WEP Warwick Evans Publishing

Transactions on Social Science, Education and Humanities Research
ISSN: 2960-1770, eISSN: 2960-2262 | Volume 17
CEPC 2026

Criminal Law Regulation of Network Rumors Generated by Autonomous Operation of Artificial Intelligence Programs

Shiwei Fu [1], Jianshu Cao [2, *]

[1] School of Information Engineering, Beijing Institute of Petrochemical Technology, Beijing, China, 102617

[2] School of Engineering, Beijing Institute of Petrochemical Technology, Beijing, China, 102617

* Corresponding Author Email: 0019980349@bipt.edu.cn

Abstract. When large language models generate text without human oversight-a capability that generative artificial intelligence has now made routine-the question of who bears criminal law responsibility for harmful output becomes acute. Online misinformation produced by autonomous programs does not fit neatly into doctrines built on the assumption of a human perpetrator. This article tackles that mismatch. It begins by unpacking the technical roots of AI-generated rumors: flaws embedded in software architecture, the volatile behavior of self-learning algorithms, and situational triggers in the deployment environment. These factors, separately and in combination, explain why fully autonomous rumor generation is not merely theoretical. The legal fallout is threefold, though not neatly compartmentalized. Pinpointing a responsible subject is elusive when the "actor" is code; subjective intent dissolves where there is no mind to read; and the causal chain from algorithm to social harm disappears into opacity. Addressing these failures, the study argues for a "program legal liability" framework predicated on "breach of program security obligations"-a standard that casts a duty net over developers, operators, and users alike. Alongside it runs a dual-track accountability structure: "actor liability" for human participants and "program liability" for the autonomous system itself. The argument culminates in "algorithm risk creation," a concept that leverages objective imputation theory to bypass the evidentiary dead-end of the "black box," drawing a defensible line between criminal and non-criminal outcomes. Together these elements reframe how criminal law should confront the challenge of self-generating misinformation.

Keywords: Artificial Intelligence; Autonomous Program Operation; Network Rumors; Criminal Law Regulation; Program Legal Liability.

1. Introduction

The quick advancement of generative artificial intelligence systems approaches have allowed large language models to create text output automatically grounded in computational logic, free from immediate human control. Represented by schemas like the GPT series, Baidu Large System, and DeepSeek, generative artificial intelligence frameworks possess strong language comprehension capabilities and independent decision-making abilities, allowing them to generate text material on their own grounded in algorithmic schemas lacking immediate human intervention. While these systems bring enormous convenience to humanity, they also create emerging legal obstacles, one, of which is the emergence of online misinformation. Compared with conventional online misinformation, AI-generated online misinformation exhibits three characteristics: fast generation speed, a wide spread range, and elaborate transmission chains. Especially, whenever such systems operate on their own and generate online misinformation, conventional criminal law concepts centered on "human perpetrators" face core problems in attributing responsibility, and the conventional criminal law imputation logic has fallen into an unprecedented deadlock [1].

Academic discussions on the criminal law governance of network rumors have already generated a certain extent of outcomes. For instance, Liu Renwen proposed a responsibility system anchored in "management responsibility," incorporating software designers and platform operators into the scope of governance obligations to achieve accountability [1]. Pi Yong carried out a comprehensive examination of AI safety and criminal responsibility, recommending feasible approaches for duty

content and responsibility creates [2]. In contrast, Zhou Guangquan, originating obtained from the viewpoint of broad criminal law idea, viewed the problems and possible countermeasures faced by criminal law in the age of AI [3]. Liu Yanhong organized theoretical resources originating extracted from the viewpoint of AI criminal law, systematically and practically explaining related crimes [4]. Zhao Bingzhi and other authors introduced novel criminal law issues into the discussion, examining the barriers presented by AI to present criminal law and responsibility measures [5]. A certain number of experts analyzed the functions of software developers, users, and operators in AI harms grounded in different responsibility undertakings; others focused on regulatory novelty itself, recommending the idea of using technical approaches to effectively control AI [6].

