Research on Copyright of AI Copycat Works

Mengqi Zhang\textsuperscript{1, a}, Cheng Liu\textsuperscript{2, b}, Zi Wang\textsuperscript{3, c}
\textsuperscript{1}Lanzhou University
\textsuperscript{2}Procuratorate of Xianyang city police
\textsuperscript{3}Lanzhou University

Abstract. The continuous maturity of big data, deep learning, intelligent algorithms and other technologies has promoted the third development boom of artificial intelligence. The Chinese government has also promoted the development of artificial intelligence as a national strategy. On the one hand, artificial intelligence technology is gradually penetrating into all aspects of life. For example, through the integration of mobile phone industry and education industry with artificial intelligence, the promotion efficiency of products and services has been greatly improved, and the audience can obtain better service quality. On the other hand, the existing legal system of China cannot solve all the problems of artificial intelligence. The author will attempt to analysis the present situation of artificial intelligence copyright infringement in our country, the protection of different law systems and the prospect of the future about artificial intelligence copyright infringement.

1 ARTIFICIAL INTELLIGENCE AND COPYRIGHT

1.1 AI definition and type

The definition of artificial intelligence (AI for short) is just like the ancient philosophers arguing about "how can people be human". Wikipedia states that "AI is a computer program that makes reasonable actions based on the perception of the environment and obtains the maximum benefit." It is considered to be a comprehensive definition of the industry textbook style, which focuses on demonstration, but is balanced and comprehensive [1]. Another group of scholars argued that "AI is a computer program that learns [2]." By learning, we mean "deep learning", a machine learning method that is flexible in its expressive power while allowing computers to keep trying until they finally get close to a goal. To put it simply, deep learning is to treat what a computer wants to learn as a mass of data, throw the data into a complex, multi-level data processing network (deep neural network), and then check whether the resulting data processed by the network meets the requirements -- if it does, the network is retained as the target model; if it does not. Again and again, perseverance to adjust the parameters of the network Settings, until the output meets the requirements. Big data has provided enough conditions for AI learning. Martin Hilbert has summarized that the big data we often talk about today is actually the data generated after 2000, because of the increase in the capacity of information exchange, information storage and information processing. Besides, big data can replace the traditional sampling survey and obtain information in real time. By mixing multidimensional information, you can also mine data in the analysis and make decisions.

According to the "smart" degree and scope of capabilities, AI is divided into three levels by the AI industry: weak, strong and super AI. Weak AI, also known as Narrow AI or Applied AI, refers to AI that focuses on and can only solve problems in specific fields. Weak AI is more inclined to tool attributes, serving or helping us with our own expertise. Strong AI, also known as artificial general intelligence or Full AI, refers to AI that can be competent for all human work. Artificial intelligence with "consciousness" is another standard of strong artificial intelligence considered by some scholars.

1.2 Artificial intelligence is a rich art form

1.2.1 AI becomes the main writer of news articles.

In the early morning of March 17, 2014, a small earthquake shook Los Angeles, causing people to wake up from their sleep. Within three minutes, a detailed report from the Los Angeles Times about the quake appeared online, including data from seismic stations and recent observations. What's surprising is that the accurate data in the report came from artificial intelligence. Back in 2011, a company called Automated Insights, led by its chief engineer, Robbie Allen, developed an artificial intelligence (AI) automated writing program. Soon after, a program called Wordsmith entered the Associated Press and other news organizations, producing 300 million news articles in 2013 alone, more than the major news media combined, and a new high in 2014. Write more than 1 billion news articles [3].
1.2.2 Second, AI is a contemporary Wang Xizhi.

On the eve of the Spring Festival in 2017, Alibaba company sent New Year greetings to people with Spring couplets of various fonts, Ali Cloud artificial intelligence ET scene, "Kyushu sky beautiful, Weiyang Yun light person Taikang" was completed in the "pen" of master ET. It is composed by the demand after on-site experience [4].

1.2.3 AI transformed the painting master

Mobile painting app Prisma became popular on wechat moments in the summer of 2016, automatically generating paintings based on photos. Prisma has learned to copy classic Chinese and Western paintings, mastering a variety of oil, watercolor and manga techniques, as well as coloring techniques, brushstrokes, wet and dry techniques, and can create more than 20 different styles of painting in her work.

1.2.4 The field of artificial intelligence has a long history of exploring music creation

In 1981, David Cope of the University of California, Santa Cruz, began experimenting with the use of computers to generate music automatically. After seven years, he wrote EMI (Music Intelligence Experiment). Experiments in Musical Intelligence, a program that has been tested to compose more than 5,000 Bach compositions in a single night. At one music festival, the organizer specially mixed EMI scores into the music, and most of the audience said they heard Bach's music. Since then, David Cope has updated EMI several times to include Beethoven, Chopin, Rachmaninoff and others.

