

Research on the Nature of Works of Artificial Intelligence-generated Content

Ye Lin

Law School, People's Public Security University of China, Beijing, China

Abstract: *Under the background that the writing boundary of "human-machine dichotomy" has become blurred, it is necessary to discuss the nature of the original content generated by artificial intelligence (AIGC). From the perspective of legal interpretation, the fruits system, the data product system and the traditional work system cannot be reasonably applied to AIGC. From the perspective of legislative theory, we should abandon the concept of "romantic author" and solve the problem of the nature of content generated by artificial intelligence, which is subject to the writing principles of human authors. The system of works is no longer applied in an all-or-nothing way, and takes originality as the key criterion for identifying works, and puts forward a new form of works.*

Keywords: *Human-machine dichotomy, AIGC, Romantic author, Originality*

1. Introduction

The so-called artificial intelligence generated content (AIGC) is the extraction, analysis and integration of massive works, information and materials under the support of a specific algorithm program, so as to produce products that conform to the characteristics of the work in form. ChatGPT, a natural language machine with Transformer neural network architecture as the core algorithm released by OpenAI in 2022, trains the model by connecting a large number of corpora, and then communicates, writes scripts, emails, copywriting and code. Technically speaking, the definition of the concept of artificial intelligence has not reached a general consensus in the natural science community, but the essence of its technical principle is "intake of data - data training - establishment of induction model - output data model" with the support of preset algorithms, with initial autonomous learning ability. The generated content generated by the artificial intelligence is identical in appearance to the human works, which also causes the qualitative problem of the generated content in the sense of copyright law [1-3].

As for the qualitative content generated by artificial intelligence, there have been many discussions at home and abroad. Scholars outside the region are more inclined not to assign the status of "works" to the content generated by artificial intelligence at the level of legislation and interpretation, while Chinese scholars hold a more open position on the copyright of the content generated by artificial intelligence and affirm the status of "works" of the content generated by artificial intelligence. From the perspective of legal interpretation, it is impossible to properly deal with the qualitative problem of artificial intelligence-generated content, but it is necessary to identify artificial intelligence-generated content as "works" with copyright incentive mechanism. Therefore, it should be made clear from the perspective of legislation that artificial intelligence-generated content with autonomous characteristics is the research object in the sense of copyright law. The recognition system of "works" in the existing copyright law is modularized and applied to create a special form of works.

2. AIGC is recognized as a necessity of the work

2.1 Requirements of copyright incentive mechanism

The recognition of artificial intelligence-generated content as "works" meets the requirements of the copyright incentive mechanism. According to Article 1 of the Copyright Law of the People's Republic of China (hereinafter referred to as the Copyright Law), the Copyright Law protects copyright-related rights and interests and encourages the creation and dissemination of works is the method, and promotes the development and prosperity of cultural undertakings is the purpose. Most of the content generated by ChatGPT artificial intelligence exists in the form of data. Some scholars have proposed that the rights and interests of data products are the whole data as a kind of intangible property. If it is original, it can

be included in the protection scope of copyright law, but the incentive issue of people who play an important role in the development of artificial intelligence should be taken into account. The practice of identifying the content generated by artificial intelligence as "works" is in line with the core intention of the copyright law and has great significance: first, it promotes the creation of high-quality works, thus promoting the prosperity of the cultural industry and meeting the public's demand for knowledge and information; The second is to classify the content generated by artificial intelligence as "works", which is a necessary process for the allocation of rights. According to the principle of externality, the externalities, especially the negative impacts, generated by individual behavior activities are transformed inward by transforming them into private costs. As a rational person, in order to pursue the maximization of benefits, private individuals will try their best to reduce private costs. Law, as one of the constraints, enables them to optimize their personal benefits to optimize their social benefits. Only in this way can they have the right allocation and facilitate the play of incentive functions. Identifying AI-generated content as "works" also meets the requirements of the cheapest form of protection, which has its path dependence[4-7].

