

AI-Generated Content and Its Legal Status Under Copyright Law

Yuchen LU

Law School, Renmin University of China, Beijing, China

lyc123@ruc.edu.cn

Abstract. As artificial intelligence (AI) has been developing at a staggering speed, arisen disputes and problems have showed how people perceive this with worries and uncertainties. Could AI-generated content be regarded as original works? Who is the author of such content? Could the regulation of AI-generated content fit into the already established system of copyright law? And will the empowerment of AI-generated content become threat to human-created works and author's legitimate interests? There are multiple issues waiting to be solved. This article discusses AI-generated contents and its significant effects. It looks at how these development clarified the nature and extent of its copyrightability and authorship, which are two legal matters crucial to determining whether and how AI-generated works should be protected by copyright. From this analysis, the paper draws three conclusions: at first, for those content that are output through complex instructions and that are significantly different from the learning data sources, they could be regarded as original works from external forms. Secondly, it is premature to grant AI the status of a legal person at this point in time. And thirdly whether and under what circumstances should its copyright be granted, and who should be recognized as the author should be left to a case-by-case judgement.

Keywords: Copyright law, artificial intelligence, AI-generated content, authorship.

1. Introduction

Though artificial intelligence was not a newly sprouted thing, the emergence of strong AI could still be astonishing; they are smart, considerate and well-rounded as humans, equipped with powerful capabilities of machine learning, computer vision and natural language processing, which enable them to view and learn things independently. Since the invention and iteration of deep learning arithmetic like Generative Adversarial Network (GAN) from 2014, the development of artificial intelligence has witnessed a huge boost in a short time. In 2022, OpenAI released ChatGPT, an artificial intelligence with capability of human-computer interaction and easy access, marking a new milestone in such field. Meanwhile, other cutting-edge AI models have followed, for example, Bard from Google, ERNIE Bot from Baidu, and SparkDesk from iFlytek, extending AI large language models far beyond its own territories and borders. The mainstream form of those models focused on AI Generated Content (AIGC), providing easy access for users to generate complicated text, image, music, and even video.

However, there are also concomitant issues, doubt of the legal nature of AI-generated content and copyright infringement by artificial intelligence has become salient. As AI models were endowed with capabilities of simulating human, especially in those fields of literature and painting, it has been more and more difficult to tell the difference between works by AI and Human. A living example was "The Sunlight that Lost the Glass Window", which was a poetry collection created by Xiaoice, an AI software developed by Microsoft. This collection has been legitimately published in China, signaling that AI could reach humans' creative level and stimulate the appearance of human-created works to some extent. The successful publication of this collection marked the fact that the extremely high similarity of AI-generated content to human works has enabled it to be recognized in an aesthetic sense. This has made the discussion on whether rights and what kind of rights should be authorized to AI-generated content a necessity.

Though China seemed to offer an affirmative answer to these problems, its opinion wasn't just remaining fixed, with the overall attitude of most countries still remaining unclear and no consensus

from most legal scholars. On one hand, a surge of AI-generated content are meeting massive problems of their legal characters; on the other hand, creators with human identity are suffering great worries of possible infringement of their rights and pre-existing status, with the fear that content generated by AI models may “free-ride” their works.

There are also harp debates about whether AI could be recognized and regulated by existing copyright law system, for example, by approaching a more flexible interpretation of authorship; or is there a necessity to enact law specially designed for AI and AI-generated contents? Some people have called for governments and relevant agencies to instigate targeted regulations, and there were examples of such measures, like the Artificial Intelligence Act (AI Act) enacted by European Union, whose final compromise context has been confirmed by COREPER Ambassadors. Besides, the U.S. Copyright Office has also been asking for broad public advice on the interplay between AI and copyright law, meanwhile seeking for the appropriate treatment and regulation of AI-generated outputs.

