

Copyright Infringement Risks and Responses of Generative Artificial Intelligence

Li Zhi

School of Law/School of Intellectual Property, Guilin University of Electronic Technology, Guilin, China
3099365396@qq.com

Keywords: Generative Artificial Intelligence, Copyright Infringement, Fair Use, Legal Use of Data

Abstract: With the rapid development of generative artificial intelligence (AIGC) technology, its wide application in literature, art, music and other fields has raised many copyright infringement issues. The use of a large number of works by AI developers for large model training will bring about copyright infringement risks, but this behavior does not cause damage to the value of the works or the legitimate rights and interests of copyright holders. Based on the goal of promoting digital technology innovation and the development of the artificial intelligence industry, it is necessary to include it in the scope of fair use of copyright. Regarding this point, although there is room for application in the interpretative approach, there are also certain disadvantages. The use of works should be limited to the replication and translation of works in the training stage and should not include the dissemination through information networks. In addition, it is necessary to modify the obligation of attribution to the obligation of AIGC marking. This article starts from the basic concept of generative artificial intelligence, analyzes the possible copyright infringement risks involved in the data collection and content generation stages, and discusses the legislative and judicial practices at home and abroad in response to this issue. On this basis, it puts forward suggestions for improving China's relevant legal system, with the aim of providing legal support for the healthy development of artificial intelligence technology.

1. Introduction

In today's rapidly developing society, with the gradual advancement of artificial intelligence technology, people have enjoyed great convenience in all aspects of life. Among them, generative artificial intelligence (AIGC) has become a hot topic in today's society. By inputting instructions to it, generative artificial intelligence will process internally and finally generate the answers and content we need. However, in the process of generating results, AIGC mainly relies on using existing works as training data. Even after reprocessing, it will involve copyright infringement to varying degrees. This article aims to explore the copyright infringement risks of generative artificial intelligence, analyze the current situation of response at home and abroad, and then put forward legislative and practical suggestions for China in this field.

2. Overview of generative artificial intelligence

2.1. Basic concepts and types

Generative artificial intelligence (AIGC) refers to an artificial intelligence technology that uses machine learning algorithms to train models with large amounts of data and can generate new texts, images, audio, and other content^[1]. Generative artificial intelligence is a new technology that absorbs and then processes and creates. It deeply analyzes data and logic, and then generates corresponding style works based on human instructions. Generative artificial intelligence is an emerging research direction, and its main goal is to model samples and mine new features with similar characteristics^[2]. The key of generative artificial intelligence is its generation function, that is, it can generate new data similar but not identical to the original. Its key is to generate some innovative and diverse new information by learning and simulating existing information. According to the generated content, AIGC is also divided into different types. For example, ChatGPT is a typical text generation artificial intelligence that can generate fluent text answers based on user input; while DeepSeek has strong natural language processing capabilities, can understand and answer questions, assist in writing code, organizing materials, and solving complex mathematical problems.

2.2. Application fields and development status

The scope of generative artificial intelligence is extremely broad, covering all aspects of life. Due to its outstanding creativity and intelligence, it can create more advanced and intelligent medical equipment in the medical field and provide better ideas for the development of medical research; in the field of market advertising, it can provide more comprehensive and intelligent solutions according to customer needs, etc. Moreover, the rapid development of the artificial intelligence industry has also injected new vitality into all walks of life, which can drive economic development and the progress of people's thinking.

3. Analysis of copyright infringement risks in the use of works by generative artificial intelligence

3.1. Using a work without permission may infringe upon the exclusive rights of the copyright owner

The learning of generative artificial intelligence often relies on a vast amount of text as samples. However, in real life, these samples are frequently used without authorization, which infringes upon the rights of many copyright holders. Copyright law grants authors various exclusive rights, including the right to reproduce, publish, and disseminate information online. When AIGC companies reproduce and disseminate works without permission, they commit copyright infringement.

3.1.1. Infringement risks in the data collection stage

During the data collection process, the working principle of generative artificial intelligence systems is to collect data through methods such as web scraping. These materials may include copyrighted content such as text, images, and sounds. If the sources of the collected materials are not properly verified or the corresponding authorization is not obtained, it may lead to illegal use of the materials. For instance, illegal web scraping of articles, images, etc. on web pages may infringe upon the copyright owner's rights of reproduction and information network dissemination.

