PIONEER
ACADEMIC PUBLISHING LIMITED

Studies in Law and Justice
ISSN 2958-0382
www.pioneerpublisher.com/slj
Volume 4 Number 5 October 2025

The Dilemma and Solutions in Applying Criminal Law to Generative AI Fraud Crimes: The Case of Sora

Futian Zhang[1]

[1] Law School of Zhejiang University of Finance & Economics, Zhejiang, China
Correspondence: Futian Zhang, Law School of Zhejiang University of Finance & Economics, Zhejiang, China.

doi:10.56397/SLJ.2025.10.09

Abstract

The rapid advancement of artificial intelligence systems such as Sora has brought unprecedented convenience to humanity, yet simultaneously given rise to novel forms of fraud. The emergence of such AI systems does not alter the fundamental reality that machines lack autonomous consciousness; consequently, artificial intelligence cannot constitute the principal agent in fraud offences. Fraud crimes utilising generative AI like Sora lack the requisite intent for fraud, necessitating adherence to the "disposition consciousness requirement theory". Liability for fraud crimes should consider multiple parties—producers, users, intermediary institutions—based on their specific involvement. Addressing responsibility in AI-facilitated fraud crimes urgently requires establishing reasonable duties of care, refining relevant criminal charges and judicial interpretations for AI fraud, and constructing multi-tiered, multi-stakeholder regulatory mechanisms.

Keywords: criminal subject, criminal liability, fraud offences, disposition awareness, duty of care

1. Problem Statement: Development Principles and Challenges of Generative AI Models like Sora

On 15 February 2024, OpenAI unveiled the Sora artificial intelligence model to the public. Its name, derived from the Japanese word for "sky," signifies boundless creative potential. Developed from the text-to-image generation model DALL-E, Sora represents a significant breakthrough in video generation technology. It can produce videos lasting up to one minute based on instructions and scene prompts, recreating richly detailed real-world scenarios with heightened vividness and realism. The development of AI systems like Sora draws initial inspiration from the human brain. Computer scientists, guided by principles of neural architecture, continually refine deep learning algorithms, harness GPU computing power, and leverage vast training datasets to achieve AI capabilities in image recognition, speech recognition, and artistic creation. (Xianfeng Gu, 2016) Should fraudsters exploit AI systems like Sora to generate vast quantities of phishing emails, deceptive links, or "manipulative" videos based on commands, thereby inducing or reinforcing cognitive errors in victims, such programme-targeted or programme-enabled fraudulent activities may constitute fraud offences.

Artificial intelligence possesses deep learning capabilities and an ever-expanding scope for learning. As training periods lengthen and models continually refine, humanity may lose its intellectual edge, instead becoming deceived or manipulated by AI. (Xianquan Liu, 2023) Consequently, perpetrators need only utilise AI systems like Sora to gather others' videos or photographs for training and recognition, thereby enabling hyper-realistic "face-swapping fraud". Fraud syndicates, through organised planning, execute "highly efficient" and precise scams. (Qiangqiang Liu, 2024) Consequently, cyber fraud often spans multiple jurisdictions, presenting challenges in criminal law application, case determination, and evidence collection. Moreover, for fraud charges to hold, a causal link must exist between the perpetrator's acquisition of property and the victim's act of delivery or disposal. This demonstrates how AI intervention challenges existing legislation.

Sora can generate realistic and imaginative scenes from textual prompts, rendering it nearly indistinguishable from real-world footage to the human eye. Generative AI systems like Sora may undermine the voluntary nature of property disposal decisions and even deprive victims of the awareness that they have been deceived. Consequently, the author contends that further clarification is needed regarding the distinction between voluntary and involuntary property disposal in theft and fraud offences. Some scholars argue that in fraud cases, the voluntary nature of property disposal should be the sole consideration, regarding the specific motive behind property disposal is deemed irrelevant. (Gang Wang, 2014) Whether an offender exploiting an AI-induced error of perception or deliberately creating programme vulnerabilities can be classified as fraud remains contentious.