However, it's still too hasty to have legislation onto agenda without considering whether there is a demand for a completely new legal area. Every law is designed to answer certain social needs, to fill certain vacancy and to solve newly emerging problems. The breakthrough of technology is one of those triggers, but the premise is that new technology does bring about legal problems that cannot be accommodated into the traditional legal system. Sometimes new technology didn't essentially change much and could still find a place in existing legal system; sometimes it just reinforced former problems or brought some potential problems into sights of the public. So before entering into the procedure of enacting, it's a must to test whether there really is such a necessity. Moreover, it's crucial to understand that it's of great significance for law to maintain stable as well as work in a timely manner. In doing research on challenges posed by a new technology, there should be scientific evidence to show that the problem as the subject matter of the research really exists [1]. The protection and protectability of AI models and AI-generated content could just be old problems arising before the development of AI algorithm and were addressed by the development of technology. So, it's necessary to definite the two types of relationship between technology and law, and to pinpoint the exact location of artificial intelligence.

This article mainly focuses on the issue of rights affirmation of AI-generated content, as this question is a prerequisite for discussing the legal nature of AI-generated content and for discussion of other issues that follow, like rights attribution and infringement, and it is also the most urgent at the present time.

Based on this background, this article will discuss the concept of artificial intelligence and its output based on its latest developing level, and will attempt to give a proper orientation of it; after that it will bring up and try to answer some questions concerning AI-generated content and its copyrightability, meanwhile seeking for solutions under existing legal system; Lastly, this article will make some suggestions about its institutional arrangement.

2. Core Concepts

2.1. Artificial Intelligence

Undoubtedly the most significant concept for this article is “artificial intelligence”. Explained from its origin and latest trends, the term “artificial intelligence” could date back to 1947, when Alan Turing firstly mentioned it in a conference in London, explaining it as “a machine that can learn from experience” which would be made possible by the same machine being allowed to change its own instructions [2]. Upon this foundation, artificial intelligence has been associated with computer or machine, but with its technological progress the latest artificial intelligence had much more association with algorithm. Taking ChatGPT as an example for latest chatbots, it's a natural language processing tool driven by specific algorithm and machine learning function, trained to follow an instruction in a prompt and provide a detailed response based on large language models.

While there are different definitions of AI, the World Intellectual Property Organization's definition is the most official and accurate one. The World Intellectual Property Organization (WIPO) eventually reached a consensus after debates and discussions, defining Artificial intelligence (AI) as “a discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence, with limited or no human intervention” [3].

2.2. AI-generated Content

AI-generated content refers to output from artificial intelligence based on its machine learning results, with its form not limited to text, picture, video, voice, etc. This article chose the word “content” instead of “work” or “creation” to describe the output by AI, as the use of work or creation presupposes the authorship of artificial intelligence and admits its output as original, which is an uncertain question remaining to be discussed.

Also, the term “AI-generated” differs from “AI-assisted”, with the latter one means the content is not created by AI autonomously, but actually the result of a human intellectual activity assisted by AI. There is no controversy that the authorship of AI-assisted content, if could be regarded as work, belongs to the human creator. So, it's crucial to distinguish the two terms, and this article mainly focus on “AI-generated” content.

3. Problems arising from AI-generated content

3.1. The legal nature of AI-generated content (Copyrightability)

The most complicated and controversial debate about AI-generated content relating copyright law is authorship and ownership. But before discussing it, there is still a prerequisite to be figured out: whether AI-generated content could be regarded as works within the meaning of copyright law? Regarding this issue, there are affirmative and negative theories, scholars with affirmative attitudes tended to advocate that the expression of AI-generated content meets the minimum creative requirements, thus it can constitute works within the meaning of copyright law; other scholars debated that AI-generated content are the result of the application of algorithms and templates, which are highly homogeneous, leaving no room for creation and cannot reflect the individual characteristics of the “creator”, thus it's impossible to meet the requirements of originality[4].

Considering legal regulations around the world, there are great similarities in the standards for identifying work within the meaning of copyright law. Taking China's Copyright Law as an example, article 3 of it stipulated the definition of work: “the term “works” means intellectual achievements in the fields of literature, art and science, which are original and can be expressed in a certain form [5].