Existing studies still have several apparent weaknesses, which can be summarized as follows: Firstly, the study objects are chiefly limited to situations where actors actively utilize AI to spread rumors, in the absence of incorporating the situation where AI algorithms automatically operate built upon their own logic and generate rumors unconsciously, thus failing to properly elaborate the new problems brought by self-operating algorithm functioning. Subsequently, the responsibility attribution procedure is outdated; conventional criminal law takes "actor's subjective intent" as the starting element for imputation, which may not meet the practical needs of self-operating software running, and the imputation idea needs to be updated. Third, the recommended countermeasures are fragmented and one-sided, lacking an unified theoretical or institutional system, and analyses obtained originating from various standpoints lack integrated conclusions. As an outcome, this paper centres on "criminal law governance of internet rumors generated by self-operating functioning of AI algorithms," analyzing the rumors, criminal law obstacles, and solutions related to self-operating AI algorithm functioning, and putting forward feasible suggestions for legislation and adjudication in a theoretical structure.

2. Mechanism Of Ai Programs Generating Network Rumors

2.1. Limitations of Software Design and Operational methodology

From the standpoint of computational logic, generative AI employs the Transformer architecture and self-attention mechanisms to predict the probability distribution of text sequences [5]. The Transformer organization combines encoders and decoders: the former encodes the input sequence and summarizes its semantic information, while the latter yields the output sequence grounded in these outcomes. The self-attention framework allows the structure to simultaneously consider all elements in the input, allowing it to handle long-distance dependencies. Nevertheless, the probabilistic generative logic thus formed is exactly one of the technical foundations for the emergence of network rumors. whenever analyzing sophisticated semantics like irony or metaphors, existing training datasets may not supply sufficient examples, resulting in the architecture misjudging the original meaning and simply "filling in" information gaps consistent with habits, resulting in realityual fabrication.

Originating from this viewpoint, the inputs itself has an uncomplicated ramification on the effects generated by the paradigm. Crucially, Nevertheless, massive amounts of text on the internet cannot all be true or accurate information, and applications bring these defects into the architecture, resulting in erroneous knowledge sets up [3]. Generative schemas require large-scale text for training, and the information they rely on encompasses news, social posts, forum content, etc., which inevitably contain harmful materials as an illustration rumors, prejudices, and stereotypes. In addition, the standpoint used is self-supervised learning, which summarizes patterns obtained originating from datasets and applies them. The erroneous parts contained in the training inputs are "learned" and output.

Additionally, the "hallucination" appearance of generative artificial intelligence is the foundational technical source of online misinformation formation [5]. "Hallucination" refers to situations where the structure produces material with erroneous facts in a logically coherent manner, which is

connected to training datasets, computational logic, or reasoning mechanisms. Given the present technical leveled, the "hallucination" manifestation cannot be completely avoided and represents a structural danger for the emergence of online misinformation. Present analyses prove that even at the majority of advanced stages of large-scale language schema advancement, the "hallucination" rate is generally inside the range of 10%-20%, along with an even higher proportion in realityual questions. Essentially, this structural defect can generate software to yield answers that appear plausible but are actually false whenever faced along with unknown problems.

Additionally that the improvement objectives of the system itself may add to false information. For generative frameworks, the conventional strategy is to undertake "generating high-quality, coherent text" as the refinement objective, instead of "generating truthful, accurate text." whenever these two objectives cannot be simultaneously achieved, the system often prioritizes the former, resulting in a loss of truthfulness. whenever an user asks the system to write details about an event that did not actually undertake place, the system may mimic the architecture learned originating drawn from persistent examples and potentially generate narratives that are superficially reasonable but essentially fictional.

2.2. Uncertainty of Autonomous Decision-Making Ability and System Evolution Law

The changeability of autonomous capabilities An inducing component for internet rumors caused by the autonomous functioning of AI programs [5]. In broad terms, conventional software programs operate built upon established rules and entail little autonomous choice-making. present generative AI can independently handle specific matters, that is, generate corresponding content grounded in computational logic. This autonomous competence is not completely reliable or fixed; in use, each choice is constrained by multiple factors, resulting in substantial ambiguity in the software's actual behavior.

From the above viewpoint, the "superficial understanding" operational mode of AI further increases variability [5]. algorithms merely yield information built upon statistical procedures for words and cannot grasp circumstance or social reality, Hence having the potential for deviation. In addition, the "understanding" of generative frameworks is pattern matching built upon probability, not actual semantic examination. at the time whenever faced together with ambiguous or conflicting instructions obtained from users, the system sometimes makes a speculative "guess" using known patterns obtained from training inputs, lacking directly stating that it does not understand. Although such "guesses" could generate logically complete text, they may not avoid the emergence of false material.