3.1.2. The legal characterization of the act of using a work and the involved exclusive rights

The use of generative artificial intelligence to create works has certain particularities. On the one hand, the replication of works by AI during the training process is a mechanical and direct copy. On the other hand, when AI disseminates generated content through the Internet, it involves the infringement of the right of communication through information networks. Because the works produced by artificial intelligence may have similarities with those of other authors in terms of substance, and if they are published on the Internet without permission, it constitutes an infringement of their right of communication through information networks. When artificial intelligence machines are learning, we need to translate the original text into a form that machines can read and establish a training data database. Moreover, the use of works by artificial intelligence is not translation. Whether the creation of a work by artificial intelligence constitutes an adaptation must be analyzed in detail. In the sense of copyright law, adaptation refers to the re-creation of the original work without changing its essence, making it a novel and completely different work from the original. The process of artificial intelligence generating works is not based on the re-creation of a specific work. There are three situations: first, the generated work is completely different from the original; second, the generated work is similar in essence to the original; third, the generated work retains the basic expression form of the original and has originality.

3.2. Other infringement scenarios in the generation of content

3.2.1. Substantially similar to existing work

The content generated by generative artificial intelligence may be substantially similar to existing works. This is because during the training process, the AI model imitates and learns various aspects of existing works, which leads to the generated content being highly similar to the original works in some respects. If this similarity exceeds the necessary range, it may be substantially similar to the original work and cause infringement.

3.2.2. Infringement of the author's rights such as the right of authorship and other rights

In addition to the aforementioned infringement scenarios, generative artificial intelligence may also infringe upon the authorship rights of copyright holders. For instance, if the content generated by AI fails to attribute the original author or attributes it incorrectly, it may infringe upon the authorship rights of the copyright holder. Moreover, if the content generated by AI distorts or alters the original work, it may also infringe upon the copyright holder's right to protect the integrity of the work^[3].

4. Practices of dealing with copyright infringement by generative AI at home and abroad

4.1. Relevant foreign legislation and judicial practice

4.1.1. American legislation and case law

In addressing copyright infringement issues related to generative artificial intelligence, the United States mainly makes adjustments within the existing framework of copyright law^[4]. The U.S. copyright law centers on the principle of "fair use", which is determined based on "the purpose of use", "the nature of use", and "the extent of use". Courts will make judgments on fair use behaviors by comprehensively considering these factors. For instance, in some disputes involving AI training data, courts will determine whether the use of works by AI constitutes infringement based on the

principle of fair use. A case in point is "Tremblay v. OpenAI": On June 28, 2023, two authors, Paul Tremblay and Mona Awad, filed a class-action lawsuit against OpenAI. The plaintiffs claimed that OpenAI's ChatGPT had used their books for training data without permission and that ChatGPT could generate summaries of their copyrighted works, thereby concluding that OpenAI's actions were for commercial profit and infringed upon their copyrights. This case reflects the stance and practice that the plaintiffs demand strict compliance with copyright laws to protect their rights over their works, while OpenAI's unauthorized use of copyrighted works to train AI is considered an infringement. This indicates that creators in the United States believe that generative AI technology should develop without infringing on copyrights.

4.1.2. Legislation and Cases in Japan

Japan deals with copyright infringement issues arising from artificial intelligence under its current copyright law. The Japanese copyright law not only stipulates "fair use", but also defines the copyright issues of content generated by AI. Taking Japan's legislation as an example, under certain specific conditions, the copyright of works produced by AI belongs to the developer or user. In practice, Japanese courts have also explored the legitimacy of AI training data through several cases to guide the handling of related issues. For instance, the Osaka District Court of Japan conducted a trial on an infringement case generated by AI regarding the legality of AI training data. The incident was as follows: An independent illustrator in Japan found that a certain company had used AI technology to produce many images very similar to his original creations, with styles, compositions, and colors very close to the original paintings, and some parts even had significant overlaps with some of his past creative styles. The illustrator believed that the AI had produced a series of works after a long study of his works, which might involve copyright issues. Moreover, the company claimed that the works produced by AI were created by AI itself, which was an independent entity, and the company could not monitor the pictures learned by AI or check the sources of these training data one by one.