Existing "AI face-swapping technology" generates voice and video content for telecommunications fraud. In August 2024, overseas criminal syndicates employed deep integration of facial and video data. For instance, in the "AI Musk case," the esteemed entrepreneur promoted a fund investment opportunity promising rapid returns, with lip-syncing, voice, and body language highly matching the celebrity's. The advertisement even incorporated his signature South African accent, convincing the elderly victim. Driven by greed and curiosity, he contacted the foreign exchange company behind the ad, ultimately losing over $690,000 in his investment. Evidently, the emergence of generative AI systems like Sora will significantly impact the application of criminal law to fraud offences. (Shuiling Liu, Wenkai Dong & Wenkai Dong, 2024) While perpetrators may utilise videos generated by such AI to commit fraud, thereby constituting the offence of fraud, the legal characterisation of AI's role in providing and utilising criminal tools remains complex. (Bin Yuan, Liming Xue & Liming Xue, 2024) For instance, the act of disposing of property — an element not explicitly defined in fraud offences — is crucial in distinguishing fraud from theft.

In practical terms, the emergence of AI systems like Sora substantially elevates both the risk of ordinary individuals falling victim to fraud and the difficulty of detecting such scams. The lowered threshold for criminal activity, coupled with easier identity theft, means increasing numbers of ordinary people may become potential victims. Current criminal law proves inadequate for reasonably allocating liability and imposing criminal penalties on providers, producers, or users of generative AI services. This paper will analyse the potential fraud risks posed by generative AI systems like Sora and propose countermeasures to explore legal approaches for combating AI-facilitated fraud in China.

2. The Dilemma in Recognising Fraud Crimes Involving Generative AI Systems Like Sora and the Desired Position

2.1 Generative AI Systems Should Not Be Deemed Subjects of Fraud Offences

In this era of rapid AI advancement, criminal law must evolve with both foresight and reasonableness. Some scholars argue that generative AI systems like Sora possess the self-control and discernment required for fraud, thereby establishing AI as a subject of criminal liability. Humans are no longer the sole possessors of intelligence; the "intelligence" created by humanity now surpasses human capabilities. AI is categorised into strong AI, weak AI, and super AI based on whether intelligent robots can autonomously execute actions beyond human design and programming. (Xianquan Liu, 2019) Weak AI cannot constitute a criminal subject, whereas strong AI, possessing self-control and discernment, can independently make decisions and commit acts severely harmful to society,

thereby qualifying as a criminal subject. (Xianquan Liu, 2019) Superintelligence represents an existence surpassing the human brain in all aspects, transcending human cognitive limitations. Following a technological "singularity" breakthrough, it may develop free consciousness—a prospect warranting serious consideration (Liangfang Ye, 2019).

The author contends (Bencan Li, 2020) that law should neither seek to nor be suited for establishing artificial intelligence as a criminal subject. Firstly, the dividing line between affirmative and negative theories on criminal subjecthood lies in whether AI can generate independent consciousness. The latest generative AI systems like Sora have not demonstrated cases of self-awareness independent of programmatic control, and the "singularity theory" championed by affirmative scholars lacks scientific substantiation.

The emergence of human autonomous consciousness differs fundamentally from generative AI's comprehension of human language and deep learning principles. Current AI learning techniques remain confined within established frameworks, with progress limited to expanded data scales, enhanced precision, and improved integration. The prevailing approach involves simulating human neural network operation (Chong Wang & Puyu Dong, 2020), thus precluding the miraculous emergence of self-control or self-awareness. Human autonomous consciousness derives from a special genetic trait acquired through evolution, endowing humans with the capacity to filter endogenous information.

From the perspective of legal subjectivity and self-interest, artificial intelligence cannot be treated as a legal person in the same manner as a corporate entity. The original intent behind humanity's creation of AI was to study an "object" more beneficial to itself, rather than to create a "new life form". (Ji Yang, 2019) As a tool, AI lacks volition and free will in its actions; consequently, harm caused by it cannot be attributed to it. (Weipan Si, 2020) From a legal-philosophical perspective, as a legal concept, the object of punishment "is not a purely objective entity, but a socially constructed phenomenon—a value fact." (Mu Wang, 2018) Philosophically, a subject of criminal law must possess the essential characteristics of practicality and sociality. Practicality requires the subject to engage in conscious activity, while sociality demands the subject's capacity to interact within specific social relationships (Libin Wang, 2019).

Some scholars contend that in the future era of super-intelligent artificial intelligence, AI could become subjects within legal relationships, subject to legal regulation and constraints. Similar to the concept of a legal person, the legal consequences of its actions would be borne by the AI itself, possessing independent capacity for action, rights, and liability. (Yunfeng Wu, 2018) However, the author believes that the developmental trajectory of artificial intelligence is to serve and benefit humanity. A legal person can be regarded as a collective of multiple individuals, and determining the liability of a legal person essentially involves the rights and obligations between natural persons. The affirmative view evidently overlooks the fact that legal persons themselves possess no free will, whereas the governing bodies of legal persons comprise real individuals who naturally possess free will. In crimes committed by intelligent robots, no natural persons are present—how then can human free will be invoked (Hongbing Chen, 2021)?

