

# Legal Analysis of Copyright of Artificial Intelligence Products in the Era of Weak Artificial Intelligence

JINGZHUO GENG

## ABSTRACT

At present, although artificial intelligence technology is still in the stage of weak artificial intelligence, it depends on the development of digital computing, which breaks the monopoly position of human beings in the field of literary creation, and then the legal subject and behavior of copyright have intelligent tendency, which puts forward new problems in judicial practice. Through the analysis of the nature of AI-generated works and the distinction between different categories, this paper makes a concrete analysis of the copyright of AI-generated works, and then draws the conclusion that the copyright of AI-generated works belongs to people. By using the criterion of "contact + substantial similarity", it is concluded that the self-editing articles produced by artificial intelligence based on deep learning do not belong to infringement. The expression and utilization of manuscript washing in artificial intelligence products constitutes infringement and does not belong to fair use. Whether the use of the original material by the ideological use of manuscript washing constitutes infringement requires a case judgment.

## INTRODUCTION

With the development of supercomputing, Internet technology and big data, artificial intelligence technology represented by deep neural networks has also made great progress. As the core force of the new round of scientific and technological revolution and industrial transformation, artificial intelligence has shown great advantages in the material fields of education, medical treatment, driving and so on. With the maturity of algorithm technology, AI has also been applied to the spiritual fields of literary creation, chess game. The progress of artificial intelligence technology has greatly enriched the material and spiritual life of human beings, but it has also made the elements of legal subjects and behaviors intelligent, which has a certain impact on the current law.

AI machines can be divided into weak AI machines and strong AI machines according to whether AI has independent consciousness and will, and whether it can make autonomous decisions and implement corresponding behaviors outside the scope of designed and programmed programs. The former can only make autonomous decisions within the scope of designed and programmed programs through deep

---

Jingzhuo Geng

School of Humanities, Social Sciences and Law, Harbin Institute of Technology, Harbin, China

learning, while the latter has independent will, and the behavior may exceed the scope of designed and programmed programs. [1] At present, the artificial intelligence technology in China and the world is still in the stage of weak artificial intelligence, limited to the scope of programming, and there is no strong artificial intelligence machine beyond the scope of programming. Therefore, this paper will analyze the related issues of copyright enjoyment and tort liability of artificial intelligence products from the perspective of weak artificial intelligence.

## **COPYRIGHT ABILITY ANALYSIS OF ARTIFICIAL INTELLIGENCE PRODUCTS**

The emergence of artificial intelligence products has caused the legal subject and object to have an intelligent tendency, subverting the legal elements of the current copyright law to a certain extent, and triggering a huge controversy in the academic community about its copyright ability. The focus of the dispute is mainly on the two major issues of subject and object.

### **The Nature of Artificial Intelligence Products**

From the perspective of the object, Article 3 of China's Copyright Law stipulates that "works refer to intellectual achievements that are original in the fields of literature, art and science and can be reproduced in some tangible form". For the identification of the originality of artificial intelligence products, the academic community has produced two different standards of "subjective and objective". The subjective standard is centered on the author, and believes that originality should reflect the author's personality, personality and spirit. [2] The objective standard is centered on the work, focusing on whether the work has a "minimum creativity" or has a "significant difference" with the existing work. [3] In China's judicial practice, the objective standard is generally used as the basic criterion for originality identification, that is, regardless of the degree of subjective aspects of the author's personality and spirit in the work, only the "minimum creativity" or the "significant difference" with the existing work can be identified as a work in the sense of copyright law.

Stuart Russell and Peter Norvig put forward eight different definitions of artificial intelligence in the book *Artificial Intelligence: A Modern Approach*, which are divided into four categories: humanized action, humanized thinking, rational thinking and rational action. [4] Although weak artificial intelligence is limited to the stage of "humanized action and humanized thinking", it simulates people's thoughts and consciousness to varying degrees according to different algorithms and program settings, and its creative ability is also different.

The nature of artificial intelligence based on self-creation generated by deep learning.