It ought to also be pointed out that the society's undue faith in AI-generated material forms a "material echo chamber," further increasing false information caused by variables connected to independent decision-making [5]. Given the existing implementation of AI in various domains, faith in AI is continuously growing, sometimes even leading to blind acceptance of AI effects. Many users employ AI-generated materials as authoritative information in the absence of proper verification, which facilitates the spread of rumors on the internet. Crucially, furthermore, the highly realistic nature of AI-generated output itself is difficult to pinpoint, thus giving rise to the large-scale emergence or spread of false information.

The changeability of independent decision-making capabilities may be examined originating from three perspectives: conflicting decision objectives, non-reproducibility of techniques, and delayed supervision [5]. whenever making decisions, AI may demand to handle conflicts between two or more objectives, i. e., simultaneously pursuing high-quality output and meeting user necessities. provided that these objectives conflict, the system may generate false material as a way to satisfy user demands. The decisions made by programs are not strictly reproducible; the same input may not create the same effects, and this indeterminacy makes algorithm functioning impossible to fully control. provided that supervision is delayed, it is elaborate to quickly pinpoint and correct the erroneous behavior of computers, and such problems all conduce to the emergence of network rumors.

Upon thorough examination, the fluid perils involved in long-term framework advancement serve as a typical example of the ambiguity of autonomous decision-making [5]. derived originating from an overall angle, once software runs or learns, its conduct shall undergo specific gradual changes; reasoning approaches may shift in implicit ways, and design objectives may be weakened or deviate derived originating from their original intent. Although such progress is slow and intricate to detect, it may generate the deployment to generate false information in individual instances. Hence, automated learning itself cannot prevent computer programs derived originating from creating bad habits via user feedback, Hence creating a vicious cycle of "computational logic evil." These evolutionary characteristics generate online misinformation to appear in nonlinear, changing structures, bringing additional challenges for the purpose of governance.

2.3. Inducing Factors of Code Execution Environment

AI programs rely on code for actual execution, and both the code itself and the external environment, in which it operates have a direct influence on the algorithm's output [4]. Hence, code execution is not a single event but consists of multiple components, including execution, storage, and network connections, each, of which may malfunction, Hence becoming an inducing component for network rumor occurrence.

From the viewpoint of code execution, software programs requirement to be analyzed and processed by compilers or interpreters, and even minor defects in the code may create software exceptions [5]. The guiding nature of input information may create the structure to "fabricate details"; ambiguous or contradictory input is often an inducing element for software to generate false information. Insufficient resources force the system to simplify reasoning instead of performing complete calculations, and prompt attacks may outcome in directional forgery of output. on the condition that the resources on which computer execution depends are not abundant enough, reasoning may be greatly simplified, and the actual effects may not be reliable, and the obtained information may be false and non-factual.

It must also be regarded that anomalies involved in system interactions may have chain impacts [7]. When designed in accordance with a distributed standpoint, communication and collaboration amidst subsystems regarding material may not be completely reliable, and misjudgments or inputs transmission errors by sub-components may produce inaccuracies in the final results. Therefore, the resulting complexity makes responsibility division elaborate, and it becomes challenging to establish which link's concern gave rise to online misinformation. Taking a few microservices of AI as an example, provided that a specific microservice supplies erroneous information, other related microservices are likely to operate based on it, and the conclusions drawn by the structure may As a result be fabricated material.

The volatility of network situations themselves is a mandatory element contributing to network rumors [8]. Various phenomena like network delays, packet loss, and network attacks may modify the normal state of software programs, causing them to yield erroneous effects. Nevertheless, For instance, whenever network latency is high, on the condition that an algorithm does not receive a responsibility for a long time, it may use timeout processing procedures and secure output information that may not be correct via some default procedures. Network attacks like DDoS or man-in-the-middle attacks may also interfere accompanied by algorithm execution and present potential threats of causing fabricated information.

Taken together, the arising of internet rumors is the repercussion of the collective consequences of technical, decision-making prediction obstacles, various inducing elements, and organization advancement. These variables are interrelated and mutually influential, together explaining how AI systems can on their own provide rise to internet rumors. A apparent examination of this emergence is the starting feature and underpinning for considering criminal law oversight.