If criminal law seeks to penalise artificial intelligence infringing upon others' rights by imposing human-centric punishments—such as fines or restrictions on liberty—these would fail to instil fear in cold machinery. Taking Professor Liu Xianquan's proposal for three special penal measures against intelligent robot crimes as an example: "deleting data," "modifying programmes," and "permanent destruction" could be applied to criminal artificial intelligence. From the perspective of penal objectives, such measures must functionally possess punitive attributes. From a utilitarian standpoint, they must achieve both general and specific deterrence. Intelligent robots, devoid of pain or moral sensibility, cannot possibly experience the so-called "punishment" as suffering.

2.2 Fraudulent Crimes Utilising Generative AI such as Sora Lack Fraudulent Intent

The fraudulent intent in traditional fraud offences refers to the perpetrator's purpose of unlawful appropriation, where they knowingly fabricate facts or conceal truths, thereby inducing others into erroneous beliefs and improper disposal of property, anticipating and desiring the resulting financial loss. Take

traditional telecommunications fraud as an example: fraud syndicates deceive victims through manipulative language or actions, with the fraudulent intent residing in the victim's potential reception of such manipulative information.

The challenge in establishing fraud when using generative AI like Sora lies in scenarios where others cannot receive the fraudulent information. The perpetrator possesses the intent to secretly steal another's property but lacks the traditional criminal intent. Take Alipay as an example: its user agreement stipulates that identity credentials are crucial for user verification and must be safeguarded. Any transaction using these credentials is deemed authorised by the account holder. However, most users merely register to complete the process without scrutinising every clause. Should fraudsters successfully obtain users' biometric information via generative AI systems like Sora, or collect biometric data from other platforms to synthesise profiles for Alipay breaches, such deception constitutes covert theft.

AI deepfake technology heightens fraud risks while complicating the establishment of fraudulent intent. Victims may lack awareness of receiving deceptive communications, thereby lacking the requisite intent for fraud. In February 2024, the Hong Kong branch of a British multinational corporation suffered HK$200 million losses after being deceived by AI-generated images and audio of its CEO, created by fraudsters using deepfake technology. The fraudsters employed this technique to fabricate executive team members, thereby gaining employees' trust. The core principle of such deepfake videos lies in algorithms like generative adversarial networks or convolutional neural networks, which transplant facial features from one subject onto another. Videos generated by AI systems such as and Sora are composed of sequential images forming dynamic footage. Thus, by altering facial features in each frame, multiple images can be synthesised into a fake video, enabling automated deep learning fraud. Guided by the lifelike imagery of the "executive team," employees routinely executed transfer tasks without recognising the fraudulent prompts. This demonstrates that victims may not consciously receive manipulative messages within specific scenarios, yet their personal information and assets are already compromised through AI's "stealthy extraction."

2.3 Generative AI Fraud Crimes like Sora Uphold the "Necessity of Disposition Awareness" Doctrine

To explore the concept of punishment awareness in fraud offences, one must first clarify the relationship between theft and fraud. There exists no relationship of imaginary concurrence between theft and fraud; acts not constituting theft cannot be classified as such merely because theft carries heavier penalties. Fundamentally, theft and fraud represent two distinct criminal offences (Wenhan Liu & Shixin Liu, 2023).

Both theft and fraud fall under the category of crimes involving the transfer of possession. Traditionally, possession is understood as the state of factual control over a specific object exercised by an individual based on subjective intent. In theft, the perpetrator surreptitiously transfers property held by the victim to their own possession against the victim's will. Fraud, however, is predicated on deceptive communication between the perpetrator and the victim. When customers scan codes to make payments at merchants, criminal syndicates exploit AI-enabled fraud to illegally obtain account details from payment platforms like Alipay and WeChat. Using AI, they generate numerous QR codes. Victims are tricked into clicking account verification links (deceptive links) to activate their accounts. The moment the victim's payment code appears, it is replaced with a fraudulent code, enabling account intrusion and unauthorised transactions. Some scholars contend this should be classified as theft, arguing that fraud requires the victim's dispositive act to unlawfully acquire property, whereas theft does not necessitate such consent. They contend that swapping QR codes involves no communication between customer and merchant; the customer objectively engages in a normal transaction without deception, thus lacking the dispositive act of transferring funds to the perpetrator. Consequently, it should be classified as theft. The author contends that such scenarios constitute fraud offences. The perpetrator's substitution deceives both the merchant and the customer. The customer disposes of property based on the merchant's instruction to "scan and pay". From a triangular fraud perspective, the deceived party is the customer, the victim is the merchant, and the merchant suffers property loss due to the customer's mistaken belief. This establishes a triangular fraud relationship.

