



Research on Copyright Protection under the Background of Big Data Application

Tan Xu^{1a}, Wei Li^{2b}(Corresponding Author)

¹Harbin Engineering University/Humanities and Social Sciences, Harbin, China

²Harbin Institute of Technology/ School of International Studies, Harbin, China

^akyotann2602@yahoo.co.jp

^br11217@163.com

Abstract. Accurately recognizing copyright protection in the era of big data is one of the important ways to protect intellectual property rights, and it is also an important issue to promote science and technology to serve the socialist economic construction with Chinese characteristics on the track of rule of law. The purpose of this paper is to interpret in detail the necessity and importance of copyright protection under the background of big data application through legal analysis and case analysis, and on this basis, to explore copyright protection methods in line with the era of big data, and to open a new chapter in the research of legal rights in the era of big data in China.

Keywords: Big data, Copyright, Intellectual property rights, Legal rights

1 Introduction

Big data is a brand-new concept arising from the continuous development of the Internet. When big data, as a pioneer of science and technology, rapidly enters all fields of our lives, intellectual property protection has been challenged unprecedentedly. This paper takes copyright in intellectual property law as the breakthrough point, and focuses on the necessity and importance of copyright protection under the background of massive data application, so as to dig deep into the implementation mode of copyright protection. It focus on the inability to use the human brain or general software and tools in a specific time. It is project that captures, manages, and processes data collections [1]. Intellectual property is an abstract concept, which is a customary name for copyright, patent right, trademark right and other property rights based on creative achievements and commercial marks [2]. After the founding of the People's Republic of China, in 1990, China passed the first Copyright Law, which was revised three times in 2001, 2010 and 2011. In 2017, the General Principles of Civil Law once again clarified that works are objects of intellectual property rights and protected by civil law. There are 9 categories of protected works, namely (1) written works, (2) oral works, (3) music, drama, folk art, dance works, acrobatic works, (4) art and architectural works, (5) photography works, (6) films and works created by similar methods, (7) graphic works and

© The Author(s) 2024

Y. Chen et al. (eds.), *Proceedings of the 2023 3rd International Conference on Modern Educational Technology and Social Sciences (ICMETSS 2023)*, Advances in Social Science, Education and Humanities Research 784, https://doi.org/10.2991/978-2-38476-128-9_52

model works such as engineering design drawings, product design drawings, maps and schematic diagrams, (8) computer software, and (9) folk literature and art works.

It should be noted here that due to different policies and reasons, China's copyright law also stipulates three unprotected objects, namely, the first, official documents and their official translations; Second, current affairs news; Third, calendar, general number table, general table and formula [2]. The following is an analysis of the case between Beijing Filin Law Firm and Baidu Netcom Technology Co.

2 Copyright Infringement Cases under the Background of Big Data Application

The Plaintiff is Beijing Filene Law Firm, referred to as "Filene Law Firm". The Defendant is Beijing Baidu Netcom Technology Co.3. The "Judicial Big Data Analysis Report of Film and Television Entertainment Industry-Film Volume Beijing" (hereinafter referred to as "Article Involved") published by the Defendant Baidu Netnews Company on Baijiahao is the plaintiff's first article published on the film law firm WeChat official account on September 9, 2018. The Defendant published the article on September 10, 2018 without the plaintiff's permission, infringing the plaintiff's right to disseminate information on the Internet; The defendant deleted the introduction, retrieval overview, annual trend chart of the number of cases in the film industry and the "note" at the end of the article involved, which infringed the plaintiff's right to protect the integrity of the work; The defendant deleted the signature of the article involved and infringed the plaintiff's right of authorship; The defendant's tort caused economic losses to the plaintiff [3]. The plaintiff requests: (a) The defendant immediately stops the infringement and deletes the articles involved; (b) The defendant apologized, eliminated the influence, and issued an apology statement on the 100 accounts of "Golden Master"; (c) The defendant compensated the plaintiff for economic losses of 10,000 yuan; (d) The defendant compensated the plaintiff for the reasonable expenses paid for safeguarding rights, 30 yuan. The defendant replied that: (a) The article involved is not original, and it adopts legal statistical data analysis software (this software is Vico Advance Legal Information Database, hereinafter referred to as "Vico Advance Database".) The data in the obtained report and report were not obtained by the plaintiff through investigation, search or collection, and the charts in the report were not created by the plaintiff through his own intellectual labor, but were automatically generated by the Vico Advance Library, so they do not belong to the scope of copyright protection; (b) The plaintiff is not the eligible subject of this case; (c) The plaintiff has no evidence to prove that the Baijiahao platform published the alleged infringing articles; (d) The plaintiff claimed that there was no factual and legal basis for the defendant to apologize.

