global Times, 2026.
37. Yicai Global Editorial Team. China's First AI Pornography Case Is Adjourned at Second-Instance Trial. Yicai Global, 2026.
38. Zhang XH. Beijing's First Case: Four Individuals Sentenced for Copyright Infringement Using Artificial Intelligence. Guangming Daily, 2025.