1.2.5 AI has become a poet.

The general public many times through the media, of course, some people have "her" company. According to reports, by the time Xiaoice turned 18 on September 17, 2016, she had had more than 20 billion conversations with more than 42 million people, including text, voice, phone and other forms. Through communication with people to accumulate a huge database, on May 18, 2017, the first poetry collection "Sunshine Lost the Glass Window" was held in Beijing Creative General Press 1949 Media Industrial Park, this is the first artificial intelligence original work, "she" wrote: "I lie in my bed again, is not rare sweet feeling, a candlelight, suddenly become lonely town".

As a new favorite in literature and art, artificial intelligence covers many art forms, such as calligraphy, poetry collection, painting and composition. It is often said that art is our final temple, and it is unique to human beings. Now it seems more or less ironic that the future will be less and less exclusively human, and more and more human space will be covered by machines.

1.3 Development status of AI copyright

1.3.1 The exploration of common law countries

As the country where Alan Mathison Turing, the father of AI, is located, the United Kingdom has issued a series of policies on AI since 1973, including the Report on the Ownership of Intellectual Achievements and Technological Innovation and the Copyright, Design and Patent Law, and issued relevant reports, such as Artificial Intelligence: Opportunities and Impacts for Future Decision Making. According to British copyright law, the "person who makes necessary arrangements" for a work generated by a computer program is regarded as the copyright owner and author of the work. Those who make "necessary arrangements" for computer-generated works may include programmers, users, or investors in artificial intelligence systems or devices. In a word, important operators can be regarded as the authors of works generated by computer information equipment under human operation [5].

Australia currently holds the same position and attitude as Britain. In 2002, we further improved the content of the legal provisions on artificial intelligence, and refined the "who is the author of the creation", that is, only those who fully control the computer program can be regarded as the author. Even though Australia pays attention to the improvement and revision of legal provisions, there are still many intellectual property protection issues that need further explanation at this stage.

Compared with the first two countries, the United States does not have much legislation on artificial intelligence.

1.3.2 The exploration of civil law countries

In recent years, Japan has promulgated the Copyright Law, the Intellectual Property Promotion Plan 2016, the Intellectual Property Promotion Plan 2017, and issued the Report of the Next Generation Intellectual Property System Review Committee. At the same time, it puts forward clear suggestions on whether the creation of new technology belongs to human or artificial intelligence. In May 2016, Japan clearly stated in the Intellectual Property Promotion Plan that it may also be necessary to give certain intellectual property protection to AI creations with certain market value [6]. However, those without relevant precise and specific description are all determined as human works.

South Korea is also a major country in the development of electronic science and technology. Its Software Policy Research Institute has made specific instructions on the copyright and intellectual property protection of creative crops during the AI development process, that is, if intellectual property disputes and issues are involved, the author of the creation must be clearly stated.
2 RESEARCH ON COPYRIGHT INFRINGEMENT OF ARTIFICIAL INTELLIGENCE

2.1 Cases of AI plagiarism

The People's Daily once published an article about the widespread plagiarism of online novels, in which there was a description: "The use of automatic writing software has become an open secret in the online literary world... The online novel" The Rise of Cold Men "was pieced together using novel writing software, and the suspected plagiarism is mainly about background knowledge, appearance description, environment description, scenery description, scene description, etc [8]." Media workers once tracked and reported the phenomenon of plagiarism of online novels and investigated writing software in depth. When searching for the word "automatic writing software" on an e-commerce platform, there would be multiple query results, including a super writing software with a price of 498 yuan named "great writer". The software provides four templates. The atomic template is the combination of the smallest elements of the article content, including letters, words, sentences, etc; The paragraph template focuses on the details, and shapes the characters' souls through character contrast, costume and scene dialogue; The sketch template is mainly used to select novel types, including mainstream novel types; The charter template belongs to the plot structure. It can automatically generate content according to users' requirements, but because it is the reproduction and integration of existing materials, the splicing trace is obvious. Of course, the material library is all the novels and writing materials that can be searched in the network. To put it bluntly, it is a kind of covert plagiarism with the help of artificial intelligence.

2.2 Definition of plagiarism

The Modern Chinese Dictionary sorts out three explanations for the word "plagiarism": (1) Publish others' articles and works as their own without permission; (2) Attack the enemy by detour; (3) Regardless of reality, copy other people's ideas, methods, experiences, etc. The Yuan Zhao Law Dictionary says: copy or steal other people's words or ideas for yourself [5]. Plagiarism as defined in the copyright law is also limited to copying or stealing the expression of others as their own works.