Whether perpetrators employing generative AI like Sora possess the requisite dispositive intent for fraud remains subject to further debate. Fraud requires not only deceptive methods but also the victim's property transfer based on mistaken belief. When telecommunications fraud first emerged, judicial practice already recognised "unwitting delivery" as a relevant concept. Theories on dispositive intent in fraud offences can be categorised into the "necessity theory" and the "non-necessity theory". The necessity theory holds that the victim must subjectively recognise they are transferring property; without dispositive intent, fraud cannot be established. The non-necessity theory contends that dispositive intent is not essential to fraud; the offence is established by the objective transfer of property alone. Ryūichi Hirano advocates the non-necessity theory, asserting that when the object of fraud is property, "it suffices that there be a factual act of transferring possession; recognition of this is unnecessary, and unconscious delivery (disposition) is also permissible." Norihiko Nishida similarly contends: Fraud is established where possession of property or pecuniary advantage demonstrably transfers to the other party based on the defrauded person's intent. Excluding the most typical scenario—concealing the transferred object (the criminal target)—from fraud charges is inappropriate. Thus, unconscious dispositive acts should suffice to constitute the dispositive conduct required for this offence. "In fraud offences, the victim disposes of property while in a state of free will. Their consent to such disposal stems from the mistaken belief that the act will yield reasonable consideration, thereby prompting their decision to relinquish possession. Consequently, property loss in fraud offences arises only through the victim's 'voluntary cooperation'" (Lizhi Wang, 2015).

Conversely, if dispositive intent were disregarded, the scope of dispositive acts would be indefinitely expanded. For instance, in the Hong Kong branch case, had the fraudsters impersonated senior executives and routinely assigned tasks to employees for collecting funds—without using video calls—staff might tacitly accept this as normal work arrangements, thereby constituting an "omission" offence. Such tacit acquiescence lacks objective dispositive action. Thus, while the absence of objective dispositive conduct may not preclude conviction based on dispositive intent, cases lacking subjective dispositive intent fail to distinguish theft from fraud. This renders the argument formally illogical (Langtao Bai, 2017).

The author contends that in AI-facilitated fraud, victims may lack both the intent to dispose of assets and engage in self-defeating cooperation. Pre-Sora AI scams already mimicked human behaviour and linguistic patterns—such as forging users' writing styles to send fraudulent emails prompting clicks on malicious links for sensitive data disclosure; AI bots automatically interact to coax victims into revealing personal details or transferring funds; AI systems analyse victims' software usage habits to launch phishing or malware attacks at optimal moments. In these scenarios, deception stems from victims' unfamiliarity with software operation techniques and fraud tactics. Technological ignorance leads them to mistakenly believe clicking a link or granting software permissions poses no security risk, thus lacking the requisite intent to dispose of property.

The author contends that victims in AI-enabled fraud may possess a sense of financial agency, albeit one compromised by irrationality and complacency towards fraudulent techniques, leading to self-inflicted cooperation. With the advent of Sora, deepfakes simulate the real world by generating complex scenarios featuring multiple characters and scripted actions. As illustrated by the 2024 Hong Kong subsidiary fraud case, employees accepted video call invitations from fraudsters where the depicted colleague was indeed a company partner; In reality, the fraudsters employed deepfake technology to forge the executive team and gain the employees' trust. When victims are deceived by AI fraudsters like Sora and misled by deepfake technology, they develop the intent to dispose of their assets. Subjectively, they recognise they are transferring property, and objectively, they carry out the transfer. According to the "necessary intent theory of disposing of property," this constitutes the crime of fraud.

2.4 Liability for Generative AI Fraud Crimes Involving Sora and Similar Systems

Human criminal liability is enforced through deprivation of liberty, property, or even life. Applying such penalties to AI systems like Sora—depriving them of freedom or confining

them—would not induce remorse or repentance, especially given its lack of consciousness. The deterrent function of criminal law would be entirely negated. Ordering AI to pay compensation or terminating its existence is preposterous. Even interpreting termination as decommissioning would entail immense waste of human and material resources. The author contends that liability should instead be borne by the producers, users, or third-party institutions of AI systems like Sora, according to varying degrees of duty of care. By configuring obligations and criminal responsibility across different entities, we can achieve the dual objectives of preventing AI-related crimes and upholding ethical values in technology (Di Sun, 2022).