Microsoft Xiao Bing released by Microsoft is a representative of creative ability based on deep learning. In order to achieve the skills of poetry creation, after tens of thousands of trainings, Xiao Bing "learned" the modern poetry of 519 poets since the

1820s. From the beginning of the verse written by Xiao Bing is not smooth, to now has formed a "unique style, preferences and writing skills".

The generation of such artificial intelligence creative behavior is roughly divided into two stages: the first stage is the deep learning stage. Artificial intelligence generates humanized thinking on the input original material through the algorithm, summarizes the source code in the code base as the original material, and establishes the language model through massive learning and training, so as to obtain creative skills. The second stage is the output stage. According to the instructions issued by the user, the action is generated, the code is automatically generated, and the independent creation is carried out. The creative behavior is through tens of thousands of learning and training to produce an independent style and expression, with significant originality. According to the three constituent elements of "works limited to the field of literature, art and science, with originality and can be reproduced in some tangible form" stipulated in Article 3 of the Copyright Law, it can be seen that the independent creation of artificial intelligence based on deep learning conforms to the works in the Copyright Law and is protected.

The nature of artificial intelligence's directional adaptation of original materials

The "adaptation" behavior of such artificial intelligence occurs on the basis of completed learning and acquired learning ability. The original material is published to artificial intelligence as an object of directional adaptation without presetting synonyms. Through article retrieval, artificial intelligence extracts ideas and directional adaptations of articles based on its own learning ability, and finally outputs articles that are partially or completely different from the original material but highly relevant. Artificial intelligence's directional adaptation of original materials is also known as artificial intelligence washing behavior, that is, using artificial intelligence as a tool to integrate new articles by tampering, deleting, and piecing together other people's articles. According to whether it is substantially similar to the original material, the artificial intelligence manuscript can be subdivided into "thought utilization type manuscript" and "expression utilization type manuscript".

"Ideological utilization type manuscript washing" is an artificial intelligence that extracts and summarizes the high repetition rate vocabulary, key content, and core viewpoints in one or more articles with similar ideas through algorithm technology. On the basis of finding out its internal logic, it retains the core ideas of the original works, uses different article forms and genres, and constructs new articles with different expressions to achieve the effect of "secondary creation". At present, both our legal provisions and international conventions reflect the principle of "protecting the form of expression, not the thought itself" [5]. Therefore, this kind of washing product can be identified as a work in the sense of copyright law and protected only if it meets the minimum creative standard. "Expression utilization type manuscript washing" is the behavior of artificial intelligence to form an article similar to the original material thought and expression by searching one or more original materials, synonym transformation, word order change, paragraph splicing and so on. The formation of this article is only through the deletion and transformation of the original article, which does not constitute a "significant difference" with the original work in expression. In fact, it is the product of complex plagiarism and cannot reflect the originality of the article.

Therefore, the product of "expression and utilization of manuscript washing" is not the work stipulated in the copyright law.