2.2 Blurred "human-machine dichotomy" boundaries

When it comes to writing, the "human-machine dichotomy" is no longer clear. From the perspective of the development law and trend of the human industrial revolution, the essence of artificial intelligence, at the beginning of the first industrial revolution, James Hargreaves invented the "Jenny machine" with the skills and production factors of human textile workers in the field of textile, although the machine is an extension of human initiative, but is it not the machine in the textile simulation of human weaving, making the production efficiency significantly improved. This has led to a leap in productivity. In this respect, artificial intelligence also has the similarity of its historical development destination, that is, simulating human beings to carry out a series of thinking activities and decision-making judgments, such as problem solving, reasoning planning and analysis, independent thinking and decision-making, deep autonomous learning and natural language expression. This type of artificial intelligence generates the following characteristics:

The generated results become unpredictable. The early computer program, when the operator gives repeated and identical instructions, can directly produce results, which can be predicted by the operator, while the current artificial intelligence in the existing database of preset algorithm intake and mining, the generated content of the output is unpredictable by the operator, because when the operator gives the artificial intelligence an abstract instruction, Artificial intelligence will choose specific goals from abstract goals to generate purpose content. When human authors create, others cannot predict the content of their creation, and the two are extremely similar in terms of final results.

The program operates with a high degree of independence. Artificial intelligence such as Dreamwriter is not yet out of the category of "tools", because the recognition of artificial intelligence generated content in the case of human creative contributions in the use of artificial intelligence in the creation process does not contribute to originality. "And the algorithms that artificial intelligence relies on are capable of operating independently, completely different from the tool-like devices mentioned earlier. As long as you give artificial intelligence an instruction, it can find correlations in massive amounts of data through associative thinking and output content, demonstrating a high level of intelligence."

Algorithmic learning results are shape-able. When developers give the same artificial intelligence algorithm training of different databases, the training results are different, which is very similar to the human brain, for the same child, when they are given different knowledge to learn, its three views, experience, feelings are different at each stage. The boundary of "human-machine dichotomy" has become blurred, so some scholars believe that the content generated by artificial intelligence, whether it is the "expression" of ideological content - works, or the "reflection" of technical solutions - inventions, is the crystallization of intellectual creation activities.

Therefore, the originality of content generated by artificial intelligence with autonomy is the sole contribution of that artificial intelligence. In short, it is content produced by artificial intelligence without human intervention. However, compared with Dreamwriter, ChatGPT has a higher degree of intelligence. It can not only write articles, but also edit video scripts and copywriting, and even translate and write code. The most important thing is that it can communicate with humans. At present, the artificial intelligence-generated content represented by ChatGPT has gradually become a new form of creation, and its generated content is recognized in terms of copyright. It is of great significance to ensure that human creators still have control over their creative content in the context of technological development, and can obtain corresponding rights and interests, so as to better encourage human creators to create more

valuable works[8-12].

3. The difficulty of qualitative content generated by artificial intelligence under the theory of interpretation

3.1 Artificial intelligence generates content and fruits system

Intangible artificial intelligence-generated content is difficult to explain into the fruits system based on tangible property. Some believe that the fruits in the field of property law can be introduced to solve the problem of identifying the nature of AIGC, which can indeed be a good way to break away from the bondage of the principle of human author and avoid causing an impact on the intellectual property system. According to Article 321 of the Civil Code of the People's Republic of China, "natural fruits shall be obtained by the owner; Where there is both an owner and a usufructuary right holder, the usufructuary right holder shall acquire it. If the parties agree otherwise, such agreement shall prevail. From the legal point of view, fruits are divided into natural fruits and legal fruits. Legal fruits are based on a certain legal subordination relationship, and the generated content is the product of artificial intelligence, and there is no legal and intentional relationship between the two. Although the generated content conforms to the characteristics of natural fruits produced by natural nature or natural laws, all kinds of fruits are derived from the theory of tangible property rights, and are independent of the original object, and can produce economic value in private law. The core point of Locke's discussion on labor and natural rights is: "The subject who exerts labor on a property should have rights over that property." Recognition of the fruits of intellectual property would be contrary to the idea of the tangibility and materiality of the object of tangible property. Later, Locke also pointed out that labor in the field of copyright is different from labor in other fields, and labor does not necessarily lead to monopoly and exclusive protection of knowledge. Therefore, it is inappropriate to qualitatively the content generated by artificial intelligence according to the theory of fruits.