The recognized constituent elements of a work are it should be the form of expression; it should have originality; and it should belong to the field of literature, art or science. As for AI-generated content, there is no substantial difference in the external form between such output and intellectual achievements created by humans. Especially with the advancement of deep learning and autonomous learning function, it's not difficult for AI to imitate the details of creating habits unique to human, making it has the ability to generated content based on a wide range of types, and it could easily meet the requirements of expression and the limitation of domains.

As for originality, considering advanced operating principles, AI-generated content are not rigid results produced by executing established algorithms and orders. The latest technology enabled AI to have thinking tactics to some extent, meaning their outputs are no longer mechanically stitched content. Though those outputs are generated based on learning from original works like novels, reports or paintings, those content has already gained substantial difference and its own originality. As a result, for those content that are output through complex instructions and that are significantly different from the learning data sources, they could be regarded as original works from external forms.

3.2. the Authorship Qualification of Artificial Intelligence

Though artificial intelligence can generate content with originality, this does not mean that artificial intelligence could automatically obtain the qualification as a creative subject and gain authorship in the sense of copyright law. It goes the orthodoxy that copyright is designed to protect human creativity, copyright laws entitle only human beings to own copyrights in virtually all jurisdictions [6]. As a result, some countries refused to give AI-generated content legal protection as they did not recognize artificial intelligence as author of their output. Even the Copyright Act (1976) of US did not ask for a ‘human’ author as a prerequisite [7], the US Copyright Office still denied registration for a painting totally produced by a highly autonomous AI model, asserting it failed to meet the requirement of “human authorship” [8]. The majority still held the opinion that artificial intelligence could not break away from the control of human and become independent in the process of generating, thus it is more rational to regard it as extension of human consciousness, which is difficult to truly obtain the qualification of a creative subject.

Could artificial intelligence be regarded as a person? To answer this, the first thing to be figured out is whether artificial intelligence has independent will resembling nature human. Free will emphasizes the consciousness, target, planning and agency of the subject. To meet this standard, artificial intelligence should be able to think and act autonomously without the whole process control from human users. While some scholars believed that with the rapid development of AI technology, AI can think and act autonomously, and even can surpass human in programming and creating with independent will [9]. Some scholars debated based on the current development of AI technology that AI cannot perform common tasks autonomously, cannot achieve the level of independent reasoning, creation or work planning. not only requires a subject to make decisions independently, but also requires a subject to exercise self-restraint; from this perspective, the latest artificial intelligence did not yet fully possess the elements of independent will. Furthermore, it’s still difficult to tell with a scientific basis whether artificial intelligence will have independent will in the future. The authenticity of this proposition remains to be tested by time and practice.

Could artificial intelligence be regarded as an “artificial human”? “Artificial human” means a legal fiction type of status, which typically allows relevant bodies, such as courts or legal experts, to ignore a fact that would prevent it from exercising its jurisdiction by simply assuming that an agency could enjoy the same rights as human beings to some extent. For example, the system of legal person. Some scholars asserted AI could be treated as person to enjoy authorship of their output applying the existing system of corporate work. From a utilitarian perspective, authorship is the currency used to obtain related rights within the meaning of copyright law. Like endowing corporation with personhood enables it to enjoy copyright of certain work directly, making AI “artificial human” could enjoy rights and bear duties, and could reduce the number of “orphan works”.

However, based on the current level of AI development, is it really necessary to empower artificial intelligence with authorship? First, there is currently no need to make AI an artificial person, as the system of legal person has its own unique social and economic background. With the outbreak of Industrial Revolution, the rapid development of industry and commerce made it necessary to organize factories, companies and other commercial organisations with the help of independent legal subject status in order to effectively manage and operate the business of these organizations. Therefore, through the law, non-natural persons are viewed as legal subjects similar to natural persons, so that they could enjoy rights, assume obligations and bear their own risks. As far as “Narrow AI” is concerned, it is still at the stage of simple procedures dealing with easy problems and has not yet reached the point where it must be authorized the status of an independent subject. Second, considering the number of it at present, the existence of orphan works could bring the benefits of cultural prosperity to some extent without causing too many problems due to unclear ownership. Finally, warranting AI the status of subject could bring about a series of controversial problems, for example, can Artificial Intelligence assume legal obligations independently, especially for criminal liability? Some scholars deemed that current mechanisms of penalty, compensation and mandatory liability could not apply to artificial intelligence [10], which may result in AI only enjoying rights

without fulfilling obligations. It is therefore premature to grant AI the status of a legal person at this point in time.