Focus of controversy: (a) Whether the plaintiff is a qualified subject; (b) Whether the defendant has committed the alleged tort; (c) Whether the defendant's defense that does not constitute infringement is established.

The Court held that the biggest point of contention in this case was whether the analysis report automatically generated by WeiKe Case constituted a work. From the perspective of the process of generating the analysis report, the computer software used by

the plaintiff in this case was provided by "WeiKe Cases", in which keywords, trial courts and other search conditions were set, and then "visualisation" was selected, and the program would automatically generate a visual report of the data, which covered curve diagrams and graphs. The program automatically generates a visual report of big data in the form of curve graphs, bar graphs, ring graphs and other forms of visual analysis graphics, and then analyses the data in the form of graphics, "the content involved reflects the selection, judgement, and analysis of the relevant data, and it has a certain degree of originality", but since "the creation by a natural person is still a work under the copyright law, it is still a work of art under the copyright law"[3]. However, since "the completion of the work by a natural person is still a necessary condition for a work under the copyright law", and the work was "created" by WK Priority Library, the work could not be considered a work within the meaning of the copyright law. The current academic views on the content of artificial intelligence generation.

At present, the scholastic circles are mainly divided into two viewpoints on whether the works generated by artificial intelligence belong to the "works" in copyright law and are protected. One view holds that the contents generated by artificial intelligence belong to the works in the sense of copyright law and can be included in the object category protected by copyright law [4]; Another point of view is that judging whether the content generated by artificial intelligence belongs to the works in the sense of copyright law should not only look at the external manifestations, but should be judged according to the generation process. Because the content generated by artificial intelligence is actually the result of algorithm and template processing, it cannot be seen as a "work" protected in copyright law [5]. The views of the above two schools reflect the scholarly views on the ownership of the rights of the content generated by artificial intelligence. Although the legal circles have not yet formed a unified opinion on this, at present, it is believed that the copyright of the content generated by artificial intelligence should do not belong to the mainstream of artificial intelligence [6].

3 The content of artificial intelligence generation from the perspective of intellectual property law

(1)The necessity of copyright law to protect the content generated by artificial intelligence

As we all know, the object protected by intellectual property law is the achievement of human wisdom, which is embodied in the protection of economic interests. Specific to copyright, copyright law should protect creators' intellectual achievements and ensure the economic benefits generated by their works, so that they have greater enthusiasm, passion and mood to create new works. As a new industry, the content generated by artificial intelligence needs high technology and a large amount of funds as the support for its development. If it cannot be protected by copyright law, it will definitely have an adverse impact on its development. For example, creators will not further invest a large amount of funds and technology in it, which will lead to the reduction or quality reduction of the content generated by artificial intelligence, which is not only detrimental to the advancement of cultural industry, but also does not take advantage of the

advancement of human civilization. Therefore, no matter from the main legislative objectives of intellectual property law or from the perspective of advancing the cause of human civilization, the content generated by artificial intelligence should be protected by copyright law.