3.2 Artificial intelligence-generated content and data products

AIGC is different from ordinary data because it has all the ingenuity of non-humans. Artificial intelligence-generated content such as ChatGPT and data property are the same in form, firstly reflected in the fact that artificial intelligence-generated content and data are both infinitely reproducible, and the generated content of artificial intelligence is different from the limited nature of traditional natural resources. Part of the content and data generated by artificial intelligence have property and personality elements, and their personality elements are mainly due to the gradual datatization of names, names, portraits, privacy and voice interests. As mentioned above, artificial intelligence uses massive data before generating content, and then outputs new content after processing by internal algorithms. The output content of artificial intelligence will contain the elements of personality, so the two are the same at this point, and there is a blend of personality elements. However, the two are very different in nature. The generation of data mainly comes from human beings, so the elements of personality in the data are directly from human beings, while the original materials for the pre-training of artificial intelligence content generation come from massive data on the network, and most of the elements of personality in the generated content are directly from the data they use and indirectly from human beings. Therefore, the generated content can be called "data of data". Artificial intelligence generated content and ordinary data are invisible, and a considerable part of the artificial intelligence generated content is also original. If the artificial intelligence generated content is processed as ordinary data, and the ordinary data is non-exclusive and repeatedly used, it will lead to a large number of original artificial intelligence generated content flooding into the market. This poses a serious threat to the existence of human authors, and does little to encourage the use of artificial intelligence such as ChatGPT to create. To sum up, compared with ordinary data, the artificial intelligence content generated by ChatGPT as an example meets the general characteristics of the latter, but it has its own characteristics: originality and certain exclusivity, so it is difficult to deal with it directly as ordinary data and is not conducive to the play of copyright incentive mechanism[13-15].

3.3 Content and work system generated by artificial intelligence

Treating AI-generated content as "current famous works" violates the principle of human authoritarianism. According to Article 2 of the Copyright Law, works of Chinese citizens, legal or unincorporated organizations, foreigners and stateless persons are protected by the Copyright Law. It is

clear that this provision limits the "works" protected by copyright law to the creations of human beings. Article 11 Paragraph 3 of the Copyright Law stipulates that only works presided over by a legal person or unincorporated organization, created on behalf of the will of a legal person or unincorporated organization, and for which the legal person or unincorporated organization assumes responsibility, shall be deemed as the author. In view of Article 3 of the Copyright Law, this provision does not indicate that as long as the content generated by artificial intelligence has "originality", the developer or user of artificial intelligence can be identified as the author, and then apply according to the creation of legal works. On the basis of system interpretation and in accordance with the relevant provisions of the copyright Law, it is necessary to first determine that the intellectual achievement belongs to the work protected by the copyright law, and then it is possible to apply the provisions of the legal person's work creation. Article 3 of the Copyright Law stipulates that a work is an intellectual achievement that is original and can be expressed in a certain form in the fields of literature, art and science, as well as a list of nine forms of work. From this article, four constitutive elements of a work can be drawn: first, the fields of literature, art and science; Second, it has originality; Third, it can be expressed in a certain form (protecting external expression, not thought); Fourth, intellectual achievements. Compared with the features of traditional works, the most fundamental difference between the artificial intelligence-generated content in ChatGPT is whether it is created by human authors, which is also the controversy in the academic field for human authors. Because the subject of creation is different, it also leads to whether the intellectual achievement reflects the human wisdom. However, there is also the biggest similarity - originality, which is the core condition of the work, and also the most difficult to distinguish between a large part of the content generated by artificial intelligence and human works. In addition, the generated content can also be copied in some form.

If the content generated by artificial intelligence is treated as "orphan works", the copyright incentive function will be invalid. However, such a delicate treatment to a certain extent solves the difficulty that the subject of creation is limited to natural persons, that is, solves the theoretical conflict that the subject of the work is limited to natural persons. However, according to Article 13 of the Regulations on the Implementation of the Copyright Law, for anonymous works, the subordinate rights of copyright except the author's name rights in the copyright rights bundle are enjoyed by the non-illegal holder who generates non-copies of the content. However, in terms of subsequent paid use, China's copyright law requires the author to be noted and consideration to be paid, which will lead to the problem of unclear subject of rights. Of course, there are also views that this problem can be solved by referring to the provisions of the American copyright law for orphan works, and the author can use the orphan works if he has fulfilled the duty of diligent search. However, artificial intelligence works are different from the scarcity of previous works, and their high productivity will cause a large number of orphan works to flood the market, which may lead to an increase in utilization costs, mainly reflected in the cost of information collection. Although certain problems can be solved according to this method, the total social cost will still increase excessively. The copyright law should encourage investors who make substantial contributions to the creation or provide substantial material support for the creation, rather than the owner of the original. Although obtaining economic benefits is not the only purpose of authors to create works, some authors are also likely to achieve personal reputation, respect, self-value realization, etc. However, from the current situation of artificial intelligence application, due to the high cost of artificial intelligence, artificial intelligence is mostly used for commercial purposes for profit. There is very little use of artificial intelligence to create for personal enjoyment. Because the incentive function of copyright law is exerted on the legal subject that contributes to originality, there is no right owner after the treatment of orphan works, which will lead to the invalidation of the incentive function of copyright.