3.3. The Copyright Attribution of AI-generated content

As transplanted from the institutional system of tangible property rights, copyright law cannot be separated from the basic legal principle that subjects and objects are not interchangeable. Therefore, the institutional design built around creators and investors should not and cannot be subverted by artificial intelligence. Since artificial intelligence itself cannot be the subject the content it generates, then should the designer or the user of AI software be recognized as the copyright owner of AI-generated content?

Some scholars debated that as both the designer and the user of AI have made necessary contributions to the creation of AI-generated content, the designer and the user should share copyright in such creation. Nonetheless, this arrangement of co-authorship may not work well in AI-generated works. The designer and the user of AI lacked the crucial elements of co-authorship, for example, not only did they lack the consensus of making joint use of the AI-generated work, but they also lacked the cooperative intent of creating such work together [4]. Besides, this model would unduly increase the cost of copyright enforcement and would lead to unclear allocation of rights and interests.

Considering the creator or the investor as subjects of works is a dichotomy that has been strictly upheld by copyright law, attributing copyright of AI-generated works to designers of the AI model has rationality to some extent. The primary theory supporting this idea is the “work made for hire” doctrine [7]. Countries around the world have adopted similar institutional arrangements, which offers protections for intellectual property involving some sort of exclusivity to prevent others from benefiting off of the creator’s work [11], for example, the “free riders”. Since the designer of AI has invested substantial money, time and creative thinking, the AI itself can be regarded as fruits of his labour, and therefore the AI-generated content has a certain degree of adjunctive nature. Thus, it could also be classified into the radiation of the designer’s copyright. There are also arguments against this arrangement. The first is the problem of double profiting by the designer of the AI. The designer has already obtained private rights for the AI itself as computer software, thus granting the designer copyright of AI-generated work would result in the designer obtaining double protection for the same act of creation. Secondly, there is a lack of subjective awareness of the designer, as the designer usually has no way of knowing the existence of the work created by users without being informed. There is no necessity of protecting such “unknown” work. Moreover, attributing copyright of AI-generated content to the designer may result in payment for use, which is extremely unrealistic at the present. This will also discourage users from using the AI models to create more content, which is contrary to the purpose of the copyright system, that is to stimulate the creation and dissemination of works.

There is also some rationality for identifying the end user of the AI software as author of AI-generated works. For example, as the content is generated following orders from the user, and this process demonstrates selectivity to some extent; The spatial and temporal relationship between the user and AI-generated work is closer; such arrangement would also eliminate the difficulty of distinguishing AI-assisted and AI-generated works compared with attributing authorship to someone else; it’s rational to regard the end user as the author as they fused their creativity with the final output[7]. But problems are also acute. Though currently very little AI-generated content was involved in commercial use, endowing users with authorship may accelerate this process as the labour of the end user is simply not worth it in the face of the possible rewards of the work. Even under pay-to-use, the reasonableness of the arrangement remains open to question.

Identifying AI itself as the author lacks basic philosophical background and practical justification, as well as contrary to the basic presuppositions of copyright law; Naming the designer or user as the author is also not conducive to commercial practices. Therefore, none of those arrangements seems to thought-out enough to satisfy all of the social needs. The mere fact that AI-generated content seems to conform requirements of a work is not a sufficient reason to bring it within the scope of copyright

law and grant rights to it. Whether and under what circumstances should its copyright be granted, and who should be recognized as the author should be left to a case-by-case judgement.