### **The Copyright Ownership of Artificial Intelligence Products**

From the perspective of the subject, it is controversial whether artificial intelligence can enjoy rights and obligations as the author stipulated in the copyright law. The second paragraph of Article 10 of the Copyright Law limits the subject of copyright to natural persons, legal persons and other organizations, and it is generally said that works are the creations of human beings and express human thoughts and feelings. [6] The results produced by human intellectual activities can be called works. At present, weak artificial intelligence does not meet the requirements of the subject and does not have the qualification of the subject. In order to solve this practical problem, some scholars advocate that the copyright protection of artificial intelligence products can refer to the provisions of duty works or employment works. [7] Give rights other than signature rights to users of artificial intelligence. Some scholars advocate that "artificial intelligence products can learn from the institutional arrangements of legal person works, which can be regarded as creative achievements representing the will of AI designers or trainers, and the copyright belongs to the owner of artificial intelligence." [8] However, due to the data attributes of artificial intelligence, the above views are slightly one-sided. With the development of science and technology, the academic community has seen the view that artificial intelligence is a tool for human creation, [9] which perfectly solves the problem that artificial intelligence does not have human identity. From an objective point of view, artificial intelligence and its products are based on human algorithms. In the final analysis, they are the embodiment of human intellectual achievements. When they meet the criteria of "minimum creativity" or "significant difference", they can be identified as works. On this basis, it analyzes the subject who enjoys the copyright of the work: when artificial intelligence is used as a tool to produce the work, the owner or user of artificial intelligence, as the original "creator" of artificial intelligence and its products, should enjoy copyright; however, the great contribution of artificial intelligence itself, which learns to acquire creative ability, to the creation of works should not be ignored. The work can be marked as an AI work when it is signed, so that artificial intelligence has its own name and can also be marked by name. This is not only an affirmation of artificial intelligence products as works, which is conducive to the dissemination and utilization of artificial intelligence products, but also the protection and maintenance of the public's right to know. Therefore, in the case that artificial intelligence products constitute works, that is, "artificial intelligence based on deep learning to generate independent creations" and "thought-using manuscript washing" mentioned in the text, the above criteria can be used as the basis for the division of copyright ownership: if the user of the work and the owner of the artificial intelligence have an agreement on the ownership of the generated work before the work is generated, the agreement is followed ; if there is no agreement, the user of the work shall enjoy the copyright, and the artificial intelligence machine may enjoy the right to indicate the identity on the work ; if there is no agreement and the user of the work is unknown or the user explicitly gives up the

copyright, the owner or investor of the artificial intelligence can enjoy the copyright, and the artificial intelligence machine also enjoys the right to indicate the identity on the work. However, "expression and utilization of manuscript washing" has no copyright ownership because it cannot constitute a work itself and is suspected of infringement in the following details.

## **ANALYSIS OF TORT LIABILITY OF ARTIFICIAL INTELLIGENCE PRODUCTS**

In the judicial practice of our country, the principle of "contact + substantial similarity" is generally used to judge whether it constitutes copyright infringement. With the continuous development of science and technology, the forms of "contact" are becoming more and more diversified, and the proof of contact is becoming more and more difficult. Therefore, the identification of "substantial similarity" has become the key to judging cases. Professor Wu Handong points out that "this rule is sometimes simply called "substantial similarity regulation" and" [10] in judicial activities, and it is more difficult to specifically identify" it is arbitrary anywhere" [11] substantial similarity. At present, the more common methods in China's judicial practice are "holistic perception method" [12] and "abstract separation method"[13]. The overall perception method is to compare the two works from the perspective of ordinary people, focusing on readers feelings about the works. This method is more intuitive and rapid, which helps to improve judicial efficiency, but the shortcomings are also very obvious. Using this method to judge whether it constitutes "substantial similarity" without screening and distinguishing the content of the work, it is easy to incorporate the ideological part that is not protected by copyright into the comparison link, thereby expanding the scope of protection of copyright law and weakening the author's creative enthusiasm. The three-step test method is to separate the parts of the work that are not protected by copyright law through the three steps of " abstraction-filtering-comparison", focusing on the similarity of expression, which is limited to a certain extent compared with the protection scope of the overall perception method. However, due to the fact that what is thought and what is expression is not absolute, the result is often based on the discretion of the judge, which is easy to produce different judgments in the same case. For the identification of substantive similarity, specific case analysis is often required.

### **Determination of Infringement**

The infringement of artificial intelligence products can be roughly classified into two categories according to the nature of the products. One is that the products meet the original characteristics and can be positioned as works in the sense of copyright law, which infringes the legitimate rights and interests of other works; the other is that the product is substantially similar to the form of the original work in the process of production, which constitutes an infringement of the original work. This type of infringement does not require the product to constitute a work. The former corresponds

to the two artificial intelligence products of "artificial intelligence based on deep learning" and "thought-using manuscript washing", which constitute the works introduced in this paper. This kind of infringement form is relatively common. It is only necessary to prove the infringement facts according to the principle of "contact + substantial similarity", clarify the ownership of copyright and according to the principle of fairness - the obligee is the obligor, and the conclusion that the copyright owner should bear the tort liability can be drawn. The latter's tort liability attribution is more complicated, which needs to be analyzed in depth and combined with specific issues.

