Science of Law Journal (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/law.2025.040405
ISSN 2616-2296 Vol. 4 Num. 4

Legal Risks and Response Strategies of Generative Artificial Intelligence

Xu Wenjie, Liao Fang, Wang Wei*

School of Law, Southwest Medical University, Luzhou, Sichuan, China
**Corresponding author*

Keywords: Generative Artificial Intelligence; Legal Risks; Response Strategies

Abstract: As a product of social progress, generative artificial intelligence has brought about significant impacts on human society. However, it has also triggered a series of legal risks, including the legitimacy of data sources, false information and improper use, and copyright disputes. In response to the aforementioned legal risks, based on the analysis of the current legal situation and deficiencies in China, this paper proposes strategies such as formulating specialized laws, clarifying the subjects of infringement liability, defining the principles of liability for infringement, and clarifying the burden of proof, with the aim of promoting the healthy development of generative artificial intelligence.

1. Introduction

Generative artificial intelligence, which summarizes and generalizes existing data based on user instructions and uses deep learning technology on that basis,[1]has brought great convenience to human life. Once launched, it has sparked extensive discussions among scholars in many fields, among which the most discussed in the legal community is the legal risks of generative artificial intelligence? How to deal with the legal risks that arise? Therefore, it becomes very meaningful for this article to conduct an academic exploration of the legal issues derived from generative AI and propose legal response strategies, thereby promoting the progress of practice.

2. An overview of generative AI

2.1. The definition of generative artificial intelligence

Generate new data by learning the distribution of input data.[2]Unlike discriminative models, generative models can not only perform tasks such as classification and regression, but also create new samples similar to the training data. Typical techniques of Generative AI include Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and Autoregressive models Models.[3]

2.2. The technical workings of generative AI

The operation of generative AI is roughly divided into three phases. The first is the data collection phase. The second is the training phase. Pre-training is conducted on a well-organized dataset, using methods such as deep learning to obtain a general and highly generalized base model,[4]which is then optimized into a new model. The third is the generation stage, where AI creations are produced.

2.3. Generative AI algorithm black box

The algorithmic model used by generative AI today is the classic black box algorithm. [5]Because the underlying technology contains manifold distribution laws and learning probability distributions, the technical theory on which it is based is not yet fully developed. Currently, there is no mature technical solution that can provide a high level of explanation for the algorithmic black box. And imperfect explanatory solutions can also reduce users' trust in the intelligent technology, mislead users and cause a series of adverse effects. Therefore, in the context of the continuous growth of training corpora and model parameters, the interpretability of generative Al models will be constrained, and unless there is a breakthrough in related interpretability technologies, uninterpretability will remain an essential attribute for a long time.[6]

3. Legal risks of generative AI

3.1. Risks regarding the legality of the data source

One of the key elements of generative AI is to have a sufficient base of nourishment - data. The sources of data for training generative AI databases can be divided into two major parts: one is data authorization obtained through contracts, and the other is massive data obtained from web crawlers. However, the legitimacy of the sources of data obtained through these two methods remains to be considered. If the data crawled is from a public database, for example: There is generally no legal risk of infringement if the public information released by the government or other parts is used within the scope permitted by law, but if the content is from a non-public database (such as works, portraits, voices, etc.), it may violate the data usage regulations of the relevant platform and may also infringe others' Copyrights, personality rights, etc.[7]

3.2. The risks of generating false content and improper use

A law professor at the University of California, Los Angeles conducted a study, asking the AI chatbot ChatGPT to answer whether there were issues of sexual harassment among law school professors in the United States; please provide at least five examples and cite relevant newspaper articles. The name of a certain law professor appeared in ChatGPT's response. According to ChatGPT, the professor was accused of sexual harassment and was embroiled in a lawsuit, describing contact with law school students during a class trip to Alaska, and citing an article from The Washington Post as the source of information. However, neither the article nor the matters mentioned in ChatGPT's response actually existed. The professor said he had never been to Alaska with his students, The Washington Post had never published such an article, and he had never been accused of sexual harassment or assault by anyone. [8]With the development of generative AI, there are many such incidents. The existence and content output of generative AI rely on huge datasets, and the content it generates is inevitably influenced by these data, which are already illegal, non-compliant, false and incorrect when obtained. Users do not identify the authenticity of the content, and in today's Internet age, the output of false content has the potential to spread widely, affecting the judgment of those who receive the false information. While generative AI brings convenience, it can also be exploited by some ill-intentioned people to achieve their inappropriate or illegal purposes. Some lawbreakers use AI to generate false information that is detrimental to national security and spread it widely, causing social unrest.[8]

