Academic Education Publishing House

Empirical Analysis of AI Fraud Crimes and Governance Strategies

Wenbo Fu*, Zhihui Ban
The National Police University for Criminal Justice, Baoding, Hebei, China
**Corresponding Author.*

Abstract: Artificial Intelligence (AI) has turned science fiction into reality, increasingly coming into public view and influencing and changing people's lives and ways of thinking. The barriers to AI usage are gradually being lowered, which raises the risk of misuse, including legal risks such as "deepfakes," data breaches, and face-swapping infringement. AI fraud has emerged as a significant branch of new types of fraud, characterized by greater intelligence, diverse methods, strong deception, easier trust acquisition, and a higher success rate. This article presents an empirical analysis of 25 cases collected from online and practical sources, examining the incidence, types, and implementation carriers of AI fraud. It discusses the challenges in governing AI fraud crimes and proposes governance strategies. These strategies include improving legislation to enhance the legal adaptability and constraints of AI, ensuring the legal and compliant use of AI technology; cultivating specialized talents at the intersection of AI technology and law; strengthening enforcement efforts and improving investigation and evidence collection methods to combat technology with technology; and facilitating efficient collaboration among various departments and industries to form a united front. Together, we can address AI fraud crimes and protect the property safety and legitimate rights and interests of the public.

Keywords: AI Fraud; Crimes; Empirical Analysis

1. Introduction

In recent years, the problem of telecommunications fraud has become prominent, with a high number of cases and significant amounts of money involved, turning it into a global challenge for law enforcement. For some time, public security agencies in China have adhered to the "four specializations and two collaborative forces" concept, organizing strict crackdowns and preventive measures nationwide against telecom and online fraud. They have continuously initiated a series of special operations to deliver a heavy-handed response to the fraud ecosystem, effectively curbing the rapid increase in cases and safeguarding social stability and public safety. With the rapid development of artificial intelligence, new forms and characteristics of telecom fraud have emerged, as many criminals begin to exploit AI technology to enhance their methods and increase their success rates. AI-driven fraud poses significant challenges to law enforcement in terms of investigation, deployment, and apprehension, due to its intelligent techniques, covert methods, and rapid technological iterations. Therefore, there is an urgent need for in-depth research into the governance of AI fraud-related issues, to predict its developmental trajectory and establish a comprehensive prevention and governance system.

2. Legal Risks Associated with AI and AI Fraud

Artificial Intelligence (AI), an important branch of computer science, was formally introduced by John McCarthy at the Dartmouth Conference in 1956 and is now considered one of the three cutting-edge technologies globally [1]. AI has become a focal point for a new wave of global revolution and industrial transformation, shifting society from "Internet+" to "AI+". Since its inception in 1956, AI has rapidly created numerous applications. In May 2017, during the Go Summit in China, Google's AI program AlphaGo defeated top player Ke Jie with a score of 3-0, demonstrating AI's exceptional

strategic prowess and vast potential applications, thus entering the public's awareness. In recent years, with the rapid development of generative AI such as ChatGPT, AI has become a tool accessible to the general public. General-purpose AI is thriving and is poised to become a crucial driving force behind the next technological revolution and industrial transformation, significantly impacting people's lives and thought processes.

2.1 Potential Legal Risks of Artificial Intelligence

Every technological advancement carries dual aspects, and AI technology faces risks of misuse. The rapid development of technology has outpaced the capacity of legal constraints, leading to widespread abuse of AI by criminals, making "AI + fraud" a significant challenge for crime governance. In 2017, law enforcement in Shaoxing uncovered the first major case of using AI technology for infringing on personal data, dismantling 43 criminal gangs and arresting 193 suspects, successfully intercepting over 1 billion stolen personal data entries and seizing over 6 million yuan in illicit gains [2]. Ulrich Beck, the sociologist who first introduced the concept of a risk society, argued that contemporary society is characterized by technological risks, with the potential legal risks of AI as follows:

2.1.1 Risk of "deepfakes"

The rise of "deepfake" technology is largely attributed to Generative Adversarial Networks (GAN) and Convolutional Neural Networks (CNN), which iteratively generate realistic outputs that closely resemble images in the original dataset [3]. The emergence of generative AI models such as Sora has disrupted traditional video generation methods, significantly enhancing the realism, smoothness, and multi-angle presentation of video content. Videos tend to be perceived as more credible than text or images, and the barriers to generating AI-produced videos are diminishing. Distinguishing the authenticity of AI-generated videos is challenging without specialized image analysis techniques, making them susceptible to misuse in spreading rumors or committing fraud, thus posing legal risks and raising copyright issues.

