Dilemmas and Solutions of Artificial Intelligence Creation in Applying Fair Use System

Xinhui Du
China Jiliang University, Hangzhou 310000, Zhejiang, China,

Abstract: Artificial intelligence creation refers to the act of acquiring massive works in various ways, digitizing them and transforming them into data, then learning and training on the acquired data, and ultimately outputting the corresponding creations. In the process of AI creation, unauthorized access and use of works may lead to the risk of copyright infringement. However, the fair use provisions in the current copyright law do not provide corresponding exemptions for the infringement risks caused by the utilization of works by AI in the process of creation, which means that there is a dilemma in applying the fair use rules to the creation of AI. Countries and regions such as the United States, Japan and the European Union have already responded and adjusted to this. Therefore, China should clear the obstacles for the further development of AI creation by adding a new fair use scenario "computer analysis information", accelerating AI-related legislation, and further clarifying the determination of the three-step test in AI creation.

Keywords: Artificial Intelligence Creation; Work Utilization; Copyright Infringement; Fair Use

1. Introduction

In recent years, generative artificial intelligence such as ChatGPT and Stable Diffusion have begun to be popularized and applied. Compared with the previous analytical AI, generative AI has stronger imitation and generation capabilities. It can generate similar works by machine learning from already digitized works, and it may also have an impact on the potential market of the original works [1]. Therefore, the question of whether generative AI's use of other people's works as a source of training data and generation of new works without the copyright owner's permission constitutes fair use has become a new issue to which copyright law needs to respond urgently. At present, the fair use system of the traditional copyright law is not yet able to provide infringement exemption for generative AI to utilize other people's works without permission in the process of creation. However, once the relevant provisions of the current copyright law are strictly followed, the further development of AI may be hindered due to the huge cost of licensing and other problems. Therefore, there is an urgent need to clarify the reasonable boundaries of the use of other works by generative AI in the process of creation, so as to reconcile the conflict between copyright protection and industrial development.

2. Principles of Generative Artificial Intelligence Creation and the Risk of Copyright Infringement

2.1 Principle of Generative Artificial Intelligence Creation

The creation process of generative artificial intelligence can be summarized in three stages: data input - machine learning - result output [2]. In the data input stage, a large number of literary and artistic works are transformed into digitized works, followed by the digitized works will be transformed into recognizable data and stored in the database. In the machine learning stage, the AI system will use algorithms to analyze and organize the data stored in the database and train, and extract the corresponding information and rules to form the corresponding model, so as to facilitate self-creation in the later stage. In the output stage, the generative AI will directly output "works" that are similar to human creations in appearance based on the previous two stages.

2.2 Copyright Infringement Risks of Generative AI Creation

Generative AI relies heavily on input data in

the creation process, but these data basically come from digitized literary works, therefore, the creation of generative AI is actually inseparable from the use of other people's works. And these copyrighted works being used involve not only public works that can be directly utilized, but also works protected by copyright law. Once a generative AI uses works that are still protected by copyright law without permission during the creation process, it may face the risk of copyright infringement. The process of AI creation is divided into three stages, and the different ways of using the works at different stages may produce different results, so it is not yet possible to identify all the AI creation and use of other people's works as copyright infringement, and it is also necessary to combine the different stages of the use of the works in a comprehensive judgment.