For artificial intelligence based on deep learning, according to the principle of "contact + substantive similarity", although this kind of artificial intelligence has contact with the original material expressed in the form of source code, the original material is only the material for artificial intelligence to carry out deep learning and obtain creative skills and does not aim at creating the same content as the original material. The creative behavior is through tens of thousands of learning and training to produce an independent style and expression, with originality, so the content of independent creation is generally not substantially similar to the original. Therefore, the self-editing articles produced by artificial intelligence based on deep learning do not infringe the rights of the original material owners and do not constitute infringement.

There are practical difficulties in the determination of whether it constitutes infringement for the products of "thought-using manuscript washing". The essence of "thought-using manuscript washing" is to use artificial intelligence to generate new articles with different expressions on the basis of retaining the original works, so as to achieve the purpose of secondary creation. In judicial practice, judging such infringing works, we first exclude the ideological and creative parts of the works, and pay attention to the similarity of the expression level, which confirms the "thought and expression dichotomy theory" that the copyright law only protects the expression without protecting the thought. Therefore, it is difficult to make a general determination on whether the ideological use of the manuscript washing product and the original material constitute infringement, which should be analyzed in combination with the specific situation.

The product of "expression and utilization" is an article similar to the original material expression formed by artificial intelligence through the retrieval of one or more original materials, only by simple synonym transformation, word order change, paragraph splicing and so on. Further analysis according to the principle of "contact + substantial similarity": when the original article is provided to artificial intelligence as the original material, the contact is completed; the formation of the new article is only through the deletion and transformation of the original article, which cannot reflect the originality of the article in expression. It can be concluded that the article is substantially similar to the original article. Therefore, the product of "expression and utilization type manuscript washing" violates the right of reproduction of the original article and should bear tort liability.

## Reasonable Use

Fair use is the right of the public to freely use copyright works, and it is also the limitation of the law on the effectiveness of copyright property rights." [14] In judicial practice, China has stipulated the "three-step test method" in Article 9, paragraph 2, of the Berne Convention for the Protection of Literary and Artistic Works. At the same time, in the specific interpretation process of the three-step test method, especially the first step-the interpretation of the specific circumstances stipulated by the law, China's judicial practice community began to learn from the "transformative use" judgment rules of the United States to judge whether it constitutes fair use. In the 1990 s, Judge Leval deepened the element of purpose and characteristics of use into transformative use from the standpoint of copyright law encouraging intellectual creation. The main points are as follows : (1) Whether the use meets the "goal of promoting knowledge and encouraging creation " of copyright law is the core issue of examining the purpose of use ; (2) If it is used in a different way from the work, or for a different purpose from the original work, it is "transformative use"; (3) "Transformative use" makes the new work different from the original work, so it is creative and belongs to fair use. [15] We can also analyze the relationship between artificial intelligence products and original materials according to this standard.

According to the "three-step test method", this paper analyzes the "expression and utilization type manuscript washing" product which initially constitutes infringement mentioned above: First, fair use can only be used under certain special conditions. However, the extensive use of original materials by artificial intelligence to generate new articles is not in line with the scope of fair use in the first paragraph of Article 24 of China's Copyright Law and the purpose of "transformative use" to promote knowledge and encourage creation; secondly, fair use shall not conflict with the normal interests of the work. The purpose of most of the "expression and utilization type manuscript washing" products is to complete the manuscript washing of the original article at a very low cost to achieve the purpose of seizing the market, which has a great conflict with the normal interests of the work. Again, shall not damage the legitimate rights and interests of the copyright owner. The product of "expression and utilization" will erase the expression of the original article as much as possible, and try to cover up the essence of plagiarism, not to mention the damage to the author's signature right and the name right of the original article. Moreover, "expression and utilization type manuscript washing" is actually a way of using artificial intelligence to plagiarize. There is no transformation in the form or thought of the work. Therefore, the expression and utilization type manuscript washing does not meet the fair use standard and should constitute the infringement of the original material and bear the tort liability.