2.1.2 Risk of data leakage

In the internet era, data breaches have become a major global social issue. Generative AI, exemplified by ChatGPT, collects vast amounts of data while performing various tasks. Users may inadvertently share sensitive information during usage, including corporate confidential data, medical privacy information, unpublished research results, and undisclosed source code. With the introduction of the GPT-4 voice version, voiceprint information also faces leakage risks. Numerous cases illustrate that information users input while interacting with ChatGPT can become iterative data, posing data leakage risks [4].

2.1.3 Infringement risks from AI face and voice cloning

Domestic social platforms such as WeChat and Douyin boast hundreds of millions of active users, generating vast amounts of information daily, including numerous images and audio recordings. With the continuous advancement of AI technology, "cloning" another person's appearance or voice online has become feasible, emerging as a new trend for generating traffic on short video platforms. Software such as Deepface and FaceNet is widely used, and models like Deep Mind and Deep Voice can even achieve real-time synthesis. However, some criminals exploit this technology without consent, using others' images or voices for their gain, leading to violations of portrait rights. In 2024, the first case of "AI voice infringement" was adjudicated in the Beijing Internet Court, finding that the plaintiff, a voice actor, had their voice rights infringed upon by the defendant’s unauthorized use for developing an AI text-to-speech product, resulting in a compensation of 250,000 yuan.

2.2 AI Fraud

AI fraud refers to the criminal act of creating and disseminating false information through means such as phone calls and online platforms, leveraging artificial intelligence technology to deceive victims remotely, inducing them to transfer money. While AI technology offers conveniences, it is also at risk of misuse and has begun to be employed as a tool for illegal activities. The era of "Internet+" crime is gradually transitioning to an "AI+" crime model, marking the advent of AI-driven criminal activities. Since 2023, cases of fraud utilizing AI technology have

surged explosively. The "2024 Artificial Intelligence Security Report" indicates a 30-fold increase in AI deepfake fraud, with reports stating a 3000% surge in AI-based deepfake fraud incidents in 2023. The Federal Trade Commission (FTC) in the United States has warned that some individuals are utilizing AI voice technology to fabricate emergencies, tricking concerned relatives into divulging cash or personal information. Analyzing similar cases of AI fraud reveals that AI, as a new type of tool and method, has significantly increased the success rate and social harm of traditional fraud, introducing numerous novel avenues for criminal activity, which markedly differs from traditional fraud and poses substantial challenges for law enforcement investigations.

3 Empirical Analysis of AI Fraud Crimes

This section presents an empirical analysis of 25 cases of AI fraud crimes collected from online and practical departments up to August 2024. The findings are as follows:

3.1 Analysis of Incident Frequency

Since the Shaoxing police uncovered the first major case of using AI technology for violating citizens' personal information in 2017, instances of AI-related crimes have emerged continuously, showcasing a variety of criminal types. Cases such as "China's first AI-generated image copyright infringement," "the first national case involving AI cheating software," and "the first AI copyright case" have become increasingly prevalent. Regarding the frequency of AI fraud crimes, as shown in the Table 1, there were a total of 5 cases before 2022, accounting for 20% of the total. Since 2023, however, the number of AI fraud crimes has surged dramatically, with incidents now comprising 80% of the total. This indicates a geometric increase in AI fraud crimes over the past two years, signaling the onset of a wave of fraud.

Table 1 Analysis of the Incidence of AI-Related Fraud Crimes

Time	2018	2019	2020	2021	2022	2023	2024
Number of Cases	1	0	2	1	1	10	10
Proportion	4%	0	8%	4%	4%	40%	40%

3.2 Analysis of Fraud Types

According to the statistical cases analyzed, the primary methods of AI fraud include AI voice cloning, face-swapping scams, using AI to filter victims, and employing AI to create false information. Among these, AI voice cloning and face-swapping scams account for the highest proportion, reaching 68%, as illustrated in Table 2.