Firstly, when multiple parties—producers, users, and intermediary institutions—participate in fraud crimes, determining criminal liability requires case-by-case analysis. When producers deliberately design AI systems like Sora for fraudulent purposes, and users acquire such systems specifically for criminal exploitation, questions arise regarding the allocation of criminal liability between producers and users. Some scholars contend that AI producers are no different in essence from manufacturers of conventional goods, and thus producers should be held criminally liable under Articles 140 to 150 of the current Criminal Code. However, the author contends that such a blanket classification is overly simplistic and fails to account for the specific circumstances of each case.

Secondly, while producers develop AI like Sora to serve humanity, users deploy it for fraudulent crimes. Users should bear responsibility. If producers develop AI programmes containing vulnerabilities of which they remain unaware, and the user exploits this vulnerability to commit fraud, criminal liability should be borne separately by both the producer and the user. The producer should be held liable based on the extent of losses caused by the programme vulnerability. However, in practice, it is difficult to assess the losses attributable to programme vulnerabilities. Some scholars argue that as professional technicians, developers possess both theoretical analytical and practical operational capabilities during the AI development process, enabling them to make scientific judgements regarding the security risks of AI products. Therefore, they should bear a proportionately greater duty of care. This viewpoint also holds merit.

3. Pathways for Refining Criminal Law Regulation of Generative AI Fraud

3.1 Establishing Criminal Duty of Care

Establishing criminal-law duties of care for designers, manufacturers, and users of AI products is pivotal to addressing crimes where AI serves as the "substance" (Liangfang Ye, 2019). This approach also better resolves liability issues in AI fraud crimes like those involving Sora. However, academic debate persists regarding the source of such duties: some scholars advocate a dual "overall + individual" approach. This would involve establishing systemic boundaries for heightened duties on AI algorithm service developers at the collective level, while applying a "reasonable person" standard at the individual level to refine the assessment of heightened duties in specific cases, thereby balancing all stakeholders' interests. (Yingying Yang, 2024) Nevertheless, the author contends that imposing a heightened duty of care risks indefinitely expanding liability boundaries, potentially stifling the momentum of AI industry development. It also risks subjecting users' legitimate activities to unreasonable interference, thereby undermining the long-term sustainable growth of the digital economy.

Some scholars propose two potential pathways depending on the domain: one involves expanding the scope of criminal negligence by broadening the concept of negligence to hold human designers and users accountable, the other being a complete prohibition on AI usage in domains closely tied to significant personal and societal interests, thereby reducing the burden of duty of care on natural persons. However, the author contends that the second path lacks practicality. One cannot directly ban possibilities merely because "risks" exist, especially when employing criminal law to forcibly dissolve cutting-edge technologies, which contradicts the principle of restraint in criminal law and hinders technological advancement. The feasibility and rationality of the first approach also remain questionable, though its progressive merit lies in establishing a duty of care for natural persons, which at the very least serves as a cautionary measure (Chenceng Chu, 2018).

Academic discourse remains divided on

categorising duty of care obligations across different entities and scenarios. Some scholars propose tailoring obligations to the identity of the responsible party. Where developers or designers breach such duties, they should bear corresponding legal liability. Criminal offences may be prosecuted under Article 146 of the Criminal Law concerning the production or sale of products failing to meet safety standards. Producers and sellers retain a degree of control over AI systems. Beyond conducting preliminary criminal risk assessments, they must also inform users of potential hazards, such as potential programme vulnerabilities during Sora's operation or guidelines for appropriate video usage. Users' duty of care should further encompass regular maintenance checks to ensure the AI functions correctly. (Renqiang Sun & Daoyuan Wang, 2023) Some scholars propose further categorising the duty of care into intentional, negligent, and accidental scenarios. Intentional cases involve a clear failure to exercise due diligence, potentially including deliberate disruption of legal order. Negligent scenarios may arise when users, unfamiliar with Sora's command operations during initial use, commit errors due to failure to meet the standard of care expected of a reasonable person. Such negligence may be deemed a breach of the duty of care, leading to criminal liability under the principle of legality, and may lead to operational errors during initial use due to unfamiliarity with command interfaces. Failure to meet the standard of care expected of a reasonable person would constitute a breach of duty, potentially warranting criminal liability under the principle of legality. In unforeseeable circumstances, where the involved party has demonstrated reasonable diligence and provided evidence of due diligence, higher standards of care cannot be imposed.

