

# *The Study of Copyright Infringement Liability of Generative Artificial Intelligence*

Yue Yang<sup>1,a,\*</sup>

<sup>1</sup>James E. Beasley School of Law, Temple University, 1719 North Broad Street, Philadelphia, PA, USA

a. 1811431122@mail.sit.edu.cn

\*corresponding author

**Abstract:** As AI continues to produce content at an unprecedented rate, it is essential to establish a robust legal framework to protect the rights of creators and appropriately assign liability in cases of copyright infringement. Generative AI technology has the characteristics of high data demand, strong human-computer interaction, weak interpretability, and weak stability. The main civil subjects in generative AI services are generative AI service providers and users. In the pre-training stage, if generative AI uses copyrighted works without authorization, it should be recognized as copyright infringement, and the relevant infringement liability should be borne by generative AI service providers. In the content generation stage, the high similarity between AI-generated content and prior works can be attributed to various factors, including flaws in the generative model, intentional design choices, and user input and guidance. Both the generative AI service provider and the user bear direct tort liability for the infringement. However, the question of whether generative AI service providers also bear indirect tort liability should be explored, considering their role and similarities to traditional ISPs. Generative AI service providers should fulfill their obligations, take appropriate actions to prevent infringement, and report any illegal activities to the relevant authorities.

**Keywords:** Generative AI, ChatGPT, Copyright Infringement Liability

## 1. Introduction

In recent years, the field of Artificial Intelligence (AI) has witnessed remarkable advancements, particularly in the domains of deep learning, large-scale language models, and content generation. These breakthroughs have greatly enhanced the capacity of machines to simulate and accomplish human cognitive tasks. Based on this capability, AI can be broadly categorized into three types: narrow or weak AI, general or strong AI, and artificial superintelligence or super AI [1]. Weak AI has already proven its efficacy in various fields, such as speech recognition, image recognition, and recommendation systems. The industry perceives generative artificial intelligence, exemplified by ChatGPT, as a crucial milestone in the journey towards attaining strong AI. This progress in AI technology has also led to in-depth legal discussions and refinements in understanding generative artificial intelligence.

Generative AI refers to a type of artificial intelligence with the capacity to acquire knowledge and skills independently, achieving a level of comprehension akin to that of humans. Moreover,

generative AI has a certain degree of creativity, which allows it to create original content upon demand. This poses a significant challenge to the current copyright legal system. The existing framework primarily focuses on human-generated content, and it is not well-equipped to handle the unique characteristics and implications of AI-generated works. The growing prevalence of generative AI technology in the field of content creation means that there is a foreseeable increase in disputes related to the infringement of others' copyrights by generative AI. Given the rapid pace at which AI generates content, it is imperative to establish a robust legal framework for protecting creators' rights and appropriately assigning liability in cases of copyright infringement.

This article aims to contribute to the discourse on liability in the context of copyright infringement by generative AI within the current legal landscape. It is divided into three parts to comprehensively address the different facets of this complex issue. Part 1 provides an overview of generative AI technology and identifies the various civil subjects involved, encompassing users of generative AI and the service providers offering generative AI platforms. Part 2 delves into the infringement and liability related to generative AI that undergoes pre-training using prior works, exploring the legal implications of using copyrighted works as a foundation for training generative AI models and analyzes the potential liability of different actors in this process. Part 3 examines the infringement and liability issues that arise when generative AI generates content that is highly similar to prior works. This section considers the responsibilities of both generative AI service providers and users in different scenarios, emphasizing the need for clear guidelines and accountability.

Through an exploration of these critical facets of liability in the wake of copyright infringement by generative AI, we aim to illuminate the current legal challenges surrounding the use of AI in content creation and contribute to the development of a comprehensive and effective legal liability framework for generative AI in the future.

## **2. Generative Artificial Intelligence Technology and the Civil Subjects Involved**

### **2.1. Overview of Generative Artificial Intelligence Technology**

Generative artificial intelligence technology is a specialized branch of artificial intelligence that utilizes machine learning techniques such as generative adversarial networks, long short-term memory networks, and Transformer models to learn and understand the main features of training data [2]. With this understanding, it has the ability to generate completely new content that possesses similar characteristics to the training data. This technology stands in contrast to analytical artificial intelligence, which primarily focuses on analyzing and making predictions based on existing data.