In the process of generative AI creation, the main work in the input stage is to digitize the literary works as creative materials, form recognizable data and store them in the database. In the process of digitizing and transforming creative materials into machine-readable "data", scanning, transcription and text extraction are often involved, and since these acts actually belong to the digital reproduction of works in China's Copyright Law, the generative AI may directly infringe the right of reproduction of works due to the above acts in the input stage of the creative process. Therefore, generative AI in the input stage of the creation process may directly infringe on the reproduction right of the work due to the above behaviors, thus facing the risk of reproduction right infringement [3]. In the creation process of generative artificial intelligence, the main work of the machine learning stage is to use various types of algorithms to analyze and train the data in the database, and to form a model on this basis for subsequent creation. In the artificial intelligence data analysis and training, this time will also involve the data "copy" behavior, but this time the copy is temporarily stored and briefly reproduce the work data, known as temporary copy [4]. At present, China's copyright law considers that temporary reproduction does not belong to the scope of the reproduction right, because it believes that once the reproduction right covers temporary reproduction, it may not only unreasonably expand the rights of the copyright owner, but also is not conducive to the access to information. Therefore, the temporary copying involved in the process of machine learning does not constitute an infringement of the copy right. In the process of generative artificial intelligence creation, the output stage mainly refers to the output of the corresponding content relying on the user's instructions after the machine learning of the original work. At this time, the output content may be an intercepted fragment of the original work, or it may be a combination of user instructions to modify and reprocess the content of multiple copies of the work. Such behavior of the generative artificial intelligence may infringe the copyright owner's right of modification, compilation and adaptation. In addition, if the generative AI does not attribute the output to the original author, it may be at risk of infringing on the copyright owner's right of attribution.

3. The Dilemma of Applying the Fair Use Regime under Copyright Law to Generative AI Creations

3.1 Impact on Author-centrism

The traditional fair use system is set up with the natural person as the center, which mainly responds to the reasonableness of the natural person's use of the work, but at present, artificial intelligence has begun to become the main body of reading and creation, and the traditional theory of "the work originates from the author" is beginning to shake, and the author's position as the creator is also being challenged by artificial intelligence [5]. On the other hand, copyright law used to pay more attention to the protection of the rights of copyright holders, but now for the public interest and the development of artificial intelligence technology, need to use a reasonable system to seek exemption from the risk of infringement of the works created and utilized by artificial intelligence. As a result, "author-centrism" has suffered to a certain extent.

3.2 Limitations on the Application of Fair Use Specific Provisions to Generative Artificial Intelligence Creations

3.2.1 Generative AI creations are not directly applicable to "personal study and research"

According to Article 24(1)(1) of China's Copyright Law, "the use of another person's published work for personal study, research or enjoyment" can be done without the permission of the copyright owner and without payment of compensation. First of all, from the perspective of the subject, the subject of "personal study and research" is limited to natural persons. Although artificial intelligence has the ability to analyze and solve problems, it is still dependent on natural persons. In addition, the subject and the object is the two basic systems of civil law, there are strict boundaries between the two, the subject of civil law is limited to natural persons and legal persons, if the artificial intelligence is directly formulated as a "human being", it may impact on the basic system of civil law [6]. Secondly, from the point of view of the purpose, individuals in the use of works, only for the purpose of "learning", "research", "enjoyment" and other non-commercial purposes, can be applied to the fair use system. In the case of AI creation, the purpose of utilizing the work is not purely for machine learning, but also for the purpose of creating valuable works with commercial value. Therefore, the copyright risks faced in the process of AI creation cannot be exempted from the provision as a result.

3.2.2 Generative AI creation cannot directly apply "proper citation"

According to Article 24(1)(2) of China's Copyright Law, "For the purpose of introducing or commenting on a work or explaining a certain issue, appropriate quotations may be made from the published works of others". First of all, AI creation is a secondary creative behavior based on learning existing works, and the purpose of quoting the contents of other people's works is to generate new works, so it can't satisfy the condition of the purpose of explaining a certain problem or introducing and commenting on a certain work [7]. In addition, AI creation cannot fulfill the "appropriateness" requirement mentioned in the Article. The requirement of "appropriateness" in the article means that the citation should be controlled within a certain limit, so as not to constitute substantial similarity or cause unfair competition, which will cause damage to the interests of the copyright owner. The utilization of other people's works in the process of creation of artificial intelligence may involve the reproduction and citation of the full text of the works, so it obviously cannot satisfy the degree of "appropriateness" of citation in this provision.