Table 2 Statistical Analysis of AI-Related Fraud Crime Types

Type of crime	AI voice mimicry, face swapping	AI creation of fake pornographic videos and images	AI victim screening	Using AI to write false information
Number of Cases	17	4	2	2
Proportion	68%	16%	8%	8%

3.2.1 AI face-swapping and voice cloning fraud

AI face-swapping and voice cloning fraud occur when perpetrators acquire facial and voice information from victims, relatives, and colleagues through various social media platforms. They then utilize AI technologies to fabricate scams, adding individuals as friends, making phone calls, or conducting meetings to deceive victims and gain their trust. A notable example occurred in Hong Kong in 2024, where a company employee received an invitation for a "multi-person video conference" and transferred HKD 200 million into five local bank accounts. According to police investigations, all participants in the video conference, except for the victim, were AI-generated faces. The decreasing costs of AI face-swapping and voice cloning software, along with increasingly realistic content, have significantly lowered the barriers to engaging in such scams, which has contributed to their explosive growth. A survey by McAfee, a leading global security technology company, found that 10% of respondents reported receiving messages from AI voice clones, with 77% of victims claiming they lost money as a result. Among those defrauded, 36% lost between $500 and $3,000, while 7% lost amounts ranging from $5,000 to $15,000 (approximately 105,000 RMB). Furthermore, McAfee's survey indicated that 53% of adults stated they share their voice data online or in voice recordings weekly, with 49% doing so up to ten times per week. All voice recordings generated from these activities could be at risk of hacker attacks, theft, or unauthorized sharing.

3.2.2 AI-Fabricated pornographic video and image fraud

Cyber pornography has long been a significant area of online crime. In 2023, law enforcement dismantled 13 criminal groups involved in obscene live streaming and shut down 25 one-on-one obscene streaming apps, with the involved sums exceeding 200 million yuan. In response to this severe crackdown, criminals have begun to explore alternative methods, such as using AI to fabricate pornographic videos and images for scams. Offenders illegally acquire personal information from citizens, impersonate "beautiful women," and lure victims into registering on "dating apps" to obtain their contact information. Subsequently, they engage in video chats and use AI face-swapping technology to create fabricated pornographic videos, leading to extortion. This type of deepfake pornography is not only prevalent in China but is also spreading internationally. Reports from several South Korean media outlets reveal that between 2021 and 2023, 527 victims of deepfake obscene imagery have reported incidents, with minors accounting for 59.8%. The ease of creating false videos using deepfake technology has led to a marked increase in offenders among minors, significantly surpassing those aged 20-29 (32.1%), 30-39 (5.3%), and 40-49 (1.1%). Recently, "deepfake obscene imagery in South Korea" has attracted widespread attention, as criminals use deepfake techniques to create illegal composite videos and disseminate them on Telegram under the name "Familiar Insult Room," with victim numbers reaching 246. One perpetrator's office was found with over 20,000 illegal sexual images, including materials related to the sexual exploitation of children and adolescents, with reports indicating that users could upload familiar photos in Telegram groups, and within five seconds of payment, obtain a composite nude image.

3.2.3 Fraud utilizing AI large models

Currently, personal privacy, including telephone numbers, addresses, consumption habits, and family situations, is severely compromised. The methods traditionally used to filter victims in telecom fraud are inefficient and inaccurate. To improve success rates and target victims accurately, criminals are employing AI large model tools to automatically filter targets, generating "profiles" for scam prospects and customizing scam scripts, often using AI voice robots to mislead victims. The allure of substantial profits has led to the gradual formation of a complete scam industry chain. Police anti-fraud departments across various regions have warned about the severity of new AI-related scams, even stating that "the success rate of these scams approaches 100%." In July 2018, a criminal gang engaged in renting smart dialers and making automated voice calls. They created automated responses based on "client" requests and recorded responses according to provided phone numbers. These "clients" would analyze call outcomes to gauge the strength of loan intentions, carefully planning their scripts to target victims with strong loan desires, ultimately leading to the promotion of customized fraudulent loan apps. During the loan application process, these apps would automatically prompt links related to "membership fees," "unfreezing bank account fees," and "verifying repayment ability," gradually extracting money from victims.