The applications of generative artificial intelligence technology are vast and span across various fields. In the realm of image generation, generative models can be trained on a large corpus of image data and subsequently generate entirely new images that differ from the original dataset. Additionally, these models can even produce images with specific artistic styles based on given conditions. Within the domain of music, generative models can learn to generate novel musical compositions and create music that matches different emotional states as instructed.

In general, generative AI technology has the following characteristics. Firstly, generative AI requires a large amount of training data to learn and understand the semantics and context relationships of inputs. For example, for natural language processing tasks, large-scale text datasets are required to train models to provide accurate language generation and understanding. Secondly, generative artificial intelligence can automatically adjust the generated response content according to the instructions entered by the user, improve the quality of the output content through continuous dialogue, and interact with the user in an unprecedented way.

Despite these strengths, generative artificial intelligence technology also manifests certain limitations. It often lacks interpretability, as it learns and generates content through extensive training data, making its internal workings and decision-making processes complex and difficult to interpret. This limitation restricts its applications in fields that require high levels of model explanation, such as medical diagnosis and financial risk assessment. Additionally, generative AI is prone to instability, as minor disturbances in the input data can result in significant deviations in the generated results. This poses challenges for applying generative AI technology in areas that demand high stability, such as autonomous driving and safety monitoring.

## **2.2. Civil Subjects in Generative AI Services**

When exploring the intricate landscape of copyright infringement liability in the context of generative AI services, it is crucial to delve into the diverse set of stakeholders recognized by the current legal framework. These stakeholders encompass both generative AI service providers and generative AI service users. At present, the legal status of artificial intelligence is still controversial. The mainstream view is that artificial intelligence has not yet posed a subversive challenge to the traditional civil law subject theory, and it is not reasonable and necessary to define artificial intelligence as a civil subject [3].

### **2.2.1. Service Providers**

Generative AI service providers refer to organizations and individuals that use generative AI technology to provide content services such as text, pictures, audio, and video to the domestic public. Under the current business model, generative AI systems are often independently invested, designed, developed and operated by Internet technology companies. As a result, service providers extend beyond mere users of the technology; they include the investors who fund these endeavors, the designers who conceive the AI systems, the developers who craft the underlying technology, and the operators who manage and deliver these services to the public.

### **2.2.2. Service Users**

Generative AI service consumers refer to consumers, enterprises, and research institutions that use generative AI services to generate content. Consumers can leverage generative AI services to get services such as personalized recommendations, automatic content generation, or ideas. Businesses can leverage generative AI services to automate the generation of text, design, music, and more to increase productivity and creativity. Generative AI can also help research institutions improve research efficiency, accelerate scientific research and innovation, and provide personalized learning and educational experiences. The application of generative AI extends its transformative potential to a wide spectrum of users, each reaping benefits tailored to their distinct needs and objectives.

## **3. Infringement and liability of generative AI using prior works for pre-training**

### **3.1. Infringement of generative AI using prior works for pre-training**

Generative AI is a cutting-edge technology that involves training very large models with vast amounts of data. The key to this pre-training phase lies in the scale and quality of the data used. During pre-training, the model learns basic semantic and syntactic knowledge by analyzing a large corpus of unlabeled data. This corpus typically consists of text, images, or videos collected from the Internet. High-quality data is crucial for creating accurate language models, resulting in more natural and believable text. Consequently, it is common for this corpus to encompass copyrighted works, offering diverse and representative content for the model to glean from.

For instance, OpenAI has openly admitted to using books that contain long, continuous text in order to teach generative models how to process long text information effectively. However, the use of copyrighted content by OpenAI has sparked controversy. In June 2023, American writers Mona Awad and Paul Tremblay filed a class-action lawsuit against OpenAI, alleging that the company unlawfully used their work to train its ChatGPT model [4]. This case has garnered significant attention, and more than 15,000 writers have signed a letter demanding that companies involved in generative AI obtain consent from writers and provide financial compensation for using their copyrighted content in the training process [5].