3.2.3 Generative AI creations not directly applicable to "scientific research"

According to Article 24(1)(6) of China's Copyright Law, "translating, adapting, compiling, broadcasting, or copying in small quantities published works for the purpose of scientific research may be used by scientific researchers, but may not be published or distributed." First of all, the establishment of the fair use clause for the purpose of scientific research is mainly based on public interest. Therefore, only the use of works in scientific research by non-commercial research institutions can be recognized as fair use, while the use for commercial purposes is not applicable to this clause. For artificial intelligence creation, it is backed by large Internet companies, which basically carry out the corresponding research on artificial intelligence creation for commercial purposes, and thus it does not fulfill the non-commercial purposes required in the clause. In addition, the AI may involve the reproduction of the full text of the work during the input stage of the creation process, and therefore does not satisfy the "small amount of reproduction" requirement of this clause.

3.2.4 Generative Artificial Intelligence Creation Not Directly Applicable to the "Underlying Clause"

In addition to the specific circumstances for fair use mentioned above, the Copyright Law also provides for "other circumstances stipulated by laws and administrative regulations" in Article 24(1)(13), which serves as an escape clause. This provision seems to leave room for exempting AI creations from the risk of copyright infringement to a certain extent, but in reality, since there are no laws or administrative regulations in China that provide for further regulations or interpretations, there are also difficulties in applying this provision to exempt AI creations. In addition, in the absence of other legal provisions or interpretations, once the AI creation of the direct application of the touting provision, will cause confusion in the application of the touting provision [8].

3.3 Uncertainty in Judicial Practice in

Determining the Application of Fair Use to Generative Artificial Intelligence Creations

3.3.1 Limitations of adopting the "three-step test" for creation

When the current copyright law was revised, the "three-step test" was introduced in order to be consistent with the Berne Convention. Unlike the Berne Convention's "three-step test", China's three-step test changes "under specific and special circumstances" to "in accordance with the relevant provisions of the Copyright Law", and "fair use shall not be inconsistent with the provisions of the Copyright Law" to "in accordance with the relevant provisions of the Copyright Law", whereas "fair use shall not be inconsistent with the provisions of the Copyright Law" is not applicable to "fair use". China's three-step test changes "under specific and special circumstances" to "in accordance with the relevant provisions of the Copyright Law", while there is no change in "reasonable use shall not conflict with the normal utilization of the work" and "shall not impair the lawful rights and interests of the copyright owner". Therefore, in practice, the "three-step test" still requires the judge to decide whether the unauthorized use of a work constitutes fair use within the specific circumstances listed in the Copyright Law. Therefore, in order for the AI to be exempted from the three-step test to determine whether the use of a work is fair use, it must meet the requirements of fair use as stipulated in the law. In this case, the use of works by AI can only be categorized into the "other circumstances" of Article 13, but due to the ambiguous wording of this provision, there is uncertainty as to whether the court can make a fair use judgment according to its own understanding of the emerging new circumstances in the judicial practice.

3.3.2 Limitations of adopting the "four elements" to create new circumstances

In judicial practice, in addition to trying to utilize the "three-step test" to determine the new circumstances of fair use, the court also tries to create new circumstances of fair use through the "four elements" under the transformative theory, in the hope of responding to the new difficulties [9]. For example, in the case of copyright dispute between Shanghai Fine Arts Film Studio and Zhejiang New Film Era Culture Communication Co., Ltd, Shanghai Fine Arts Film Studio considered that Zhejiang New Film Era Culture Communication Co., Ltd. used part of the company's art works in the publicity posters without permission, which was an infringement of copyright. However, the company argued that the "Huluwa" and "Black Cat Sheriff" used in the posters were different from the works of art of the Shanghai Fine Arts Film Studio, and it considered that the proportion of quotations belonged to a small part, which should be regarded as fair use. In the trial, the court of first instance adopted the "four elements" judgment standard, which was based on the nature of the original work, the purpose of use of the citation, the proportion of the cited portion, and whether the citation would be given to the Shanghai Fine Arts Film Studio on the "Sergeant of the Black Cat", "Huluwa" art works from four perspectives, under comprehensive consideration, it was considered to constitute fair use. It can be seen that under the "four elements" theory, the court can expand the scope of fair use of copyright beyond the statutory circumstances. However, since China is not a case law country, even if the "four elements" have created a new fair use in individual cases, it cannot be taken as a general phenomenon, thus providing a reference for the fair use of AI creative acts.