3.3 Analysis of Scamming Platforms

As illustrated in Table 3, the internet has become the primary medium for telecom fraud. Analyzing the platforms used for AI fraud crimes reveals that WeChat and QQ account for more than half of the incidents, serving as the main channels for AI fraud. Additionally, customized apps, phone calls, and multimedia messages are also significant platforms for these crimes. According to published data, WeChat has 1.37 billion monthly active accounts, and QQ has 570 million, making their vast user bases critical platforms for AI fraud. As technology advances and cybersecurity measures strengthen, the security of these instant messaging tools is relatively high, making it difficult to steal accounts. Criminals often resort to alternative methods; in collected cases, they impersonate family members, friends, or relatives to add victims using different accounts. After verification, they further gain trust through AI face-swapping and voice cloning methods. For example, in an AI fraud case in Inner Mongolia, Ms. Li received a WeChat friend request from someone named "Zhuo Yu Bu Yu," whose profile picture was that of her old

classmate "Jia." She accepted the request and moved to QQ for further conversation. After some pleasantries, "Jia" claimed to be participating in a project bid and urgently needed funds, hoping to use Ms. Li's account as a bridge for fund transfers. Subsequently, this supposed classmate requested Ms. Li's bank account number and claimed to have transferred 1.96 million yuan to her account. Without confirming the transfer, "Jia" tried to persuade Ms. Li to cover the funds. Trusting her old classmate, and after video calling to verify the identity, Ms. Li transferred 400,000 yuan without confirming the deposit. Afterward, the scammer continued to pressure her for more transfers. It was only then that Ms. Li realized she had been defrauded and promptly reported the case to the police.

Table 3 Statistical Analysis of Main Carriers for Implementing AI-Related Fraud

Carrier for implementing fraud	We Chat, QQ	Various customized apps	Phone calls, text messages	Other social platforms	Web pages
Number of Cases	13	5	3	3	1
Proportion	52%	20%	12%	12%	4%

4. Characteristics of AI Fraud Crimes and Governance Challenges

4.1 Characteristics of AI Fraud Cases

4.1.1 Increased intelligence and deceptiveness

With advancements in science and technology, particularly in artificial intelligence, high-tech tools have become focal points for fraudsters. Big data, deep learning, and blockchain technologies are increasingly employed to facilitate fraudulent activities, leading to continuous iterations and updates in criminal methods. Empirical analysis reveals several key aspects: First, the precision of victim selection has significantly improved. Traditional fraud typically involved the illegal acquisition of personal information and a tedious process of individually calling victims, characterized by clichéd scripts that were easily recognizable. This approach often targeted a large proportion of elderly individuals, resulting in low efficiency and success rates. In contrast, AI fraudsters utilize AI algorithms to analyze and mine victim data, employing web scraping tools to gather publicly available information and illegally purchase personal data. This enables them to generate customized scam scripts in a short time, enhancing the effectiveness of their schemes. Second, the deceptive nature of the scams has intensified. Fraudsters utilize advanced technologies such as voice synthesis and face-swapping to create a strong illusion of trustworthiness, appealing to a broad range of victims, including well-educated professionals and business owners. Third, the speed of fund transfers post-scam is rapid and difficult to trace. To evade oversight and tracking, criminals increasingly resort to virtual currencies in financial crimes, complicating subsequent evidence collection and recovery efforts. Overall, the emergence and proliferation of AI technologies lower the barriers for fraudsters, making fraudulent activities more "intelligent," "efficient," and "precise," thereby increasing the likelihood of successful scams.

4.1.2 Diverse and rapidly evolving scam techniques

Analysis of collected cases indicates that AI fraud has led to the emergence of numerous new scam techniques, beyond the conventional face-swapping and voice synthesis methods. For instance, the recent launch of the domestically produced AAA game "Black Myth: Wukong" garnered significant attention and became a key avenue for fraud. Scammers exploited players' desires to quickly obtain superior game equipment and enhance their gaming skills by creating "AI simulation gaming equipment draw" simulators in live-streaming rooms. They attracted players through tactics such as mutual role-playing and hype generation, eventually luring victims to add customer service contacts on other communication platforms under the pretext of sensitive information, enticing them to transfer money as part of the scam. Currently, the rapid pace of AI technological innovation and the increasing number of application scenarios, particularly in open-source code, make it challenging to prevent fraudsters from exploiting these advancements for illicit gains.