The legal implications of using prior works in the pre-training phase are still being debated. One crucial aspect is whether pre-training constitutes an act of reproduction that infringes on the right of reproduction of the original works. Copying encompasses not only permanent forms, such as replication and printing, but also temporary forms that occur briefly during computer transmission. Different countries have varying perspectives on the legal nature of temporary reproduction, with China, for example, excluding it from reproduction rights. If pre-training is deemed a form of temporary reproduction, it may not be construed as an infringement of reproduction rights. Temporary copying occurs during the technical processes of reading, storing, and briefly reproducing network information or work content in computer memory or cache. It should be temporary, automatic, and automatically deleted after operation, without causing substantial harm to the copyright owner's interests. However, in the context of generative AI pre-training, where the model and its parameters remain stored on servers for an extended period using distributed file systems, such as Hadoop HDFS, this can be considered a permanent copy of the work. In the absence of the author's consent, this act indeed violates the right to reproduce the original work.

Another important consideration is whether pre-training infringes upon the right to adapt an earlier work. The right of adaptation refers to the ability to modify a work and create a new work with originality. Some argue that the "learning and training behavior" of generative AI models, in terms of ideology and style derived from others' works, falls outside the scope of existing copyright laws [6]. They claim that as long as the generative AI system's learning and training processes remain internal and do not involve the output of expressive content, it would not harm public communication and would not constitute copyright infringement [7]. However, this viewpoint lacks sufficient grounds to counter allegations of pre-training infringement. Generative AI pre-training is specifically designed for subsequent expressive use, and it should be viewed as different stages of one continuous process that may impede the existing or potential market for the original work. Generative AI models that learn and extract the core features, integrate them, and reuse them without permission from the original work are likely to infringe upon the right to adapt the original work.

### **3.2. Liability for infringement of generative AI using prior works for pre-training**

When generative AI uses prior works for pre-training and infringes the right of reproduction and adaptation of earlier works, it signifies that the AI system utilizes existing works as a foundation for generating new content without obtaining the consent of the original creators. This behavior directly violates the rights of the original creators who hold the copyright to those works. Consequently, generative AI service providers who knowingly engage in this practice can be held liable for engaging in direct copyright infringement.

It's important to note that generative AI service providers have the autonomy to make independent decisions regarding whether or not to infringe on others' works. They are not obligated or compelled to use copyrighted materials in their pre-training process, which means it is a conscious choice that they make. Furthermore, these service providers possess the necessary technical capabilities to avoid copyright infringement. They have access to various tools and

algorithms that can scan and analyze the collected data to identify any unauthorized works that may be present and eliminate or filter them out. After gathering the data, generative AI service providers typically go through a meticulous cleaning process. This involves the removal of any duplicate, noisy, or irrelevant data to ensure the accuracy and quality of the training data. During this step, automated tools and algorithms play a crucial role in detecting any copyrighted materials that should not be included in the pre-training process. Through the deployment of these automated filters, service providers have the means to warrant that the pre-training data they employ does not incorporate any resources that would infringe upon copyright.

However, it is essential to emphasize that these actions do not absolve generative AI service providers of liability in cases where copyright infringement does occur during the pre-training phase. Even with meticulous filtering and cleansing, it remains their responsibility to confirm that the data used for pre-training does not contain copyrighted materials without authorization. Failure to adhere to these responsibilities and to knowingly use copyrighted content without permission may expose generative AI service providers to legal consequences for their involvement in direct copyright infringement. Thus, the availability of technical tools and algorithms is not a safeguard against potential infringement, but rather a resource to facilitate compliance with copyright laws.

#### **4. Infringement of content generated by artificial intelligence that is highly similar to prior works and liability for them**

##### **4.1. Infringement of content generated by artificial intelligence that is highly similar to the prior work**

Generative AI, with its unique operating mechanism, undergoes a pre-training process where it thoroughly "chews" and "digests" the provided corpus. This process involves analyzing and understanding the structure, patterns, and nuances of the corpus, allowing the AI model to gain a deep knowledge and representation of the language present in the corpus. It learns to recognize and comprehend the various writing styles, sentence structures, and vocabulary used in the provided texts. Consequently, when generating content, the AI model does not simply replicate a specific work from the corpus. Instead, it leverages the acquired knowledge and generates original content by drawing upon the learned patterns and structures. It weaves together ideas and language in a manner that can resemble the style and characteristics of the corpus but presents a unique and novel composition.