4. Extraterritorial Provisions on the Application of Fair Use Regimes to Artificial Intelligence Creations

4.1 The United States: The Legal Application of the Transformative Theory

The United States passed legislation in 1976 to establish the four elements of fair use developed in the judicial process [10]. It believes that in order to constitute fair use, the following four elements must be considered, namely, the purpose as well as the character of the act of use, the nature of the work used, the amount of the portion of the work used and the effect of the act of use on the potential market for the original work. However, since the U.S. copyright law does not specify the boundaries of the application of the four elements, it has led to rigidity in practice. The turning point came in 1990, when Judge Pierre proposed the transformative theory for the first element, "the purpose and character of the use". He believed that "the use must be productive and

use another's work in a manner or for a purpose different from that of the original", and emphasized that this is the most important element [11]. And then, the theory was accepted by the U.S. Federal Supreme Court. Under the guidance of the theory, the U.S. courts have made judgments determining that the act of copying works in the field of artificial intelligence constitutes fair use. For example, in the case of Writers Guild of America v. Google, Google scanned a large number of books without the author's permission and provided the public with searches of these digitized works as well as small snippets of content [12]. The judge of the Court of Second Instance held that "Google's copying of a work for the purpose of criticizing, commenting on, or providing information not found in the original work makes it abundantly clear that it is in fact satisfying the transformative purpose required in the first of the four elements. Therefore, in this case, we hold that Google's profit-making purpose does not negate its highly persuasive transformative purpose, and thus the overall profit-making purpose will not be a reason to exclude the application of fair use.

4.2 Japan: Creating a Computer Analysis Exception

Japan was the first country to provide a fair use pathway for the exploitation of works in the creative process of AI. Japan believes that for the purpose of computer information analysis, a computer can be used to conduct audiovisual linguistic analysis of a work or information within the necessary limits, and if adaptations and storage are made on this basis, then the behavior is fair use and does not constitute infringement. In addition, in order to meet the new round of scientific and technological revolution and build a copyright system in line with the development of artificial intelligence technology, Japan with 2018 again revised the Copyright Law. In this revision, Japan, after borrowing the theory of transformative use from the United States, instead of limiting the definition of computer analysis to statistical analysis, it relatively expanded the scope, and Japan extended the scope to provide new information or knowledge for the purpose of providing new information or knowledge. It also provides that an information processor may make necessary copies of a work and make them available to the public without prejudice to the interests of the right holder.

4.3 EU: Adding Text Mining Exception Rules

The EU adopted the Single Digital Market Copyright Directive in 2019 and provided for copyright exceptions for text and data mining in Articles 3 and 4. It provides in Article 3 that only text and data mining for the purpose of scientific research falls within the statutory exception to copyright, and that any deviation from this provision by way of terms of service or license agreements is invalid. Unlike Section 3, Section 4 provides a limited exception for text and data mining. While Article 4 stipulates that users are allowed to reproduce and extract legally obtained works in the course of text and data mining, it also provides an exception that copyright holders may retain the right to reproduce and extract for the purpose of text and data mining. In other words, if the user wants to use text and data mining for commercial purposes, he/she must respect the will of the copyright holder, otherwise it may constitute infringement. The Copyright Directive issued by the EU, with its new addition of text and data mining, not only allows the use for scientific research purposes, but also takes into account the interests of copyright holders in addition to the use for commercial purposes, which to a certain extent has contributed to the development of artificial intelligence technology.