The Osaka District Court of Japan pointed out in its judgment that although the content generated by AI differs from conventional creation in terms of its generation mechanism, in this case, there is a high degree of similarity between the AI-generated work and the plaintiff's work, which is by no means a coincidence. During the training of the artificial intelligence, the plaintiff's creative content was widely used, resulting in a strong substantive similarity with the plaintiff's creation, which seriously infringed upon the plaintiff's copyright. The court also held that although the training and generation of works by artificial intelligence are carried out by computer programs, the selection and use of the data required for the program's calculations are still determined and ultimately responsible by the company.

4.1.3. EU legislation and case law

The EU has taken relatively proactive legislative measures in addressing copyright infringement issues related to generative artificial intelligence. Through legal documents such as the "Copyright Directive for the Digital Single Market", the EU has regulated the legality of AI training data^[5]. The directive stipulates that AI developers must obtain the permission of the copyright holder when using works as training data,^[4] except in cases of fair use. The LAION case in Germany involved relevant provisions of German copyright law regarding temporary reproduction and text data mining for scientific research purposes and their application in AI training data. The German court ruled that the establishment of a qualified training dataset did not increase economic harm to copyright holders and actually promoted information access in society, allowing temporary reproduction for scientific research purposes. This judgment provided a legal basis for the acquisition and use of AI training data under specific conditions, further clarifying the application of the temporary reproduction and

data mining exceptions in EU copyright law in the AI field. It helps to find a balance between technological development and intellectual property protection, provides certain legal guidance for the application of AI technology in scientific research and other fields, and also provides a reference example for other EU member states in handling similar issues, promoting consistency and coordination in the implementation of AI and copyright-related laws across the EU.

4.2. Domestic Response to Current Situation and Problems

4.2.1. The fair use system is inapplicable

With the rapid development of generative artificial intelligence technology, the protection of copyright is facing an increasingly severe situation, especially regarding the "fair use" rule for copyright infringement. This is difficult to apply well in real life. During the learning and training of generative artificial intelligence, a large number of copyrighted works are usually used^[6]. However, among the "fair use" and "legitimate authorization" situations listed in the current "Copyright Law", there is no provision that can include the training materials of artificial intelligence. This results in the inability to consider the use of others' creations for training as "fair use". Considering the actual situation, the reasons for copyright infringement by generative artificial intelligence specifically include: first, the use of works by generative artificial intelligence often has a commercial nature, which does not meet the non-commercial use requirement in fair use; second, the way it uses works is usually large-scale batch use, which will have a significant impact on the market interests of copyright holders, not meeting the condition of having a small impact on the market in fair use; third, the creative process of generative artificial intelligence is relatively complex, making it difficult to accurately determine whether the works it uses fall within the scope allowed by fair use.

4.2.2. The legal licensing system is rather difficult to implement

Firstly, the application conditions of the statutory licensing system in China's current Copyright Law are rather strict, and the creative process and usage scenarios of generative artificial intelligence are difficult to meet these conditions^[7]. Statutory licensing is usually applicable to specific or pre-defined scenarios, such as compiling textbooks and reprinting in newspapers and periodicals. However, the training and usage scope of generative artificial intelligence is too broad to be classified into specific scenarios. Due to the significant differences between its massive information collection and processing behaviors and the application scope of traditional statutory licensing, it is difficult to directly apply statutory licensing.

Secondly, the data sources for generative artificial intelligence are complex and vast, involving a huge number of copyright holders and works. Under the statutory licensing system, the identities of copyright holders must be clearly identified and payment made, but in practice, this has encountered significant difficulties. Firstly, it is very difficult to determine the copyright holder of each work, which is almost impossible when the data sources are extensive. Additionally, even if the copyright holder can be identified, how to make the payment to meet the reasonable calculation of licensing costs remains a challenge. The complex training scenarios and high transaction costs make it difficult to effectively implement the statutory licensing system in the field of generative artificial intelligence.

Furthermore, the implementation of the statutory licensing system also requires a sound registration and supervision mechanism, but the current laws and regulations are not yet complete. Due to the partial absence of the copyright registration system, many works have not been registered, which brings great difficulties to the identification of copyright holders. Moreover, as China has not yet established a reasonable and fair cost-sharing system and a sound supervision and management system, even if the legal permission is theoretically feasible, it is very difficult to implement in

practice.