4. The conception of the new work form of artificial intelligence-generated content under the theory of legislation

4.1 Reflections on the concept of "romanticism author" and "modularization idea of institution"

The idea of the "romantic author" is to regard the "author" as a "craftsman", capable of genius creation, but it is contrary to the reality of creation, and no author can create completely apart from the works of others, after all, "cultural production is always a question of appropriation and transformation". In addition, the idea of the "romantic author" was mostly advocated by publishers during the period of Enlightenment thought. This concept is reflected in the field of "copyright law", and one of the criticisms is that it overprotects copyright and exaggerates the originality of works. Therefore, there is no absolute one-to-one correspondence between the author and the work. The modular system was first proposed by Lee Fennell, who emphasized the integrity and indivisibility of the system. Only a work that fully meets

the requirements of copyright protection can be protected by copyright; otherwise, it will fall into the public domain and cannot be protected. Therefore, copyright is given exclusivity, which is conducive to the reduction of social transaction costs. However, the history of the development of law is a process of gradual individualization of law, and the development of digital technology provides new possibilities for the individualization of law. Personalized law has significant advantages, which can adapt to the reality of heterogeneity and provide personalized legal protection for individuals, so as to be more conducive to the realization of the purpose of legal norms. Too much emphasis on wholeness, when the emergence of new things such as artificial intelligence-generated content will inevitably have an impact on the traditional copyright system, so it is necessary to change the traditional concept and treat the copyright works system as a "toolbox", which cannot be applied in an all-or-nothing way[16-17].

4.2 A new form of work after the selection of traditional work standards

It is reasonable to set up a new work form for the content generated by artificial intelligence or take the content generated by artificial intelligence as the object of copyright. Just because the author of the content generated by artificial intelligence is artificial intelligence, and because it is not a natural person subject under the existing work type standard, it should not hold a negative position on the work inclusion of the content generated by artificial intelligence. Under the premise of treating the copyright system as a toolbox rather than an all-or-nothing concept, the key to the recognition of works in the sense of copyright law lies in how much to return to the essence of works. As one of the intellectual products, a work can obtain copyright protection because it is different from the object of material property ownership and cannot be a mechanical or assembly-line product. It needs to have innovation and breakthrough, and originality is its core feature. A work is a kind of intellectual creative achievement expressed by the author. In the process of subjective understanding of the objective world, the creative subject expresses his thoughts and feelings in the form of words, symbols, etc. Therefore, copyright protects the form of the work rather than the idea. Ideas, like the "law of gravitation", generally exist objectively and can be expressed in different ways, and when the symbols that make up these forms are arranged and combined according to a certain sequence of logic, they have creativity. Therefore, the key to the identification of works is whether they have originality and should be expressive forms. From the perspective of the working principle of artificial intelligence creation process, it selects, splits and combines the ingested data to obtain generated content, which is similar to human creation. The "boundary of human-machine dichotomy" of creation is fuzzy, and it is difficult for ordinary people to distinguish artificial intelligence-generated content from human works without specifying the source, so in form, artificial intelligence-generated content meets the requirements of originality conditions.

5. Conclusions

Author-centrism leads to the problem of content identification generated by artificial intelligence, and the solution lies in "seeking commonality and maintaining individuality". The essence of intelligent creation is to make artificial intelligence realize the simulation of some human thought expression process and physiological intelligent behavior. The content generated by artificial intelligence is based on the storage and intake of massive data under the preset deep learning algorithm framework, and based on the analysis of internal procedures and big data situations, so the generated content of artificial intelligence and human works are in common and overlap lies in originality. As Ulbian once said, "Law is authoritative, but not eternal." It is reasonable to separate the artificial intelligence products from the conditions of work identification under the current copyright law, discard the standard threshold restriction that the creation subject of a work is a natural person, and set up a special form of work different from traditional works to give it the protection status. Therefore, some scholars believe that from the perspective of results (objective criteria), the content generated by artificial intelligence has the appearance form and information consumption function of copyright works, which conforms to the nature of copyright law works. In addition, in terms of the model architecture that deep learning algorithms give artificial intelligence progressive perfection and the iterative update of deep learning algorithms, it is desirable to assign the term "creation" to the artificial intelligence-generated content that increasingly fits the characteristics of human works. However, the name of the new work form should be "artificial intelligence creation", so as to distinguish between "instrumental artificial intelligence generated content" and "artificial intelligence generated content with autonomy", the latter being "artificial intelligence creation", "artificial intelligence creation" will be a proper term for the creation of a special work form. In the future, the content generated by artificial intelligence should be included in the scope of protection through the amendment of the copyright system. Some scholars believe that it is