(2)The content generated by artificial intelligence conforms to the definition of "work" in China's copyright law

Works are the achievements of intellectual creation, the expression of thoughts, the expression in the field of literature, art and science, and must be original [7].For example, Article 2 of China's "Regulations for the Implementation of Copyright Law" defines works as "intellectual achievements that are original in the fields of literature, art and science and can be reproduced in a tangible form"; Article 102 of the Copyright Law of the United States stipulates: "Works created by the author and fixed in tangible media-whether existing or future inventions-so that they can be perceived, copied or otherwise disseminated directly or by means of machinery or devices shall be protected according to law"; Article 2 of Japan's Copyright Law stipulates: "Works are original expressions of thoughts or feelings in the fields of literature, art, science and music"; Article 2, paragraph 1, of the Berne Convention provides that "the term literary and artistic works includes all achievements in the literary, scientific and artistic fields, regardless of their form or manner of expression"[8] .

Looking at the contents generated by artificial intelligence, they are also expressed by words, pictures, tables, notes, data, etc., which can be objectively perceived by human beings. Therefore, they meet one of the requirements of works in China's Regulations for the Implementation of Copyright Law, that is, "intellectual achievements that can be copied in a tangible form", and also meet the requirements of American Copyright Law "fixed in tangible media". Therefore, as long as the content generated by artificial intelligence is "original" and belongs to the ideological expression in the fields of literature, art and science, it should be protected by copyright law.

(3)The ownership of content copyright generated by artificial intelligence

At present, there are three main ways of obtaining copyright in various countries, namely, the principle of obtaining copyright by registration, the principle of obtaining copyright by marking and the principle of obtaining copyright automatically.

(a) Principle of Acquisition of Registration

The principle of obtaining registration requires that the copyright can only be obtained after the creation of a work is registered with the copyright management department. Its advantage is that the evidence is obvious in the litigation of confirming the right, but its disadvantage is that the procedures are complicated and it violates the Berne Convention. Therefore, most countries that originally implemented the registration system to obtain copyright have simplified the procedures or abandoned the registration system. Even if the registration system is retained, most countries only use registration as a means to prove the enjoyment of copyright, not as a condition for obtaining copyright [9].

(b) The principle of obtaining annotation marks

The principle of obtaining annotation marks requires that copyright marks should be marked on every copy when a work is published. According to the provisions of Universal Copyright Convention, the Copyright mark includes three contents. The first one

is similar statements such as "No Copyright", "Copyright Retention" and "Copyright Ownership", or it must be abbreviated to the letter C (the initials of English Copyright) and added with a circle. If it is an audio product, add a circle outside the letter P (the handwritten letter of English Phonogram); Item 2. The name of the copyright owner; Third, the publication year of the work. The characteristic of the principle of obtaining copyright by marking is that if the above three marks are marked or the above marks are not recorded in proper places, copyright cannot be enjoyed. Because adding copyright marks is simple and easy to implement, and can be used as preliminary evidence to obtain copyright, even countries that adopt the principle of automatic acquisition of copyright generally add copyright marks to copies of works [10].

(c) Principle of automatic acquisition

The principle of automatic acquisition of copyright means that after the creation of a work is completed, the copyright is automatically generated without performing any formalities. This principle was established in the Berne Convention of 1908, and is currently adopted by most countries in the world. The advantage of the principle of automatic acquisition lies in the high level of protection of works, that is, once the works are completed, the copyright will not be lost because of any human factors, and it can better protect the interests of copyright owners; The disadvantage is that once copyright disputes occur, it is difficult to confirm the ownership of copyright. China implements the principle of automatic acquisition of copyright. In addition, the National Copyright Administration has also promulgated the Trial Measures for Voluntary Registration of Works. Authors can voluntarily register their works. This registration is not a legal condition for copyright acquisition, but it has the function of preliminary evidence, which is conducive to confirming rights in case of copyright disputes [11].