4.2.3. The current situation of industry self-discipline and regulation

Currently, the issue of copyright infringement by generative artificial intelligence has drawn widespread attention. While the current state of industry self-regulation and supervision is gradually improving, it still faces challenges. In terms of industry self-regulation, in August 2024, the "Self-Regulatory Initiative for the Generative Artificial Intelligence Industry" was released, emphasizing the legal and compliant use of data and algorithm models. It requires enterprises to strictly abide by laws and regulations, ensure the legality of data sources and the clarity of their uses, and adopt technical measures to prevent data leakage and abuse. The initiative also calls for promoting the construction of a content ecosystem, preventing the generation of inappropriate content, driving technological innovation and quality improvement, adhering to ethical and moral standards, and facilitating exchanges, cooperation, and open collaboration^[8]. This initiative was jointly compiled by several well-known enterprises and universities, providing clear self-regulatory guidelines for the industry's development.

In terms of the current regulatory situation, China has initially established a regulatory framework based on the Cybersecurity Law, the Data Security Law, and the Personal Information Protection Law, supplemented by regulations such as the Interim Measures for the Administration of Generative Artificial Intelligence Services, which clearly defines the legal boundaries for the use of generative artificial intelligence^[9]. However, there are still issues in China's copyright management, such as unclear copyright ownership and difficulties in determining liability for infringement. Additionally, due to insufficient technological means, it is also challenging to accurately identify and effectively supervise cyber attacks.

5. Suggestions for Dealing with Copyright Infringement by Generative Artificial Intelligence

5.1. Legislative progress: Adding rules for the fair use of works used in training artificial intelligence

To effectively address the issue of copyright infringement by generative artificial intelligence, China should consider adding a fair use rule for AI training works in the Copyright Law. The addition of such a rule for AI training works requires a comprehensive consideration of the demands of technological innovation and the protection of copyright holders' rights. Firstly, based on the revision of Japan's Copyright Law, "text and data mining" can be included in the scope of fair use, and its applicable scope should be defined. Secondly, the "Regulations for the Implementation of the Copyright Law of the People's Republic of China" should add special provisions, clarifying that AI developers can use published works for training under certain circumstances, such as for non-profit scientific research and limited reproduction. At the same time, the "purpose of use", "proportion of use", and "whether it affects the market price of the original work" should also be defined. Additionally, this project will establish an open information source and regulate it to reduce the risk of information leakage.

5.2. Progress in data regulations: Improving the legal utilization mechanism of data for artificial intelligence training

Improving the legal utilization mechanism of data for artificial intelligence training is the key to addressing copyright infringement issues. China should clarify the legal requirements for AI training data through legislation, stipulating that AI developers must obtain the permission of the copyright

owner when using works as training data, except in cases of fair use. Additionally, a mechanism for reviewing the legality of data sources should be established, and the data collection process should be strengthened in supervision to ensure the legality of AI training data. Improving the legal utilization mechanism of data for artificial intelligence training is an important guarantee for promoting the healthy development of the artificial intelligence industry^[10].

Firstly, the legality of data sources should be clarified to ensure that the collection, storage and use of data comply with laws and regulations. The Interim Measures for the Administration of Generative Artificial Intelligence Services require that platforms use data from legal sources and follow the "notice-consent" principle, especially for data related to users, which should be given high priority. Secondly, a hierarchical and classified management system based on data should be established, and differentiated regulatory measures should be adopted for different types and levels of data^[11]. For example, when it comes to high-risk data, the transparency of the data and data management regulations must be strengthened to ensure the quality and representativeness of the data^[12]. On this basis, the norms for anonymous authentication should be further improved to ensure that users' rights and interests are not infringed upon in applications. At the same time, technological and legal methods should be organically combined. For instance, blockchain can be used to record the source and use of data to ensure the traceability and transparency of the data.

5.3. The Inevitable Scheme for Establishing a Reasonable Use Rule of AI Training Works in China

5.3.1. Clarify the subjects and scope of the reasonable use of works for training artificial intelligence

When formulating the rules for the fair use of works in the training of artificial intelligence, it is necessary to clearly define them and specify the scope of their application. The subjects of fair use should be the developers and users of artificial intelligence, and the objects should be the acts of copying and disseminating works during the training of artificial intelligence. Therefore, the characteristics of artificial intelligence should be fully considered, and corresponding adjustments should be made around these characteristics. It should be correctly defined without overly restricting it to promote the healthy development of artificial intelligence.