The basic attribute of the text is intertextuality. In the sense of literary text creation, there is no development of literature without mutual influence, penetration and transformation between existing literary works. Not only literature, but also music, painting, calligraphy and other forms of artistic expression are different from existing elements, resulting in new expression techniques. The expression of the legal definition of "plagiarism" is too general, and the identification of plagiarism in the practical circle is also controversial. Copying the content of other people's works is plagiarism? The social restriction on plagiarism is not to prohibit copying, but to prevent some people from violating the interests of others by plagiarism. In contrast to the traditional meaning, such behaviours as "copying" and "copying privately" are absolutely unacceptable to the society. When identifying human plagiarism, we mainly adopt the thought text dichotomy, which only discusses the plagiarism of texts and ignores the ideas that are difficult to determine. However, they are different in AI and lack the ability of independent thinking, and the ideas expressed in their works are undoubtedly not their own ideas. Therefore, the definition of AI plagiarism should be: stealing the spirit results of others as an unfair act.

2.3 Identification of plagiarism

2.3.1 Identification method of artificial plagiarism

In current practice, plagiarism is no longer a "word for word" copy, but a selective rewriting. Take the novel as an example. The role setting and plot development direction of one work and another work echo each other, and the possibility of similar relations leading to plagiarism is getting smaller and smaller. It is difficult to determine whether there is plagiarism. At the same time, the detailed division of art types and styles also solidified the expression of a certain style, and the emergence of a large number of similar works also brought trouble to the identification of plagiarism, so there is no uniform standard for the identification of plagiarism. Scholars can divide the identification method into two aspects: actor and work. From the perspective of the actor, it can be divided into four elements of infringement and illegal act; From the perspective of works, it can be mainly divided into "two-step method" (by proving "contact and similarity", and then proving whether it is suspected of illegal appropriation or practical similarity) "Syllogism" (including three methods: "abstract method", "filtering method" and "comparative method"). The specific operation is to separate ideas from "expression of ideas", delete the content belonging to the public domain, compare the rest, and check whether the infringing works are substantially similar to the infringed works), "two-step method and syllogism".

2.3.2 On plagiarism of artificial intelligence creations

Artificial intelligence creations need to learn from other people's works, but must maintain respect for the original author's work. Malicious tampering with the content of another person's work must be recognized as plagiarism. This is not only opportunism, but also an insult to literary and artistic creation. The academic papers need to be checked for duplication, which is the key issue for Qianqian's sub papers to pass. Such strict standards are to ensure the originality of the article, and the research topic is more meaningful. AI creations can also set the corresponding text repetition rate, which limits the ability to only capture and integrate information. If the AI text repetition rate exceeds this repetition rate, it should constitute plagiarism.
3 PROSPECTS FOR FUTURE AI PLAGIARISM

Establishing the main body of AI plagiarism infringement can help the infringed to recover their losses through litigation, but it cannot stop AI plagiarism. I have my own ideas on the governance of AI plagiarism in the future. First of all, set up a professional regulatory agency to adapt to the rapid development of high-tech industries, formulate appropriate industry standards, strictly test the application programming function, pay special attention to AI products with deep learning ability, do a good job of post release monitoring, and timely deal with problems; Secondly, the real name system and credit reporting management system are introduced. Literary and artistic works created by artificial intelligence must be authenticated by real name system, otherwise partial functional experience cannot be achieved, which further facilitates the identification of the responsible person of artificial intelligence infringement and the maintenance of the legitimate rights and interests of others. The credit reporting system is similar to a punishment method. For the responsible person of artificial intelligence infringement, it is included in the system blacklist, and the rights of the subject will be limited for a period of time; Finally, promote the mutual restriction of artificial intelligence, hide the plagiarism of artificial intelligence, and the effect of traditional identification methods is not good. Borrow similar duplication checking software, "attack the shield of the son with the spear of the son", simplify the identification method, and achieve technological innovation through confrontation, hoping that the plagiarism of artificial intelligence will not occur in the future.

4 CONCLUSION

The arrival of the era of artificial intelligence has liberated human hands, not only bringing impact to various sectors of society, but also bringing new challenges to the legal system. First of all, the existing legal system lags behind the development of artificial intelligence technology. Artificial intelligence does not have the qualification to become a legal civil subject, but its creations meet the conditions for reproduction and originality of works as stipulated in the Copyright Law, and should be copyrighted. Human beings have subjects that can be prosecuted to safeguard their own interests. Artificial intelligence creations are protected by copyright. The interests of the two have reached the best balance, so as to promote the coordinated development of artificial intelligence and human beings and create a new era of literary and artistic innovation.

References

7. Artificial intelligence strategy. "artifact" writing the first appearance of artificial intelligence, network novels can production [EB/OL].