4. Proposals for Regulation of AI-generated Content

4.1. Balance of Interests

One of the design philosophies of copyright law is to delineate clear attribution of work ownership and to regulate the interests of the copyright industry of which the work is an element. While rationally assessing the role of intellectual property law as an incentive for creativity, it is important to know that creation can be driven by nature, while investment can only be driven by profit. The functioning of copyright law is mediated by the market, and copyright law is designed to ensure that industries in which knowledge is a factor can run smoothly. From this conception, a general user-centred structure concerning copyright attribution for AI-generated content could be established, taking the protection of investors' interests into account [4], meanwhile respecting the relevant agreements on the distribution of rights and obligations reached among different subjects on a case-by-case basis.

This arrangement goes as follows: In most instances, Copyright of AI-generated content should belong to the end user of AI software. The end user intellectually created the work with AI as the mechanical generator, and his strings of instructions meet the minimum standard of labor [7]. Attributing copyright to AI users promotes the production and realizes the incentive goal of copyright system, while better adapting to the current situation and development law in the field of AI. Though it is not a perfect solution, deficiencies can be remedied in a number of ways, like increasing the rigorousness of recognizing AI-generated content as works by considering user's training time for the model, the complexity and creativity of the instructions, etc, which could reduce the number of speculative users who aim to make a profit.

Besides, the interests of investors and designers should be taken into consideration, as the principle of investment is an important addition to the copyright attribution system.

This maybe not prominent in situations where the AI-generated content is not for commercial use, but in occasions when AI-generated works transacted at a certain price, the legitimacy of designers and investors benefiting from this should be recognized. In the absence of prior agreement between the parties, the designer and investor should be allowed to have a rightful claim to distribution of benefits afterwards.

4.2. Suggestions and Possible Solutions

First, with regard to the nature of rights of AI-generated content, there are mainly two ways of resorting legal interests and legal rights, and the latter can also be subdivided into three: copyright (narrowly defined), neighbouring rights and ad hoc rights. Nonetheless, allowing AI-generated content entering into the public domain or protecting them only as a legal interest through laws against unfair competition, may only provide incentives for parties to pass off AI-generated content as works created by humans in order to seek and obtain a higher level of copyright protection. Therefore, the way of authorizing rights to it remains to be the best solution.

One way is to protect AI-generated content through copyright (in the narrow sense, not including neighbouring rights). However, as noted above, while AI-generated content can be considered as works under certain conditions, the attribution of authorship is highly controversial and needs to be judged on a case-by-case basis, thus making it difficult to establish a more unified institutional arrangement to regulate it.

Another way is to protect it through neighbouring rights. Neighbouring rights, as part of copyright in the broader sense, are the rights of the distributor of a work to the creative labour and investment he or she has made in distributing the work. This viewpoint argues that the inclusion of AI-generated content in the system of neighbouring rights helps to avoid the theoretical obstacles and logical conflicts arising from the protection of AI-generated content through copyright [12]. The application

of neighbouring rights can also avoid the difficult task of determining the attributes of AI-generated content, as neighbouring rights are designed for objects related to but not works. Meanwhile, the arrangement can solve the problem of greatly varying quality of AI-generated content, with its property rights and the period of protection could be reasonably stipulated in accordance with characteristics and the real needs analyzed case-by-case.

Neighboring rights were designed to balance interests. It was created late, after the invention of recording technology, radio technology, etc. It emerged in the context of the unemployment of a large number of performers caused by the advent of recording technology. Such background does have some similarities with the current onslaught of artificial intelligence on human creations. Nonetheless, this may work well for special types like AI-generated music, AI-generated news, AI-generated videos, etc, but may not be well adapted to all kinds of content. In order to fully regulate it under neighboring rights, a "data processor's right" could be added [13]. From a utilitarian perspective, the process of AI-generating can be viewed as a positive circular chain, where AI-generated products can only benefit society through dissemination. Users of AI models are critical to the commercialization of AI-generated content. As this arrangement confers the identity of right-holders to users, users are also subjects of liability when AI-generated content constitutes an infringement.