4.2 Governance Challenges of AI Fraud Crimes

4.2.1 Legal adaptation and insufficient constraints

Current laws and regulations provide a broad framework for the healthy development of AI

technology. The "Deep Synthesis Management Regulations for Internet Information Services," released in December 2022, explicitly prohibit actions such as creating, replacing, or altering faces and synthesizing voices. Enforced from January 10, 2023, these regulations clarify that no organization or individual may use deep synthesis services to produce, replicate, publish, or disseminate information prohibited by law or administrative regulations. Overall, the regulatory scope of relevant laws and documents is somewhat limited, resulting in inadequate enforcement. Criminals often exploit the internet as a cover, collaborating under the guise of providing services to evade legal responsibilities when fraud cases arise. The lag in legal development means that related regulations and judicial interpretations have not kept pace [5]. Many major platforms, such as Douyin and Toutiao, are widely reported to unlawfully utilize AI face-swapping technologies to manipulate others' works, leading to pervasive infringement issues.

4.2.2 Outdated investigative techniques and difficulty in evidence collection

AI fraud cases often involve audio and video content that is either entirely AI-generated or partially altered, such as fabricated pornographic images, making it challenging to discern authenticity. The high concealment capabilities of generative AI technology complicate the accurate identification of alterations in existing evidence, increasing the difficulty of evidence validation. Even when relevant evidence is retrieved, issues arise regarding its recognition. Identifying false videos and audio created by fraudsters often requires considerable manpower and resources, and verifying the authenticity of the content can be time-consuming. Additionally, factors such as user reports, rapid technological iterations, and efforts to reduce criminal risk motivate fraudsters to frequently update their methods and software, making the retrieval of electronic data challenging and potentially compromising the integrity of the evidence.

4.2.3 Lack of collaborative governance in AI fraud crime

AI fraud crimes involve multiple departments and industries, including internet, public security, and finance. The initial step in committing these crimes involves extensive gathering of publicly available and private information. According to the "2016 Report on the Protection of Online Rights in China" by the China Internet Association, in 2016, 688 million internet users suffered economic losses exceeding 91.5 billion yuan due to fraudulent messages and information leaks. On average, each individual's personal information was leaked at least five times. Despite the implementation of foundational policies aimed at protecting user data, there remain individuals and businesses that, in pursuit of short-term gains, violate laws and regulations by selling customer information for illicit profit. Protecting personal privacy is not solely the responsibility of one department; it requires a collaborative effort among multiple entities to effectively curtail the sources of AI fraud. Furthermore, the rapid proliferation of various AI software, some of which excessively obtain user permissions or even cross legal boundaries, has made these tools the primary instruments for fraudsters. Screening and regulating AI software is also a crucial aspect of combating AI fraud.

5. Governance Strategies for AI Fraud Crimes

5.1 Improve the Legal Framework in the AI Field

For effective governance of AI fraud crimes, a solid legal foundation is essential. Recent laws such as the "Personal Information Protection Law" and the "Cybersecurity Law" have been introduced. The "Interim Measures for the Management of Generative AI Services," published in July 2023, specifies that AI providers are responsible for the legality of training data and optimization data sources. The "Artificial Intelligence Law" has been included in the "2023 Legislative Work Plan of the State Council." In early August, the EU's "Artificial Intelligence Act," the world's first comprehensive regulation of AI, came into effect and will be gradually implemented over the next three years. The governance of AI fraud crimes requires the formulation of forward-looking laws and regulations to provide a solid legal basis for addressing these issues. Additionally, there should be a comprehensive legal framework addressing the definition of criminal acts, investigation, and evidence collection in AI fraud cases to ensure swift and effective implementation.