In addition to the above three principles, some countries require that works must be fixed in writing or other forms before they can enjoy copyright, for example, Britain. The biggest advantage of this practice is that it is easy to obtain evidence when copyright disputes occur in the future. However, if it is an oral work, its creation is not fixed, and it cannot meet the requirement of "fixed", so it cannot obtain copyright in Britain [12].

As for the content generated by artificial intelligence, because the content generated by artificial intelligence involves developers, users and investors of artificial intelligence, they have more or less direct or indirect contributions to the content generated by artificial intelligence. For example, investors of artificial intelligence have invested a lot of manpower, material resources and financial resources in the process of developing artificial intelligence, and have to bear a series of risks brought about by development failure; The developers of artificial intelligence have invested a lot of energy, contributed their own wisdom and carried out mental work; The trigger of the user contributes to the final generation of the content. Therefore, based on the principle of fairness, China's current copyright law adopts the contractual principle for the copyright ownership of the content generated by artificial intelligence, that is, the copyright of the content generated by artificial intelligence is obtained by the agreement of the developers, investors and users of artificial intelligence. In the absence of agreement or unclear agreement, reference should be made to the right ownership arrangement of film works [13].

4 Conclusion

The birth and wide application of new technologies such as big data and cloud technology have driven the rapid development of artificial intelligence industry, and artificial intelligence has been rapidly applied in literature, painting, music and other cultural and artistic fields. The emergence of the content generated by artificial intelligence not only promotes the rapid development of human civilization at a high speed, but also breaks many people's cognition of inherent things. At the same time, the birth and progress of this new technology also brings certain impact to the law. Intellectual property law, as a legal system that escorts the birth of new technologies, new technologies and new wisdom achievements, how to adjust, change and update the contents of copyright law in order to keep pace with the development of science and technology and the times is a new topic that must be considered by the legislative circles in China.

Acknowledgment

This paper is the phased research results of the following four topics.

1. General Project of National Social Science Fund (22BDJ054), 2. General Project of Philosophy and Social Science Research Planning in Heilongjiang Province (21TQB115), 3. Heilongjiang Province Philosophy and Social Science Research Planning Special Project (21ZSD275), 4. Heilongjiang Province Red Cultural Resources Development and Utilization Think Tank (ZKKF20220277).

Reference

1. Guo Jinyong. Research on security protection methods and technologies of big data. *Digital technology and application*. 2023, 41 (06): 225-227
2. Intellectual Property Law Writing Group: *Intellectual Property Law*, Higher Education Press, 2019, pp. 44-45.
3. Beijing Internet Court (2018) Beijing 0491 Minchu No.239 Civil Judgment
4. Wu Handong. Institutional Arrangement and Legal Regulation in the Era of Artificial Intelligence. *Legal Science (Journal of Northwest University of Politics and Law)*, 2017 (5): 128-136.
5. Wang Qian. On the characterization of the content generated by artificial intelligence in copyright law. *Legal Science*, 2017 (5).
6. From Lixian. Copyright and copyright ownership of content generated by artificial intelligence. China Publishing, 2019 (1).
7. Intellectual Property Law Writing Group: *Intellectual Property Law*, Higher Education Press, 2019, pp. 37-38.
8. Intellectual Property Law Writing Group: *Intellectual Property Law*, Higher Education Press, 2019, pp. 37.
9. Intellectual Property Law Writing Group: *Intellectual Property Law*, Higher Education Press, 2019, p. 47.
10. Intellectual Property Law Writing Group: *Intellectual Property Law*, Higher Education Press, 2019, p. 47-48.

11. Intellectual Property Law Writing Group: Intellectual Property Law, Higher Education Press, 2019, p. 48.
12. Intellectual Property Law Writing Group: Intellectual Property Law, Higher Education Press, 2019, p. 48.
13. Sun Yurong and Zhou Hwei: Intellectual Property Law: Jurisprudence of Case Rules, Intellectual Property Publishing House, 2021, p. 65

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