3.3. The risk of data privacy breaches

From the operation mode and characteristics of generative AI, it can be seen that "data privacy leakage" has always been a major problem with this technology as it has become widespread in various fields. First, in the process of human-machine interaction between users and generative AI, records such as instructions input by users, personal information, private content, and business secrets may be retained. This increases the risk of user privacy leakage. Second, the systems of generative AI may have algorithmic flaws or be hacked, and the attackers can obtain information from them, leading to the leakage of users' privacy. For example, less than 20 days after Samsung Electronics introduced ChatGPT, it was exposed that internal confidential information was leaked, involving semiconductor equipment measurement data, product yield, etc. Despite the growing emphasis on data security and the maturation of protection techniques, it is still difficult to withstand the evolving cyber malicious attack techniques, resulting in a non-negligible risk of data privacy leakage for generative AI systems.

4. Legal risks of copyright infringement

The technical principle of generative AI models is to learn from existing data and then generate new data or complete tasks based on these patterns and characteristics, a process known as data "feeding". From the perspective of this characteristic of generative AI models, the first major risk of copyright infringement is concentrated at the input end of the training data. Generative AI requires a large amount of data collection during the training process, including copyrighted works, which are not all authorized and are collected through web crawlers for use in generative humans According to existing laws, artificial intelligence does not fall under fair use of works. On the other hand, the infringement of copyright is due to the potential risk of infringement in data collection, the content it generates may be the same as or substantially similar to others' works, and due to the lack of infringement review and filtering mechanisms. There is a possibility that the generated content may directly apply unauthorized text or patents. When users utilize generative AI, the possibility of infringing others' intellectual property rights will be at a high level if there is no effective compliance review of the relevant content or if it is simply difficult to conduct an effective compliance review.[9]

5. The causes of legal risks associated with generative AI

5.1. High algorithmic trust

The agile processing of decision requests by algorithms makes decisions efficient and real-time. When faced with algorithms that can both reduce decision-making costs and improve decision-making efficiency, it is difficult not to trust and rely on them. Based on reliable trust in generative AI, users input information that should not be made public, such as personal privacy or business secrets, and use AI to generate content without doubt based on trust. However, algorithms are not always reliable, and the credibility of the algorithms themselves remains questionable. [10]And for some questions, generative AI "may produce answers that seem reasonable but are actually wrong." That is to say, the misinformation it outputs can easily lead users to have a wrong perception that it is correct because of its complete logic and coherence, and over time, human subjectivity will be threatened.

5.2. Legislation lags behind technological development

The social problems caused by technology drive the renewal of legal norms, and the existence of legislative lag makes existing legal norms naturally unable to keep up with technological development. [11]The existing legal framework is unable to adapt to the rapid development of its technology and gradually loses control of new developments. This phenomenon is to some extent due to the lag of the law. In the process of the evolution of the law, it is usually not until new problems or more serious social phenomena occur that adjustments are made through legislation or judicial interpretations. However, this "diagnosis after illness" approach is difficult to maintain social stability in most cases, but it seems inadequate in the context of the current rapid development of generative artificial intelligence.

6. The current state of legal regulation of generative AI

At the macro level, The State Council, the National Committee for the Governance of New Generation Artificial Intelligence, and others have successively issued several policy documents on the governance of artificial intelligence, including the "New Generation Artificial Intelligence Development Plan" and the "Principles for the Governance of New Generation Artificial Intelligence - Developing Responsible Artificial Intelligence". [12]At the micro level, in recent years, China has successively introduced the "Regulations on the Administration of Deep Synthesis of Internet Information Services" and the world's first specific legislation on generative artificial intelligence, the "Measures for the Administration of Generative Artificial Intelligence Services", which shows that China has been constantly optimizing and improving the governance system of artificial intelligence and exploring actively.

7. Legal response strategies for generative AI

7.1. Develop specific laws for generative AI

While existing legislation can regulate the development and application of generative AI and regulate the risks it brings, it is still difficult to deal with concrete and realistic problems. Therefore, it is necessary to continue to explore feasible risk regulation paths. Specialized laws on AI governance should be developed to regulate the legal risks associated with generative AI from the source.

7.2. Clarify the subject of tort liability for generative artificial intelligence

To effectively address the legal risks of infringement by generative AI, first, copyright infringement. Second, generation and dissemination of false information. Third, data privacy leaks. From the three types of infringement legal risks mentioned above, it can be concluded that the liability subjects that may constitute infringement by generative AI are the generative AI itself, the service provider and the user. Therefore, it is necessary to first discuss whether the AI has legal subject status. Currently, there are the following theories in the academic community. Some scholars believe that AI is still an artificial creation. Unlike other tools used by humans, AI can only be a tool and not a subject. Scholars who oppose the tool theory argue that AI has developed beyond human intelligence in some areas and its tool attribute has weakened significantly, but this does not mean that AI can replace humans to carry out its role in social relations and have legal subject status. It does not have self-awareness. All its actions are the result of the evolution of algorithmic big data. If it is given legal personality to achieve the effect of liability limitation, it does not meet the basic requirements of fairness and justice. Therefore, because it does not have legal subject status itself, the main responsible subjects involved in its infringement are service providers and users.