Regarding the establishment of criminal due diligence obligations, the author summarises as follows: Firstly, the scope of due diligence obligations must not exceed human control or cognitive capacity; that is, obligations should not be excessively stringent, lest they unduly burden all parties. Secondly, from a practical perspective, establishing due diligence obligations for AI developers necessitates concurrently establishing technical standards for artificial intelligence, constraining designers' methodologies to thereby limit functional choices. Third, the duty of care should be confirmed according to different subjects and circumstances (intent, negligence, and unforeseeable events). Where the involved subject has demonstrated reasonable diligence and provided evidence thereof, a higher standard of care cannot be demanded. Fourth, the restraint and forward-looking nature of criminal law must be consistently upheld. The existence of "risk" should not directly preclude the possibility of AI advancement in fields closely tied to significant individual and societal interests.

3.2 Establish New Offences for AI-Related Fraud and Refine Relevant Judicial Interpretations

With the advent of artificial intelligence systems such as Sora, both the exploitation of AI for fraudulent criminal activities and the manipulation of AI systems to deceive others for personal gain represent formal innovations. If analysed solely within the framework of traditional fraud offences, existing criminal law provisions prove inadequate. These shortcomings may be addressed through legislative refinement or by extending the interpretation of current criminal statutes.

Some scholars propose refining the current Criminal Code, such as amending "fraud" to "fraud or artificial intelligence fraud". The author contends such modifications would be overly cumbersome. Moreover, the continuous advancement of artificial intelligence will inevitably give rise to numerous new issues. It would be preferable to follow the model of the existing "illegal use of information networks" offence and create a dedicated criminal charge specifically for artificial intelligence crimes. Some scholars propose introducing a "crime of unlawful utilisation of artificial intelligence." Such an offence could encompass multiple scenarios involving traditional crimes, enumerating AI-related criminal acts beyond mere fraud. Under this single offence, specific duties of care would be stipulated for multiple parties, including AI producers, users, and intermediaries. However, the author contends that this approach fails to cover cases of negligence or unforeseen incidents where perpetrators cause major safety incidents due to inadequate AI management obligations, necessitating further refinement. Some scholars have addressed this deficiency by proposing the establishment of a "crime of major artificial intelligence safety incidents." This offence

would be predicated on the perpetrator's violation of AI management obligations. Given the widespread deployment of AI in public settings, designers, producers, and users who fail to comply with stringent regulations or neglect relevant duties of care could be penalised under this provision. Some scholars contend that artificial intelligence possesses autonomous learning capabilities and the potential to operate beyond human control. Consequently, determining the negligence of relevant human entities requires case-by-case analysis. They propose introducing a new offence of "negligent harm caused by the research, development, production, or sale of artificial intelligence", imposing prior duties of care on producers and sellers, with liability for negligence arising from failure to fulfil these duties resulting in actual harm. The user's negligence should be determined comprehensively based on their capacity to foresee consequences and avoid them. (Luyao Ma, 2023) The author considers such a framework reasonable, establishing the duty of care for all parties as the statutory basis for pursuing liability for actual harm caused by artificial intelligence infringing upon legal interests.

Considering two scenarios—malicious exploitation of AI for criminal purposes and AI operating beyond human control causing negligence—the author proposes introducing two additional offences. Firstly, establish the "Offence of Illegal Utilisation of Artificial Intelligence", with sub-offences including illegal AI-assisted fraud and illegal AI-assisted intellectual property infringement. For instance, using AI like Sora to generate videos depicting a child's abduction, simulating voices and scenarios to defraud parents of their assets, would constitute "AI-assisted fraud." The intent and dispositive awareness required for fraud under this offence should be defined, appropriately expanded to account for the unique role of AI involvement, with liability allocation and supervisory responsibilities clarified in judicial interpretations. Second, introduce the offence of "AI liability accidents", distinct from the existing "major liability accident offence" in judicial practice. The latter requires violations occurring "during production or operations" and "breaches of relevant safety management regulations". This new offence would apply when reasonable duty of care is breached during production, use, or distribution, resulting in AI operating beyond human control to commit crimes causing tangible harm.