Copyright law, aligning with the traditional copyright tradition, primarily safeguards the "external expression" of ideas rather than the ideas themselves. This means that while the AI-generated content may draw inspiration from the corpus, it does not infringe on the copyright of the original work as long as it does not reproduce its external expression verbatim.

However, if the AI-generated content bears significant resemblance to a prior work, it may be considered copyright infringement according to the "contact + substantial similarity" rule. This rule requires that there is both direct contact with the copyrighted work and a substantial similarity between the generated content and the earlier work. In cases where the copyright holder struggles to prove direct contact with their work, the presence of obvious similarities between the generated content and the earlier work can indicate unauthorized access and utilization of the work. These similarities may include shared characteristics, themes, plot structures, or even specific phrasings that are distinctive to the prior work. The style and manner in which the content is presented can also play a role in establishing the absence of originality and unauthorized usage.

Within the context of generative AI, copyright holders often encounter challenges in proving "contact + substantial similarity" due to the extensive use of prior works in machine learning processes. This is because the AI system, during its pre-training phase, digests a vast amount of



existing copyrighted material, blurring the direct connection between the generated content and any individual prior work [8]. To address this difficulty, judicial authorities can adopt a two-step approach. Initially, they can assess general copyright disputes through the "substantial similarity" criterion. If the generated content demonstrates substantial similarity to a prior work, it can serve as an indicator of potential contact between the generative AI system and prior works. This consideration of substantial similarity allows the court to make an initial determination and subsequently investigate the existence of actual contact with prior works brought about by the generative AI system.

#### **4.2. Liability for infringement of content generated by artificial intelligence that is highly similar to the prior work**

The high similarity between AI-generated content and prior works can be attributed to various factors. One possible reason is that the generative model used in AI generation may have flaws that result in the similarity. For example, the pre-trained model might not have sufficiently analyzed and learned from the corpus of prior works, leading to limited originality in the generated content. This could be due to limitations in data availability or inadequate training methodologies. Additionally, the generative model may intentionally be designed to correspond to a particular artist's style, resulting in a deliberate similarity to prior works. This attribute can be a desirable feature for specific applications, such as content creation that emulates renowned artists' styles or replicates a particular genre or historical period. In such cases, the resemblance to prior works is a purposeful design choice rather than a flaw in the AI model.

Another factor contributing to the high similarity is the influence of consumers of generative AI services. Users of these services may input material that infringes upon existing works themselves, knowingly or unknowingly. They may also deliberately guide the AI to generate content that closely resembles others' prior works through language training. This user input and guidance can also influence the similarity between AI-generated content and prior works.

According to Paragraph 1 of Article 1165 of China's Civil Code, an actor who through his fault infringes upon another person's civil-law rights and interests shall bear tort liability [9]. Therefore, in the two scenarios mentioned above, both the generative AI service provider and the user bear direct tort liability as the perpetrators of the infringement.

The assumption of direct tort liability is generally accepted and does not face much controversy. However, the focus of the discussion should shift towards determining whether generative AI service providers also bear indirect tort liability when their users engage in infringement.

In traditional online copyright dispute cases, China has established rules for determining the indirect infringement liability of internet service providers (ISPs) based on their involvement in aiding infringement. Two main rules, known as the "notice takedown rule" and the "know rule," govern the liability for online infringement. Under the notice takedown rule, if an ISP fails to take timely measures after receiving a notice from the right holder, they bear joint and several liability with the network user for the expanded part of the damage. The know rule states that network service providers who know or should know that their services are being used for infringement and fail to take necessary measures are also liable for joint liability with the network users. Although generative AI service providers differ in technical foundation, data requirements, and business models from traditional ISPs, they can be considered a special type of ISP in cases where their services are used for infringement. Generative AI service providers might face challenges in controlling the generated content when users input infringing materials or guide the AI to produce infringing content through language training. In such cases, the role of generative AI service providers is similar to that of ISPs. If generative AI service providers are aware of infringing

activities and do not take appropriate actions, such as removing the infringing content or disabling user access, they can be held jointly liable for the damages caused.