5. Suggestions for Improving the Fair Use Regime for Generative Artificial Intelligence Creations

5.1 Adding "Computer Analysis Exception" to the Fair Use Scenarios

The United States, Japan and the European Union have all responded to the question of whether the use of other people's works in the process of generative AI creation can be used to defend against infringement through the fair use system. Currently, China's existing copyright law does not provide applicable provisions for the use of works created by generative AI, therefore, in order to cope with the challenges posed by the development of AI technology, China's copyright law needs to take corresponding measures. Japan and the

European Union both chose to add a fair use scenario to their legislation, namely "computer analysis exception" and "text data mining", respectively. The United States, as a case law country, mainly utilizes the transformative use theory to determine whether AI creations constitute fair use in judicial practice. The United States, as a case law country, mainly utilizes the transformative use theory to determine whether the creation of artificial intelligence constitutes fair use in judicial practice [13]. Since China is not a case law country, it can choose to follow the practice of the EU and Japan, through legislation, in addition to the existing legal circumstances of fair use to add a special field of generative AI creation of fair use, that is, "text data mining" or "computer analysis exceptions". ".

5.2 Improvement of AI-related Legislation

In addition to considering the addition of new fair use circumstances, the situation of generative AI creation and utilization of works can actually be included in the underlining clause of the fair use circumstances stipulated in the Copyright Law, i.e., "other circumstances stipulated by laws and administrative regulations". As this provision is semi-open-ended, its application should be based on other laws and regulations. Therefore, in the event that other laws or administrative regulations do not provide for it, it may not be applicable. Currently, although China has adopted the Interim Measures for the Administration of Generative Artificial Intelligence Services, the document has not yet mentioned whether the utilization of works by generative AI in the process of creation belongs to the category of fair use, and therefore AI creations are not yet able to apply the underpinning clause. Therefore, in order to be able to apply the underpinning clause, it is necessary to introduce laws and regulations related to the creation of generative artificial intelligence as soon as possible, so as to provide a corresponding legal basis for the application of the underpinning clause to the creation of artificial intelligence, so as to avoid the corresponding risk of copyright infringement.

5.3 Clarify the Determination of the Reasonable Use of the Three-step Test in Generative AI Creation

The "three-step test" consists of three elements, and there is a logical relationship between the three elements, which should be carried out in a progressive manner. The "certain specific circumstances" refers to the 12 specific circumstances stipulated in the current copyright law and the bottom clause. Since the creation and utilization of works by generative AI does not belong to the specific circumstances of fair use stipulated in the law, it needs to rely on the court's previous judicial experience to determine whether it is fair use according to "whether it belongs to the special circumstances that are really necessary to promote technological innovation and commercial development". The second element is that "it must not conflict with the normal utilization of the work". First of all, "normal utilization" refers to the use of the normative sense, once the economic competition with the author, will constitute "conflict with the normal utilization of the work". At present, judicial practice lacks specific criteria for determining whether the normal use of the original work is jeopardized, which makes it very difficult to make a judgment. Therefore, when judging whether the acquisition and use of the work in the process of AI creation can constitute fair use, it is necessary to further combine the content of the output phase with whether the existing and potential market of the original work is alternative use, and also combine the expectation of the copyright owner and whether the real economic interests are harmed to make a specific judgment [14]. "The third step of the three-step test is that "no legitimate interest of the author shall be unreasonably prejudiced". This "legitimate interest" covers "legitimate economic interests" as well as "some non-economic interests such as personality interests". "Unreasonableness" is a judgment about the extent of use of a work or the division of economic interests, and is intended to bring within the scope of fair use relatively small economic harms within a given range of interests. This standard permits the use of the copyright of the copyright holder within a certain range, which restricts the scope of the use of the work by the user, and also restricts the rights of the copyright holder to a certain extent, in order to achieve a balance between the reasonable use of the work by the public and the protection of the

corresponding rights of the copyright holder. Therefore, when judging whether the AI creation can meet the reasonable use when utilizing the work, it can observe whether its use method adopts the method that minimizes the damage to the author's rights, and it can also be considered in combination with the nature of the work to be used, the proportion of the original work's expressive value being occupied, and other factors [15]. At the same time, it can also be considered whether the social public interest brought by the creation of AI outweighs the impact on the market of the original work.