7.3. Define the principle of liability for infringement by generative artificial intelligence

In the tort liability system for generative AI, a balance should be struck between innovation incentives and risk prevention. If strict liability principles are adopted and producers are required to unconditionally bear the damage caused by inherent defects, it will dampen the willingness to release technology and hinder development. Given that AI is essentially an information tool that provides a reference of value rather than a direct threat to personal and property safety, users still have the obligation to judge. Unless there is a significant infringement of rights and interests, the law should apply the principle of fault liability.

7.4. The burden of proof for generative artificial intelligence is statutory

The burden of proof for infringement by generative AI should be distinguished from that for product liability. Due to the black box operation mechanism within generative AI models and the lack of technology to make them explainable at present, the difficulty of collecting evidence to prove the subjective fault of the infringer for ordinary users outside the industry can be imagined. Due to the inherent disadvantage of the victim, the process of presenting evidence is difficult, and considering the unexplainability of generative AI algorithms, the principle of inversion of burden of proof should be adopted in terms of burden of proof for generative AI infringement.

8. Conclusion

To sum up, this paper first understands generative artificial intelligence and discusses the specific characteristics of the two major types of legal risks of generative artificial intelligence, and proposes countermeasures in the legal aspect. However, the countermeasures constructed inevitably have problems of insufficient foresight and adaptability. Therefore, the solutions proposed in this paper are not a one-off effort and need to be constantly adjusted and optimized in accordance with the development trend of generative AI and the needs of its risk response in order to ensure the healthy development of "technology for good" of generative AI.

References

[1] Liu Yuenan, Qian Yi, Wang Ping, et al. Challenges and Prospects: The Impact of DeepSeek on Archival Work and its application prospects [J]. Zhejiang archives, 2025, (02): 5-13. DOI: 10.16033 / j.carol carroll nki/g2.20250305.0 33-1055 01.

[2] Wenzuei chang. The theory of emergent artificial intelligence service providers tort liability [D]. Hebei university, 2024. The DOI: 10.27103 /, dc nki. Ghebu. 2024.000828.

[3] Zhang Zheng. Research on security protection technology for generative artificial intelligence and Data privacy [J]. Network Security Technology and Application,2025,(04):75-77.

[4] He Qiongqiong, Yue Chunxiao. Copyright risks and their resolution in generative AI creation [J]. Journal of Zhejiang Shuren University,2025,25(01):70-81.

[5] Liu Xuerong. Technical governance and hierarchical Accountability of generative artificial intelligence [C]// Political and Legal Affairs Commission of the Hebei Provincial Committee of the Communist of China, Office of the Hebei Provincial Committee for Comprehensive Law-based Governance of the Communist of China, Hebei Law Society. Jilin University; 2023:344-349. DOI: 10.26914 / Arthur c. nkihy.2023.118502.

[6] Jin Longjun. Emergent AI cannot be interpreted and the rule of law dealing with [J]. The rule of law research, 2025, (02): 42-53. DOI: 10.16224 / j.carol carroll nki cn33-1343 / d. 20250303.004.

[7] Su Zilong. Research on Data Security Risk Prevention and control and Legal Regulation of Generative AI [J].

Communications and Information Technology,2024,(05):95-98+110.

[8] Liang Xuntong. On the Legal Risks and Regulation of Generative Artificial Intelligence [J]. Journal of Anhui Police Vocational College, 2024, 23(05):38-44.

[9] Sun Na, Bao Yiming. Research on the Legislative Model of Generative Artificial Intelligence in China from the Perspective of Comparative Law [J]. Journal of International Economic Law,2024,(04):53-70.

[10] Liu Yonghong, Li Wenying. Emergent path of legal risk and the regulation of the artificial intelligence [J]. Journal of neijiang normal college, 2024, 33 (09) 6:94-100. The DOI: 10.13603 / j.carol carroll nki. 51-1621 / z. 2024.09.015.

[11] Sun Guorui, Liu Weibin. Artificial intelligence technology governance and legal governance [J]. Science and technology and the law (both in English and Chinese), 2025, (02) : 1-10. DOI: 10.19685 / j.carol carroll nki cn11-2922 / n. 2025.02.001.

[12] Wang Xiumei, Zhang Xue. Criminal Responsibility and punishment for crimes involving artificial intelligence [J]. People's Procuratorate, 2023,(19):11-16