To fulfill their obligations, generative AI service providers should promptly take measures like stopping the generation, transmission, and elimination of illegal content, as specified in Article 14 of the Interim Measures for the Administration of Generative Artificial Intelligence Services [10]. They should also employ measures such as model optimization training to rectify the situation and report to the relevant authorities. Furthermore, if generative AI service providers uncover instances of users engaging in unlawful activities, they should implement corrective measures, including warnings, and promptly notify the authorities regarding the infringing conduct.

## 5. Conclusion

Generative AI technology has the characteristics of high data demand, strong human-computer interaction, weak interpretability, and weak stability. The main civil subjects in generative AI services are generative AI service providers and users. In the pre-training stage, if generative AI uses copyrighted works without authorization, it should be recognized as copyright infringement, and the relevant infringement liability should be borne by generative AI service providers. In the content generation stage, the high similarity between AI-generated content and prior works can be attributed to various factors, including flaws in the generative model, intentional design choices, and user input and guidance. Both the generative AI service provider and the user bear direct tort liability for the infringement.

However, delving deeper, it is imperative to explore the potential realm of indirect tort liability for generative AI service providers, given their roles and parallels to traditional ISPs. While generative AI service providers operate on distinct technical foundations and within unique business models, their involvement in acts of infringement necessitates a thorough examination of their indirect liability. By conforming to established rules for ISPs' indirect infringement liability, generative AI service providers can foster an environment of accountability and mitigate their legal risks.

Recognizing the issue of copyright infringement that arises from the use of generative AI services doesn't imply opposition to AI as a whole. Instead, it involves adapting the legal aspects of these services to strike a fair balance between different interests. In the era of AI, the existing rules governing intellectual property face new challenges, particularly due to the rapid progress of generative AI. To tackle this issue effectively, it is crucial to establish a copyright framework that is not only scientifically robust but also implemented at an institutional level. Achieving this objective necessitates thorough observation, meaningful discussions, and legislative actions to ensure a sustained equilibrium of values and interests in the long run.

## References

- [1] Fjelland, R. (2020). *Why general artificial intelligence will not be realized. Humanities and Social Sciences Communications*, 7(1), 1-9.
- [2] Sarker, I. H. (2021). *Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Computer Science*, 2(6), 420.
- [3] Liming Wang. (2018). *New Challenges to Civil Jurisprudence in the Age of Artificial Intelligence. Eastern Jurisprudence*, 3.
- [4] European Innovation Council and SMEs Executive Agency. (2023). *OpenAI sued for copyright infringement. Retrieved from [https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/openai-sued-copyright-infringement-lana-del-rey-settling-plagiarism-dispute-2023-07-04\\_en](https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/openai-sued-copyright-infringement-lana-del-rey-settling-plagiarism-dispute-2023-07-04_en)*
- [5] Authors Guild. (2023). *More than 15,000 Authors Sign Authors Guild Letter Calling on AI Industry Leaders to Protect Writers. Retrieved from <https://authorsguild.org/news/thousands-sign-authors-guild-letter-calling-on-ai-industry-leaders-to-protect-writers/>*

- [6] *Lixian Cong, Yonglin Li. Copyright risks and governance of content generated by chatbots - from the perspective of ChatGPT application scenarios. (2023), China Publishing, 5, 16-21.*
- [7] *An Li. Analysis of copyright law for machine learning works - non-work use, fair use and infringement use. (2020), Electronic Intellectual Property, 6, 60-70.*
- [8] *Handong Wu, ““Substantial Similarity + Access” As the Rules for Deciding Infringement. (2015), Law Science, 8, 63-72.*
- [9] *Civil Code of the People's Republic of China (promulgated on May 28, 2020, and came into effect on January 1, 2021).*
- [10] *Interim Measures for the Administration of Generative Artificial Intelligence Services (promulgated on 13 July, 2023, and came into effect on 15 August, 2023).*