Besides, this article argues that the rights of AI-generated content should only contain property rights within the meaning of copyright law. Personal rights such as the right of attribution, the right of publication, the right of revision, etc, should not be included. The right of reproduction in property rights should also be limited, since the content itself comes from deep learning of various information and published works, thus other subjects should not be unduly restricted from re-learning and utilizing such results. In terms of the period of protection of the right, the term of protection of such data output should be shorter than the term of protection of property rights and neighboring rights in general. The reason for this is that AI-generated content is inextricably linked to the AI model itself, which may already gain protection through copyright or patent, therefore it is inappropriate to grant an excessive term of protection to AI-generated content.

5. Summary

As a new product brought about by technological development, AI is the crystallization of human wisdom, and its impact on society is evident to all, however, AI-generated content did bring new challenges to human creators and the copyright industry. The mindset of dealing with the child of new technology has a direct impact on the institutional choices made to regulate AI-generated content. It should be recognized that the original purpose of the broad copyright system was not only to encourage and facilitate the output of new works, but also to balance the interests of copyright industry according to the needs of society. Therefore, in dealing with AI and its products, it is appropriate to adopt the mentality of "blocking is better than dredging" and include AI-generated content into the broad copyright system to achieve free circulation of it, with giving more freedom of choice to the recipients of information, so that the public can participate in the training of AI machine learning. Meanwhile, this can also add more vigour and vitality to the fixed system of copyright.

The development of technology is progressing rapidly. Whether artificial intelligence can really become the subject of rights in the future is still hard to forecast at the moment, but if people can continue to adopt a tolerant and open mind in deciding the path of regulation, it will be a great help to the progress of technology.

References

- [1] Wang Qian. How can the issues raised by new technologies for the legal system be examined? --Using the impact of artificial intelligence on the intellectual property system as an example. *Oriental Law*, 2019, (5): 20-27
- [2] B.J. Copeland artificial intelligence in *Encyclopedia Britannica*, 20.03.2023, available at <https://www.britannica.com/technology/artificial-intelligence>, last time consulted on 25.02.2024.

- [3] WIPO, Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence, 21.05.2020, available at https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_2_ge_20/wipo_ip_ai_2_ge_20_1_rev.pdf, last time consulted on 25.02.2024, p.3
- [4] Yang Lihua. Exploring Copyright Issues in AI-generated content. *Modern Law Science*, 2021, (4): 102-114
- [5] Copyright Law of the People's Republic of China (2020)
- [6] Dornis, supra note 89, at 17 (“Copyright protection in virtually all jurisdictions depends on the quintessential element of human creativity. Both civil-law and common-law copyright have an anthropocentric foundation.”) [extracted from Redesigning Copyright Protection in the Era of Artificial Intelligence]
- [7] Caldwell, M. (2023). What Is an “Author”? --Copyright Authorship of AI Art through a Philosophical Lens. *Houston Law Review*, 61(2), 411-442.
- [8] U.S. Copyright Off., Copyright Rev. Bd., Opinion Letter on Second Request for Reconsideration for Refusal to Register a Recent Entrance to Paradise
- [9] Liu Xianquan, Hu Hejia. On the Criminal Responsibility of Intelligent Robots in the Age of Artificial Intelligence. *Law Science*, 2018, (1): 40-47.
- [10] Zhang Wenxian. *Jurisprudence*. [M]. Beijing: Higher Education Press, 1999
- [11] Oleksy, E. (2023). That Thing Ain’t Human: The Artificiality of “Human Authorship” and the Intelligence in Expanding Copyright Authorship to Fully Autonomous AI. *Cleveland State Law Review*, 72(1), 1-38.
- [12] Xiang Bo. On the Protection of Neighbouring Rights in Artificial Intelligence Generated Results. *Science-Technology & Publication*, 2020, (1): 70-75
- [13] Tao Qian. On the Protection of Artificial Intelligence Generated Results under Copyright Law-- Demonstration of Data Processor's Rights as Neighbouring Rights. *Law Science*, 2018, (4): 3-15