3. Criminal Law Application Challenges

3.1. Ambiguity of Behavioral Subject

Classically, the answerable party for network rumors is a natural person, and the rule of "no obligation in the absence of fault" demands an immediate correspondence between behavior and accountability [5]. In conventional network rumor crimes, judicial authorities might establish the identity of the perpetrator via technical means like IP tracking and account identification, and then pursue their criminal accountability. This characteristic of "determining the perpetrator" comprises the foundation for conventional criminal obligation pursuit, and the tenet of "no obligation in the absence of fault" mandates crimes to be perpetrated by apparently identifiable actors. In the scenario of independent algorithm running, the software makes its own decisions built upon computational logic in the absence of immediate human participation in material creation, and the "human" influence element is greatly weakened [7]. This outcomes in the ambiguity of the behavioral liable agent. More exactly, the ambiguity of the behavioral party is reflected in three perspectives: is a lack of an immediate causal connection between computer algorithm programmers and harmful consequences [8]. The duty of programmers is to design and evolve software frameworks, and their actions generally arise prior to algorithm deployment, with temporal and spatial gaps between their actions and the fabricated inputs generated during independent algorithm running. Platform operators locate it sophisticated to efficiently monitor the independent behavior of computer software frameworks; they do not immediately participate in the algorithm model design and choice-making running of software frameworks. Third, ordinary users lack substantial control over computer algorithm output; they cannot predict that software frameworks would automatically generate false output, let alone control the algorithm's output outcomes. These three extents of liable subjects all have mediated connections with harmful consequences. but none meet the demands of "immediate perpetrators" in conventional criminal law.

The ambiguity of the agent discussed in the In addition itself gives rise to a severance in the responsibility chain. Thus, computer software systems cannot have criminal accountability ability and cannot serve as answerable subjects; present accountability formats (criminal penalties or non-criminal penalties) are mostly designed for natural persons or units and are not straightforwardly applicable to self-operating software execution [9]. In terms of the present criminal system, accountability formats are focused on criminal penalties and non-criminal functionings, but such accountability strategies are comprehensively designed for natural persons or units and may not be applicable to self-operating AI runnings. How to establish a responsibility procedure suitable for AI society shall be an underlying issue that criminal legislation strives to resolve.

3.2. Lack of Behavioral Intent

Internet rumors are intentional acts perpetrated by natural persons, with all four elements fulfilled [2]. In established internet rumors crimes, natural persons or units with criminal obligation intentionally fabricate and spread false information to infringe on citizens order, personal reputation, and other legal interests. Their subjective intent to comprehend the harm and hope or allow the outcome to undertake place is the foundation for responsibility pursuit. The "autonomous running" of AI software structures is fundamentally a mechanical execution of computational logic. Software structures lack the ability to recognize and control, making it impossible to format subjective intent or negligence [10]. The behavior of computer software structures is mechanical and non-cognitive; they do not understand the meaning and outcomes of the content they generate. There is a core difference amidst AI algorithm behavior and established criminal behavior, which makes it difficult to incorporate algorithm behavior into the established criminal legal evaluation structure. Computer software structures lack the ability to recognize the nature of their behavior and cannot understand the legal meaning and social outcomes of the content they generate.

On the whole, the lack of behavioral intent is especially evident in three aspects: Computer programs lack the ability to recognize the nature of their behavior and cannot understand the legal meaning and

social consequences of the information they generate; software programs lack the ability to foresee harmful consequences; even provided that programs are endowed together with certain danger identification capabilities during design, such capabilities are limited to preset rules and patterns; software programs lack the ability to control their behavior; program running is automatic and continuous, and once started, it executes in line together with preset algorithm paradigms [11]. These characteristics create it impossible to evaluate AI program behavior as "intentional" or "negligent," and the lack of subjective intent deprives the tenet of guilt responsibility of its utilization underpinning.