6. Conclusion

The development of AI technology has brought convenience to society to a certain extent, but at the same time, it has also brought challenges to the current copyright law in China. As the current fair use system of the current copyright law is not yet able to respond to the legitimacy of the works utilized by AI, it leads to the risk of copyright infringement when it is created by AI. As a cutting-edge scientific technology of common concern to countries around the world, AI has been responded to by the United States, Japan and the European Union with new measures. Therefore, in order to solve the dilemma of applying the fair use system to China's AI creations as soon as possible, it is necessary to add a "computer information analysis" clause to the specific statutory circumstances of the fair use system, formulate AI-related laws, and clarify the application of the three-step test for fair use of AI creations in judicial practice, so as to include the use of works created by AI into the fair use system. Artificial intelligence creation and utilization of works into the scope of the fair use system. This will not only clear the obstacles of infringement risk for AI creation and promote the accelerated development of AI technology, but also protect the legitimate interests of copyright holders and further promote cultural prosperity.

References

[1] Jiao Heping. Copyright Risks of Data Acquisition and Utilization in Artificial Intelligence Creation and the Path to Resolve Them. Contemporary Law. 2022(04): 128-140

[2] Wu Handong. The Question of Copyright Law of Artificial Intelligence Generated Works. Chinese and foreign jurisprudence. 2020(03): 653-673

[3] Zhang Jinping. Fair Use Dilemma of Artificial Intelligence Works and Its Solution. Global Legal Review. 2019(03): 120-132

[4] Dai Xin. Selection of Copyright Infringement Exception System for Artificial Intelligence Creation. Journal of Hubei Institute of Science and Technology. 2022 (05): 38-46

[5] Lin Xiuqin. The reshaping of copyright fair use system in the era of artificial intelligence. Legal Research. 2021(06): 170-185

[6] Wu Handong. Institutional arrangement and legal regulation in the era of artificial intelligence. Legal Science (Journal of Northwest University of Politics and Law).2017(05):128-136

[7] Gao Yang, Hu Danyang. Challenges and responses of machine learning to copyright fair use system. Electronic Intellectual Property Rights. 2020(10): 13-25

[8] Wang Wenmin. Challenges and Responses of Artificial Intelligence to Copyright Limitation and Exception Rules. Law Application. 2022(11): 152-162

[9] Wan Yong. Dilemma and way out of fair use system of copyright law in the era of artificial intelligence. Social Science Series. 2021(05): 93-102

[10] Benjamin L.W. Sobel, Artificial Intelligence's Fair Use Crisis, 41 Colum. J.L. & ARTS 45 (2017)

[11] Hua Jie. The Dilemma and Way Out of the Fair Use System Applied to Artificial Intelligence Creation. Electronic Intellectual Property Rights. 2019(04): 29-39

[12] Stephen McJohn & Ian McJohn, Fair Use and Machine Learning, 12 N.E. U. L.R. 99 (2020).

[13] Xuan Zhe. On Copyright Fair Use of Artificial Intelligence Creation under the Perspective of Classification Protection. Publishing and Distribution Research. 2022(03): 81-87

[14] Xiong Qi. Interpretation of judicial determination standard of copyright fair use. Law. 2018(01): 182-192

[15] Liu Youhua, Wei Yuanshan. Copyright infringement problem of machine learning and its solution. Journal of East China University of Politics and Law. 2019(02): 68-79.