5.2 Enhance Investigative Techniques and Train Specialized

Personnel Regions across the country have developed effective governance experiences tailored to their local circumstances. For instance, the public security authorities in Shandong Province have established a closed-loop prevention and control system combining "AI + big data + human resources" to tackle AI fraud. This includes utilizing technology to sever "data feeding." Precise early warnings and preventive measures should be implemented through modern technology, ensuring that even if criminals possess citizens' personal information, relevant departments can employ technical measures to prevent citizens from accessing fraudulent websites, downloading scam software, or transferring funds to fraudulent accounts. Additionally, an AI fraud risk early warning prevention model should be developed to expand the monitoring coverage of anti-fraud calls and encourage mobile manufacturers to enhance fingerprint and facial recognition technologies to assist users in accurately identifying AI voice and face scams [6]. Furthermore, developing AI detection software to automatically recognize AI-generated deepfakes, and establishing AI firewalls to dynamically monitor risky websites and links in cyberspace, can help shield potential victims. To effectively counter the rapid evolution of AI crimes and technological advancements, it is crucial to leverage AI's intelligent characteristics for data collection and analysis related to fraud cases, using this data to train AI anti-fraud models and improve the recognition and interception capabilities of AI algorithms. These data can also be utilized to analyze fraud rings and predict trends in AI fraud crimes, enabling the quick identification of similar cases and enhancing investigative efficiency.

The core of investigating AI fraud cases lies in the personnel involved. The complexity and diversity of various cases necessitate higher professional standards for investigators. Rapid resolution of AI fraud cases requires talent, particularly those with knowledge of relevant laws, case-handling experience, and expertise in AI technologies. Investigators must continuously enhance their professional skills, acquire knowledge of AI, and master the latest investigative techniques to create tools for identifying and collecting evidence in AI fraud cases, thereby achieving intelligent data collection.

5.3 Collaborative Efforts to Improve Governance Efficiency

AI fraud cases span multiple regions and countries and involve various industries and departments. Effective governance of AI fraud requires the establishment of a high-efficiency collaboration mechanism. Law enforcement agencies should intensify their crackdown on AI fraud through a series of special operations. Internet platforms must strengthen regulation by enhancing the development of detection tools and marking AI-generated content, promoting verification of large AI products regarding privacy protection, output responses, and user demographics, thereby increasing the effectiveness of detecting false information [7]. Additionally, personal information protection mechanisms should be established to bolster the safeguarding of personal privacy data, including images and voices. Financial institutions should enhance their capacity to identify suspicious transactions, employing big data and AI technologies to monitor and analyze transaction behaviors in real-time, thereby improving the accuracy of risk assessments and taking strict action against illicit activities. Public awareness campaigns should establish effective warning mechanisms to timely discover and report new dynamics and methods of AI fraud, creating a preemptive defense system by disseminating warning information and advising the public on preventive measures to mitigate the risk of victimization.

Currently, fraud-related crimes remain a significant issue in China. In light of the rapid development of artificial intelligence and its continually evolving fraud techniques, there is an urgent need to strengthen the construction of relevant laws and regulations, establish a comprehensive prevention and investigation system for AI fraud, enhance investigative technologies, and build an efficient collaboration mechanism across law enforcement, sectors, and regions to regulate the healthy development of the artificial intelligence industry, safeguarding public property and security.

Acknowledgments

This work was supported by the “Liaoning Collaboration Innovation Center For CSLE Funded Projects (Number: XTCX2024-007)”.

Reference

[1] Zou, L., & Zhang, X. (2012). Artificial Intelligence and Its Development Applications. Journal of Information Network Security, (02), 11-13.

[2] Beck, U., & Wang, W. (Trans.). (2003). From Industrial Society to Risk Society (Part I): Reflections on Issues of Human Existence, Social Structure, and Ecological Enlightenment. Marxism and Reality, (3), 26-45.

[3] Wang, H., & Cai, S. (2021). The Legal Risks of "Deepfake" Technology and Its Collaborative Governance. Research on Science and Technology Management, 41(11), 156-163.

[4] Zhang, X. (2023). Data Risks and Governance Paths of Generative Artificial Intelligence. Legal Science (Journal of Northwest University of Political Science and Law), 41(05), 42-54.

[5] Lü, P., & Zhang, W. (2024). Research on the Governance of AI-related Telecom Network Fraud Crimes. Journal of Fujian Police College, 38(01), 23-31.

[6] Li, D., & Song, J. (2024). An Exploration of the Situational Characteristics and Preventive Mechanisms of AI Fraud from the Perspective of Artificial Intelligence. Journal of Wuhan Public Security Cadre Institute, 38(01), 53-57.

[7] Yang, S. (2024). The Impact of AI on the Governance of Telecom Network Fraud Crimes as Seen from ChatGPT. Research on Crime and Rehabilitation, (05), 26-3