5.3.2. Clarify the requirements for the legality of the source of works

To ensure the legality of the training data for artificial intelligence, it is necessary to legalize it. When developers use it as training material, they must first conduct a legal review of the data source to ensure its legitimacy. For instance, by signing licensing agreements with copyright holders or protecting published works through copyright, such measures can provide legal protection.

5.3.3. Amend the obligation of attribution to the obligation of "AIGC" labeling

In addressing copyright infringement issues related to generative artificial intelligence, the obligation of attribution should be revised to an "AIGC" labeling obligation. AI developers must clearly label the content as generated by artificial intelligence and indicate the source of the works used when generating content. Revising the attribution obligation to an AIGC labeling obligation is an important measure to address the legal challenges brought about by technological development^[9]. This revision is not only an inevitable choice for technology to adapt to the law but also an effective path to balance technological innovation and rights protection.

In practice, the AIGC labeling obligation should be clear and operational. First, it should be marked through prominent methods, such as adding a label like "Generated by AI" in the produced

content, to ensure that users can easily recognize it. Second, the annotation should cover the entire process of the generated content's release, including adding information about the comprehensive attributes of the generated content in the metadata of the document. Additionally, technologies such as digital watermarking and blockchain should be utilized to enhance the credibility and traceability of the labels.

5.3.4. Pursuing the liability for infringement of network users and network service providers

When network users utilize artificial intelligence tools, there is a risk of publishing illegal information or creating content through AI. For instance, if a user uploads copyrighted works to an AI platform without permission or uses AI technology to generate content that is substantially similar, it will be regarded as direct infringement. According to the Copyright Law, users must bear the responsibility of ceasing infringement and compensating for losses. If the user's actions are intentional or involve gross negligence, they may face higher compensation liability. For network service providers, their liability for compensation is determined based on the type of business they offer and whether they have fulfilled reasonable due diligence obligations. As for websites that directly engage in creation or provide infringing products on the internet, they may be held liable for direct infringement. For example, in the "Ultraman AI Model Infringement Case", the platform was found to have helped infringement as it failed to review the infringing images uploaded by users and generated infringing content through technical means. If a service provider merely offers technical assistance without direct involvement but fails to fulfill necessary due diligence (such as not establishing an infringement complaint mechanism or not taking necessary measures to stop infringement)^[13], it may result in indirect infringement. For instance, in the "AI Text-to-Video Infringement Case", the company's failure to take reasonable precautions and preventive measures constituted an infringement of the right of information network dissemination.

According to the Civil Code, the Copyright Law and relevant judicial interpretations, network service providers shall bear corresponding responsibilities when they know or should know about users' infringement. In addition, the Interim Measures require service providers to establish and improve complaint and reporting mechanisms and handle infringement complaints in a timely manner. If service providers fail to fulfill this obligation, they may be deemed to have been at fault and bear the liability for assisting in infringement.

To effectively protect their rights under the law, copyright holders should proactively collect evidence such as the fixation of infringing content, users' usage behavior and the operation mode of the platform^[14]. Meanwhile, service providers should improve their internal management mechanisms and establish infringement early warning and handling mechanisms to reduce the risk of infringement^[15].

6. Conclusion

To achieve a balance between technological innovation and copyright protection, China should take multiple measures and address the issue of copyright infringement by generative artificial intelligence through a systematic legal and regulatory framework. Firstly, the rapid development of generative artificial intelligence technology poses a severe challenge to the existing copyright system. The training of artificial intelligence is based on a vast amount of data, which often includes copyrighted works. Therefore, it is necessary to consider how to effectively utilize such data without infringing upon others' copyrights. China should consider adding a fair use rule for AI training works in the copyright law, which indicates that the current copyright law has certain gaps and deficiencies when dealing with this emerging technology. The addition of fair use rules will provide clear legal guidance for AI developers.

Secondly, improving the legal utilization mechanism of artificial intelligence training data is the core to addressing infringement issues. This not only involves clarifying the legal requirements for AI training data but also defining the subjects and scope of reasonable use. Requiring AI developers to obtain the copyright owner's permission when using works reflects respect and protection for the copyright owner's rights. Additionally, clearly stipulating the legality requirements for the source of works is an important step in preventing infringement. By enacting legislation to clarify the legality requirements for the source of works, it can prompt AI developers to be more cautious when obtaining data, ensuring that the data used is legally authorized, thereby reducing the occurrence of infringement from the source.