## Shoulder Responsibility

For the tort liability of "expression and utilization of manuscript washing", the artificial intelligence product can be regarded as a special job work, that is, the users who use artificial intelligence to wash manuscripts constitute direct infringement and bear no-

fault liability. At the same time, the auxiliary position of artificial intelligence algorithm designers in infringement should not be ignored. Its original intention of designing this algorithm is to facilitate the washing behavior through the algorithm and to make profits. Knowing that the algorithm will lead to the occurrence of infringement, still adhere to the development and application of this algorithm, with a strong subjective intention of infringement, it can be determined that there is "indirect infringement" and should also be responsible for the consequences of infringement.

## PLATFORM RESPONSIBILITY

As a variety of network platforms that build a bridge between the news media and readers, it should be responsible for supervision and review. However, as to whether the platform should be responsible for the infringement of the media, I believe that the principle of safe haven should be applied. According to the provisions of Articles 1195 and 1196 of the Civil Code of China and Article 23 of the Regulations on the Right to Network Dissemination of Information of China, "Internet service providers provide search or link services for service objects. After receiving a notification from the right holder, if the link with the infringing works, performances, audio and video recordings is disconnected according to the provisions of this Regulation, it shall not be liable for compensation; however, those who know or should know the infringement of the linked works, performances, audio and video recordings should bear the joint tort liability." It can be seen that when the infringement occurs, the network platform should timely and accurately fulfill the obligation of notification and take necessary measures to prevent further damage to the right holder. In this case, the platform cannot bear the joint liability of tort liability. However, the platform should not relax the review of the content of the dissemination. If the infringement is obvious and the platform has not taken any measures, the platform cannot use the safe harbor principle as a defense to shirk responsibility.

## REFERENCES

1. Liu Xianquan: The View of Criminal Law in the Era of Artificial Intelligence, Shanghai People's Publishing House, 2019 Edition, Page.
2. Zongye. Research on copyright infringement of artificial intelligence products [D]. Shandong University, 2021.DOI: 10.27272/d.cnki.gshdu.2021.006270.
3. Lu Binghong. On the selection of criteria for judging the originality of artificial intelligence creations [J]. Inner Mongolia Science 2021. DOI: 10.27272/d.cnki.gshdu.2021.006270.
4. Russel S.J. Norvig P. Artificial Intelligence: A Modern Approach (Fourth Edition) [M]. London: Pearson, 2021:20-22.
5. Guide to the Berne Convention for the Protection of Literary and Artistic Works (1971 Paris Text).
6. Lu Haijun: "Works in the sense of copyright law"—taking artificial intelligence products as the starting point, Seeking 2019, No. 6, p. 75.
7. Wu Handong: Institutional Arrangement and Legal Regulation in the Era of Artificial Intelligence, Legal Science (Journal of Northwest University of Political Science and Law) No. 5, 2017, p. 131.

8. Lu Binghong. On the copyright protection of artificial intelligence products [D]. Jilin University, 2021.DOI: 10.27162/d.cnki.gjlin.2021.007564.
9. Artificial Intelligence, Copyright, and Copyright Infringement. Zack Naqvi. Marquette Intellectual Property Law Review. 2020.
10. Wu Han Dong. On the "substantial similarity + contact" infringement identification rules [J]. Law, 2015(08):63-72.
11. Liang Zhiwen. Judgment of substantial similarity in copyright law [J]. Jurist, 2015(06):37-50+174.DOI: 10.16094/j.cnki.1005-0221.2015.06.004.
12. Pamela Hobbst. Methods of Determining Substantial Similarity in Copyright Cases Involving Computer Programs, [12] University of Detroit Law Review No. 67, 1990.
13. Computer Associates International, Inc. v. Altai Inc, 982 F. 2d 693, (2nd Cir.1992).
14. Wu Han Dong. Questions on Copyright Law of Works Generated by Artificial Intelligence [J]., 2020, 32(03):653-673.
15. Pierre N. Leval, Toward a Fair Use Standard, Harvard law review, Vol. 103, No. 5, 1990, pp. 1111-1112.