of great reference significance to propose how to adjust the existing intellectual property theory and legal system to provide a reasonable and necessary rights protection system for AI texts or other AI works, including AI art works and music works.

References

- [1] Wu Handong. *Questions on the Copyright Law of works generated by artificial Intelligence* [J]. *Chinese and Foreign Law*, 2020, v.32, No.18(03):93-11
- [2] Wu Handong. *Institutional Arrangement and legal Regulation in the era of Artificial Intelligence* [J]. *Social Science Abstracts*, 2017(12):76-78.
- [3] Wu Handong. *Intellectual Property Law* [M]. Beijing: Law Press, 2014.
- [4] Xiong Qi. *Legal Structure of Copyright Incentive Mechanism* [M]. Beijing: China Renmin University Press, 2011.
- [5] Shi Jinchuan, Chen Chunliang. *Law and Economics* [M]. Beijing: Peking University Press, 2014.
- [6] Feng Xiaoqing, Pan Baihua. *Research on the identification of Artificial Intelligence "creation" and the Protection of property rights and interests -- Comment on the "First case of Copyright infringement of Artificial Intelligence-Generated Content"* [J]. *Journal of Northwestern University (Philosophy and Social Sciences Edition)*, 2020, 50(02):39-52.
- [7] Yi Jiming. *Are artificial intelligence creations works?* [J]. *Legal Science (Journal of Northwest University of Political Science and Law)*, 2017, 35(05):137-147.
- [8] Jing Jing. *On the protection of Adjacent rights of Artificial Intelligence-generated Content* [J]. *Jinling Law Review*, 2020(01):259-278.
- [9] Huang Yuye, Malaysia Airlines. *Rights ownership of works generated by artificial Intelligence from the perspective of Fruits* [J]. *Journal of Henan Normal University (Philosophy and Social Sciences Edition)*, 2018, 45(04):23-29.
- [10] Yang Lihua, Wang Shitong. *Reflections on the Objectiveness of copyright in artificial Intelligence-generated Content -- and on the selection of originality criteria for judging works* [J/OL]. *Journal of Beijing university of aeronautics and astronautics (social science edition): 1-17* [2023-10-24]. DOI: 10.13766/j.b HSK. 1008-2204.2023.0862.
- [11] Wang Qian. *On the qualitative analysis of artificial intelligence-generated content in copyright law* [J]. *Political Science and Law Forum*, 2019, 41(04):16-33.
- [12] Zhou Xuefeng. *Analysis on tort liability of generative Artificial Intelligence* [J]. *Comparative Law Research*, 2023(04):117-131.
- [13] Xu Jiali. *Copyright ownership of artificial intelligence products* [J]. *Jinan Journal (Philosophy and Social Sciences)*, 2023, 45(04):37-49. (in Chinese)
- [14] Feng Xiaoqing, Hao Mingying. *Generate invention patent protection system of artificial intelligence research* [J]. *Journal of hunan university (social science edition)*, 2023 ((02): 143-152. The DOI: 10.16339/j.carol carroll nki HDXBSKB. 2023.02.019.
- [15] Yu Xingzhong, Zheng Ge, Ding Xiaodong. *Six issues in Generative Artificial Intelligence and law: A case study of ChatGPT* [J]. *China Law Review*, 2023(02):1-20.
- [16] Yin Zhiqiang. *Why Artificial Intelligence is "human": Civil Law response in the era of Artificial Intelligence* [J]. *Social Science Research*, 2023(01):60-70.
- [17] Feng Xiaoqing. *Intellectual Property Law in the Digital Age* [J]. *Digital Rule of Law*, 2023(03):25-45.