3.3. Difficulties in Proving Causation

Conventional network rumors possess manifest causal connections, with performance and outcomes largely directly associated [11]. According to the conditional architecture, the perpetrator's fabrication and dissemination performance is the necessary condition for the purpose of the emergence and spread of rumors; built upon the system of equivalent causation, the performance and outcomes are equivalent, and the causal relationship might be elaborated or verified via an testimony chain.whenever software frameworks operate independently, the "black box effect" hides causality [12]. The unexplainability of paradigms makes abnormal performance lack regulatory targets, and algorithms themselves may generate new dangers. The decision-making procedure of LLM paradigms is similar to a "black box"; even program designers may not fully understand its logic. This unexplainability hinders the timely identification and correction of anomalies in software frameworks and also prevents judicial authorities originating extracted from making causal correspondences amidst particular runnings and false output. So-called algorithmic bias is the inherent system deviation in decision-making, which may be caused by insufficient training materials or algorithm design defects, Hence promoting the formation of false information. The unexplainability of schemas presents a major challenge for the purpose of responsibility division, and this "black box" effect further drags causal verification into problems. Fundamentally, on one hand, it is impossible to locate an exact causal relationship amidst a specific design or performance and false inputs; conversely, it is impossible to determine which part of the issue contributed to the erroneous speech. conventional causal relationship concepts possess shown their deficiencies in dealing with such intricate modern technological frameworks. Taking AI as an example, on the condition that it creates fabricated datasets, courts shall locate it sophisticated to verify whether this datasets resulted drawn from the designer's negligence, the service provider's improper management, or the user's misuse.

Originating from the viewpoint of proof law, causal relationships must be proven by sufficient proof. once the emergence of AI, the acquisition of such proof has become extremely challenging. The determinations made by programs are the effect of the joint execution of a large number, or even an extremely large number, of procedures, and the changes of each technique format a highly elaborate changing state. It is not easy for courts to acquire complete functioning logs or reasoning explanations; even provided that relevant materials are obtained, it may not be possible to identify the direct source proof of false inputs. Hence, the above obstacles in proof acquisition further plunge the determination of causal relationships into greater difficulty.

4. Criminal Law Response Pathways

4.1. Criminal Subject Status of AI and Criminal Liability Pursuit Logic

A debate exists in academic discourse between the affirmative structure and the negative structure regarding the party status of AI in crimes [13]. In particular, the affirmative system recommends recognizing AI as a legal person party, stating that, given existing technological advancement, software might already generate choices along with a specific degree of autonomy and must Hence be recognized as having legal personality. The negative structure maintains that even provided that AI has a high degree of independence, it is ultimately a human design and a human-controlled object,

and cannot possess true shall and subjectivity [14]. This paper adopts the negative architecture for the following reasons: Criminal obligation competence consists of the ability to recognize and control, and AI programs cannot recognize the nature of what they do nor foresee harmful consequences, Hence lacking subjective fault [10]. Criminal punishment itself implies punishment and prevention; imposing Criminal penalties on programs has neither punitive significance nor helps prevent crimes. Treating AI as a Criminal entity would undermine the restraint spirit of criminal law and may give rise to an inappropriate expansion of the scope of Criminal sanctions. extracted from a basic criminal law viewpoint, Criminals shall be persons (or units) along with autonomous shall or organizational capabilities, whereas AI programs are essentially human creations.

To investigate criminal accountability for network rumors caused by the autonomous execution of AI programs, it is appropriate to follow the "obligation-danger-responsibility attribution" strategy [15]: clarify the safety obligations of software designers, managers, determine whether the parties possess failed to fulfill their duties and caused the legally prohibited danger to materialize; objectively verify whether the danger has a realistic foundation along with constitutive element significance; and, ultimately, determine the structure of responsibility based on the fault perpetrated and the magnitude of harm. This logic is not anchored in "actor's subjective intent" but is based on "duty violation and danger creation," offering a new direction for thinking approximately criminal accountability in AI scenarios.

For implementation, criminal responsibility must be reasonably distributed founded on the part and position of different parties in algorithm functioning. As the designers of software programs, developers must bear the duty of design reliability, ensuring that effective technical measures are taken during the design phase to stop the generation of fabricated output; as the managers of applications, service providers shall bear the duty of operational reliability, establishing effective monitoring procedures to promptly handle abnormal performance; as the direct users of computer programs, users shall bear the duty of safe use, using programs reasonably and avoiding misuse that creates false information. whenever these parties violate their respective safety duties and generate internet rumors, they ought to bear corresponding criminal accountability.