On this basis, by legally regulating Internet users and Internet service providers, the normal operation of the copyright system can be guaranteed. Due to the development of the Internet, the spread of infringement is getting faster and wider. Therefore, it is necessary to define the responsibilities of Internet users and website service providers so that they can compensate for their actions when infringement occurs. This can not only effectively curb infringement but also enable copyright holders to obtain more rights^[16].

In conclusion, in dealing with copyright infringement issues caused by generative artificial intelligence in China, it is necessary to improve a series of systems and rules through legislation. This includes adding fair use rules, clarifying the requirements for data legality, defining the subjects and scope of fair use, and pursuing liability for infringement, etc. Only through these systematic measures can the protection of copyright be ensured while promoting the healthy development of artificial intelligence technology and achieving a positive interaction between technological innovation and legal protection.

References

- [1] Xu Enping. *Risks and Solutions of Generative AI in Judicial Applications* [J]. *China-Arab Science and Technology Forum (Chinese and English)*, 2024, (0¹): 149-155.
- [2] Huang Gaole. *Empowering the Creation of Popular Science Content with Generative Artificial Intelligence Technology* [J]. *Science Education and Museum*, 2024, 10(01): 22-27.
- [3] Wang Qian. *On the Right of Modification of Software Works: A Review of Recent Cases Including the "Rainbow Case"* [J]. *Jurist*, 2013, (01): 135-147 + 179-180.
- [4] Chen Liming. *Copyright Infringement by Generative AI: Types, Determination and Liability* [J]. *Knowledge and Action*, 2024, (01): 52-62.
- [5] Cai Di. *Research on the Reasonable Use of Copyright in AI Creation* [D]. *South China University of Technology*, 2022.
- [6] Zhang Tao. *Legal Risks and Inclusive Prudent Regulation of Generative AI Training Data Sets* [J]. *Comparative Law Studies*, 2024, (04): 86-103.
- [7] Wang Shengjie. *Reasonable Boundaries and Infringement Regulation of Text and Data Mining by Generative AI* [C] // Shanghai Law Society. *"Smart Rule of Law" Journal 2024, Volume 1 - Proceedings of the 2024 World Artificial Intelligence Conference Rule of Law Forum*. Zhongnan University of Economics and Law; 2024: 213-225.
- [8] Wu Qiong. *Technology for Good: The Core Meaning and Theoretical Value of China's Global AI Governance Initiative* [J]. *Peace and Development*, 2024, (05): 25-49 + 179-180.
- [9] Luo Shijie, He Guorong. *Legal Responses to Data Compliance in Fintech during the Generative AI Era* [J]. *Financial Development Research*, 2023, (11): 82-90.
- [10] Chen Bing, Fu Xiaou. *Legal Tone and Development of Data Training for Generative AI* [J]. *Journal of Liaoning Normal University (Social Sciences Edition)*, 2024, 47(03): 1-10.
- [11] Zhou Yejun, Li Shan, Wang Conghu. *On the Construction of Data Element Circulation and Trading System in the Field of Public Resources Transactions* [J]. *Bidding & Procurement Management*, 2023, (10): 35-38.
- [12] Wang Dan, Zhang Ding. *A Preliminary Analysis of the Construction Ideas of Virtual Teaching and Research Rooms in Colleges and Universities Based on Blockchain Technology* [J]. *Journal of Hubei University of Economics (Humanities and Social Sciences Edition)*, 2024, 21(05): 114-116.
- [13] Wang Xiaoqing. *Research on the Improvement of Civil Liability System for Live-streaming E-commerce Platforms* [D]. *Hebei Normal University*, 2023.

- [14] Liu Yunkai. *The Reasonable Use Approach of Copyright for AI Training Works* [J]. *Journal of Northeastern University (Social Sciences Edition)*, 2025, 27(01): 117-126.
- [15] Fan Xinghui. *On Copyright Infringement Risks and Responses in AI Creation* [J]. *Henan Science and Technology*, 2024, 51(24): 105-108.
- [16] Xie Yizhang. *Copyright Disputes and Resolution in the Training of Generative AI Works* [J]. *Chinese Editor*, 2024, (11): 38-46.