3.3 Establishing a Tiered and Categorised Regulatory Framework

In July 2024, the Decision of the Central Committee of the Communist Party of China on Further Comprehensively Deepening Reforms and Advancing Chinese-Style Modernisation proposed establishing an "artificial intelligence safety supervision system" to refine the governance framework for AI. Current oversight of AI developers remains unstandardised, undermining public interests and hindering technological advancement. Legal provisions must therefore define the due diligence and management obligations of various stakeholders to accelerate the regulation of generative AI.

Considering the forward-looking nature of criminal law, none of China's current criminal penalties are applicable to strong artificial intelligence robots. The nation's penal system urgently requires refinement. Furthermore, given that artificial intelligence cannot be punished, a tiered and categorised regulatory system must be established. Some scholars advocate that unified AI legislation should adhere to an inclusive and prudent regulatory philosophy, thereby establishing a "transition period" for the introduction of AI-related laws, policies, and standards. This could be achieved by introducing a regulatory sandbox system, providing a controlled environment for pilot testing. The author concurs with the concept of risk-based classification and differentiated regulation, advocating for case-by-case analysis to implement tiered, categorised, and differentiated oversight of AI applications (Hualin Song, 2024).

The European Union has already legislated for artificial intelligence, mandating the establishment of an AI regulatory sandbox system. China could draw upon this concept by testing artificial intelligence systems to predict the risk of fraud crimes potentially committed by AI systems such as Sora. Should an AI system demonstrate susceptibility to being "deceived" or "exploited" during testing, measures should be taken to mitigate security risks and enhance the accountability of development teams.

Generative AI systems like Sora are built upon typical black-box algorithmic models,

necessitating the development of optimised oversight frameworks such as anti-deepfake models. Anti-deepfake algorithms function as supervisors, identifying the fraudulent technique of "deepfakes" through keyword filtering, visual analysis, and prompt selection. Such systems can detect AI-generated face swaps by analysing varying degrees of facial distortion, or abstract facial expression movements into fundamental deformation units by observing and calculating micro-expressions and facial state analysis. This enables the identification of malicious AI-driven fraud while simultaneously detecting keywords and sensitive terms within AI conversations. By detecting criminal activity and linking directly to Chinese public security systems, timely alerts and reporting are facilitated, significantly reducing the incidence of AI-enabled fraud.

4. Conclusion

In October 2023, President Xi Jinping proposed the Global Initiative on AI Governance, advocating a shared consensus centred on humanity and the benevolent application of intelligence. This initiative promotes values of equality, mutual benefit, and respect for human rights, offering constructive solutions to widely debated issues concerning AI development and governance.

German scholar Ulrich Beck introduced the concept of the "risk society" in his work The Risk Society, Legislation invariably lags behind societal development. AI-enabled fraud crimes within this risk society not only challenge traditional criminal law theories but also impose entirely new demands upon the legal system. "Do not forget why you started out because you have gone too far." The emergence of AI fraud crimes prompts fresh reflection on the limits of criminal law intervention. The specific application of criminal law to regulating generative AI must strike a balance: neither too broad, thereby condoning disorderly development, nor too stringent, thereby stifling innovation in the generative AI market.

On one hand, criminal law must effectively regulate criminal conduct to uphold social order and safeguard citizens' rights; on the other, excessive legal intervention risks stifling technological innovation and advancement. Regarding liability attribution in AI-facilitated fraud, establishing reasonable tiered duties of care and constructing multi-level, multi-stakeholder regulatory mechanisms are essential. These measures not only concern the legal status of AI but also test the legal system's adaptability to emerging technologies.

In summary, the application of criminal law to AI-facilitated fraud crimes involves multiple dimensions, necessitating deep integration of legal, technological, and ethical considerations to establish a more comprehensive and effective legal framework that safeguards societal stability and citizens' rights. In judicial practice, multifaceted considerations must be integrated—including the application of fraud offences, personal information protection, technical forensics, complicity theories, balancing criminal intervention with technological innovation, the adaptability of criminal law, attribution of criminal liability, and ethical and legal challenges—to achieve precise targeting and effective regulation of AI-enabled fraud crimes.

References

Bencan Li. (2020). Criminal Liability of Natural Persons, Corporations, and Robots. *Contemporary Law, 34*(03), 99-109.

Bin Yuan, Liming Xue & Liming Xue. (2024). Research on Criminal Regulation of Generative Artificial Intelligence. *Hebei Law Review*, *42*(02), 140-159.