4.2. Establishing and Improving Program Security Guarantee Obligations

The conventional criminal law system based on "actor's subjective negligence" as the underpinning for accountability pursuit is challenging to apply in scenarios where AI programs operate on their own. In the example of self-operating software operation, the software itself does not maintain subjective intent, and the causal connection amidst the subjective fault of software designers and users and harmful outcomes is often indirect and sophisticated. This paper supports the idea of using software legal accountability as one of the solutions to the obstacles, replacing actor intent with violation of software security obligations as the foundation for accountability attribution [16]. The reasons are as detailed below: On one hand, the progression of AI techniques have allowed software programs to possess a particular degree of self-operating operation capabilities, and the existing accountability pursuit chain has broken owing to "ambiguity of behavioral liable subjects"; the self-operating operation of programs is not completely divorced extracted originating from human control but is closely connected to the design of software designers, the management of operators, and the conduct of users, and these parties bear unshakable responsibility for the safe operation of programs.

To examine the solutions to the problems included in program legal accountability, with respect to distinct structures, it is to establish and upgrade program security guarantee duties, which are elaborated in three aspects [17]: The safety duty during design demands programmers to employ methodology to avoid the emergence of fabricated information in the design; the operational safety duty is borne by the platform, monitoring executions and promptly handling abnormalities; users shall meet basic safety requirements whenever using software structures. These three duties structure a complete guarantee organization, clearly and specifically regulating how each relevant party must assume responsibility.

Based on program security guarantee obligations, drawn originating from the viewpoint of institutional design: the specific content and benchmarks of the program security obligations of each liable entity ought to be defined, so that each party obviously knows which obligations they ought to fulfill; effective supervision procedures ought to be established to guarantee that all parties efficiently fulfill their security guarantee responsibility; Last but not least, strict accountability pursuit executions must be established to severely punish violations of security guarantee obligations, and it is recommended to add a "crime of violating arrangement security legal obligations" in criminal law [18].

4.3. Human-Machine Hybrid Subjective Psychology and Objective Imputation

The "black box" hurdle casts the validation of subjective fault into a quandary of human-machine hybridity. The choices made by computer programs are intertwined with the designer's ideas or usage approaches, making it impossible to discover an evident attachment aspect or reference for subjective intent. In the AI context, relevant psychological states have characteristics of "human-machine hybridity" and cannot be fully clarified by established subjective fault theories. Given the present situation, the black box impasse ought to be primarily addressed via the notion of objective responsibility attribution [19]. The logical arrangement built upon "creating hazards-realizing dangers-scope of constitutive element effectiveness" might avoid the hurdles in confirming subjective negligence and directly party the legal dangers involved in performance to objective examination. Objective responsibility attribution shifts the center of imputation originating from the subjective to the objective leveled. More distinctly, the implementation of the objective imputation viewpoint should divid into three steps: determine whether the actor has violated software guarantee obligations and caused prohibited perils; determine whether the peril has arisen inside the scope permitted by constitutive elements; and exclude other non-imputation components. The analysis Hence obtained might better address the proof problems described by the "black box."

In implementation, the acceptance of the objective obligation attribution notion needs to be accurately executed in combination with the characteristics of AI techniques. whenever examining "hazard creation," components including design defects of software algorithms, abnormal operating conditions, and improper user operations ought to consider; whenever examining "danger realization," components including the creation procedure of false inputs, dissemination procedures, and harmful consequences ought to consider; whenever examining "exclusion of accountability attribution components," variables for instance force majeure, accidental events, and victim fault shall consider. through these specific assessments, the criminal obligation of each party might be ascertained more accurately.

4.4. Reasonably Setting Criminalization Standards and Applying Relevant Charges

Social harmfulness is the indispensable quality of crimes, but its meaning in the AI environment faces dual problems. The Initially is the paradox of legal acts aggregating to make up crimes: individual acts are legal, but together they may violate the law; the Subsequently is the quantification difficulty—the scope or harm involved cannot be evidently counted. The tenet of restraint in criminal law ought to be applied to handle instances where software systems automatically generate false information, and criminalization shall not be rash. Social harmfulness shall reach a high leveled prior to criminal penalties may be applied, and technological progress ought to be given appropriate room for progress. The distinction amidst crime and non-crime shall analyz comprehensively grounded in the scope of dissemination, harm, and subjective state. Regarding the obligations of computer software designers or platforms, they must measur by the criterion of reasonable care; as for users, it shall evaluat whether they knowingly publish false information. At the same time, a graded strategy may be evaluated to mitigate the issue, integrating the information of false information, the breadth of dissemination, and the actual harm to determine whether criminal procedures shall be initiated.

On the whole, online misinformation may primarily make up the crimes of libel and slander, fabricating and intentionally spreading false inputs, picking quarrels and provoking trouble, and

infringing on citizens' personal information. The key to distinguishing among these crimes lies in the different legal interests they protect. Hence, in the implementation of charges for online misinformation, it is urgently mandatory to centre on the distinction of legal interests. The crime of libel and slander protects citizens' reputation rights, which belong to personal legal interests; the crime of fabricating and intentionally spreading false information protects citizens order in specific situations; the crime of picking quarrels and provoking trouble protects general public order in comprehensive situations. In the AI scenario, the utilization of charges for online misinformation ought to pay attention to the following issues: for online misinformation on their own generated by applications, the charge ought to be determined grounded in the actual legal interest infringed; for the acts of different actors, the charge should determin based on their position in the generation and dissemination of false datasets, like designers and users; attention shall be paid to the boundaries among charges to avoid double evaluation.

5. Conclusion

This study advances three theoretical innovations for criminal law governance of AI-generated network rumors: (1) a "program legal liability" framework built upon "violation of program security obligations"; (2) a dual-track accountability system distinguishing "actor liability" from "program liability"; and (3) the "algorithm risk creation" notion to advance the objective imputation construct in the AI era.

For implementation, this paper recommends: (1) a multi-degree security obligation system for developers, operators, and users; (2) clarified identification standards for "algorithm crimes" with explicit criminalization thresholds; and (3) a specialized judicial process with technical expert participation.

This study has three constraints: (1) theoretical analysis without judicial empirical information; (2) no differentiation between general-purpose and industry-specific AI models; and (3) no discussion of cross-border jurisdiction. Future investigations shall address these gaps through empirical scenario studies, differentiated regulatory strategies, and international cooperation.

References

[1] Liu RW. Criminal law threat and imputation of AI technical entities[J]. Political Science and Law Forum, 2024, 2(2): 5-20.

[2] Pi Y. On AI security management obligations and criminal liability of bearers[J]. Criminal Law Review, 2024, 1(1): 23-45.

[3] Zhou GQ. Theoretical transformation of criminal law in the AI era[J]. Science of Law, 2024, 1(1): 34-56.

[4] Liu YH. Research on AI criminal law[J]. Peking University Law Journal, 2024, 3(3): 789-812.

[5] Zhang BS. Research on AI legal liability[J]. Social Sciences in China, 2024, 5(5): 67-86.

[6] Zheng G. Analysis of legal subject qualification of AI[J]. China Law Journal, 2024, 2(2): 345-368.

[7] Wang ZX. Criminal law governance of network rumors[J]. Political Science and Law, 2024, 3(3): 78-95.

[8] Tian HJ. Causation judgment in the AI era[J]. China Legal Science, 2024, 1(1): 156-178.

[9] Lin XX. On the construction of AI legal personality[J]. Modern Law Science, 2023, 6(6): 112-130.

[10] Chen JH. Legal status of AI[J]. Chinese Journal of Law, 2024, 1(1): 56-78.

[11] Xia Y. On the case of Deng Yu-jiao and the theory of objective imputation by Claus Roxin[J]. Northern Legal Science, 2009, 3(5): 31-39.

[12] Bai LT. The delayed consequence of cases and the question of the conclusion about accomplished offense[J]. The Jurist, 2016, 31(1): 161-174.

[13] Sun YL. The objective imputation of intentcrime and defectcrime[J]. Journal of Henan University of Economics and Law, 2012, 27(3): 94-103.

[14] Mitchell A. Artificial intelligence and criminal responsibility: A challenge to the traditional framework[J]. Journal of Criminal Law & Criminology, 2023, 113(4): 789-812.

[15] Fincan M. AI and legal issues: A review of AI-based legal impasses in terms of criminal law[J]. Journal of Law and Technology, 2023, 15(2): 201-218.