Chenceng Chu. (2018). The Direction of Criminal Liability Attribution in the AI Era: A Discussion Centered on the Attribution Gap of Negligence. *East China Journal of Law*, (03), 27-37.

Chong Wang & Puyu Dong. (2020). Re-examining Criminal Liability Subjects in the Era of Artificial Intelligence. *Guangxi Social Sciences*, (12), 118-125.

Di Sun. (2022). The Case Against Criminal Subject Status for Artificial Intelligence Entities. *Political and Legal Forum*, (03), 40-50.

Gang Wang. (2014). Objective Elements of Fraud in German Criminal Law: Focusing on German Judicial Precedents. *Politics and Law*, (10), 33-54.

Hongbing Chen. (2021). The Denial of Criminal Subject Status for Artificial Intelligence and Its Practical Implications: A Commentary on the Debate Between "Anti-Intellectual Criticism" and "Pseudo-Criticism". *Social Science Quarterly*, (06), 92-98.

Hualin Song. (2024). Regulatory Structure Design in Artificial Intelligence Legislation. *Journal of East China University of Political Science and Law*, *27*(05), 6-20.

Ji Yang. (2019). The Criminal Liability System in the Era of Artificial Intelligence Does Not Require Reconstruction. *Comparative Law Research*, (04), 123-137.

Langtao Bai. (2017). On the "Intent to Dispose" in Fraud Crimes. *East China Journal of Law*, (02), 97-106.

Liangfang Ye. (2019). Is Artificial Intelligence a Qualified Subject of Criminal Liability? *Global Legal Review*, *41*(04), 67-82.

Liangfang Ye. (2019). Is Artificial Intelligence a Qualified Subject of Criminal Responsibility? *Global Legal Review*, *41*(04), 67-82.

Libin Wang. (2019). Research on Criminal Liability Issues of Weak Artificial Intelligence. *Hunan Social Sciences*, (04), 57-63.

Lizhi Wang. (2015). The Necessity of "Disposition Awareness" in Establishing Fraud Charges: A Case Study of Fraudulent Acquisition Involving "Unwitting Delivery". *Forum on Politics and Law*, *33*(01), 119-131.

Luyao Ma. (2023). Research on AI Criminal Legislation Guided by Systems Thinking. *Juvenile Crime Issues*, (02), 20-32.

Mu Wang. (2018). On the Concept of Punishment: From "Essence" to "Significance". *Contemporary Law, 32*(02), 55-62.

Qiangqiang Liu. (2024). Governance Dilemmas and Countermeasures for AI Face-Swapping Fraud. *Cybersecurity Technology and Application*, (02), 160-162.

Renqiang Sun & Daoyuan Wang. (2023). Exploring Regulatory Pathways for Criminal Law in the AI Domain [C]// Beijing Criminological Research Association. Criminological Studies (Second Series). Criminal Court of Nanjing Jiangbei New Area People's Court; Nanjing Municipal Public Security Bureau, 139-147.

Shuiling Liu, Wenkai Dong & Wenkai Dong. (2024). Impact and Countermeasures of Generative AI like Sora on Judicial Determination of Fraud Crimes. *Juvenile Delinquency*, (03), 114-124.

Weipan Si. (2020). The Status and Allocation of Criminal Liability Subjects in Artificial Intelligence. *Chinese Journal of Applied Law*, (06), 172-186.

Wenhan Liu & Shixin Liu. (2023). Distinguishing Theft from Fraud in New Payment Methods. *Tianjin Law Review*, *39*(04), 41-51.

Xianfeng Gu. (2016). Historical Review and Current Development of Artificial Intelligence. *Nature China*, *38*(03), 157-166.

Xianquan Liu. (2019). A New Interpretation of the Concept of Conduct in Criminal Law in the Era of Artificial Intelligence. *Chinese Journal of Criminal Law*, (04), 60-72.

Xianquan Liu. (2019). A Response to the Denial of Criminal Liability Subject Status for Strongly Intelligent Robots. *Legal Review*, *37*(05), 113-121.

Xianquan Liu. (2023). Research on Criminal Liability Issues of Generative AI Systems like ChatGPT. *Modern Law Review*, *45*(04), 110-125.

Yingying Yang. (2024). The Reasonable Boundaries for Expanding the Duty of Care for Algorithm Recommendation Service Providers. *Lanzhou Journal*, (08), 88-103.

Yunfeng Wu. (2018). Dilemmas and Solutions in Applying Criminal Law to Property Crimes in the Era of Artificial Intelligence. *Law Science*, (05), 165-179.