Proceedings of the 3rd International Conference on International Law and Legal Policy
DOI: 10.54254/2753-7048/57/20240124

Legal and Economic Analysis of Fair Use in the Process of Artificial Intelligence Creation in Data-Driven Era

Jiehui Mai[1,a,*]

[1]*International College Beijing, China Agricultural University, Beijing, 100000, China*
a. 13924260801@163.com
**corresponding author*

Abstract: In the era of digitalization, AI-generated content has emerged as a pivotal component within the creative industries, thereby instigating extensive deliberations concerning copyright protection and fair use. This study aims to scrutinize and analyze the equitable use of AI-generated content from both a legal and economic standpoint. This paper does a thorough review of the literature to introduce basic ideas about AI content creation and fair use. It then divides existing research into four separate areas: technical analysis, legal case analysis, economic impact analysis, and policy adaptation research. Despite notable advancements in both technological capabilities and legislative frameworks surrounding AI-generated content, there remains a dearth of comprehensive understanding, particularly with regard to its economic implications and policy implementation. Consequently, this study posits a significant research question on how to strike a balance between the fair usage of AI-generated content and upholding copyright protection in order to foster sustainable development within the creative industries. The findings presented herein offer novel perspectives and insights that can inform future policymaking endeavors as well as academic investigations.

Keywords: fair use, AI-generated content, intellectual property, law and economics.

1. Introduction

With the rapid advancement of artificial intelligence (AI) technology, AI-generated content has gradually emerged as a pivotal component within the cultural and economic domains. This technological transformation has not only presented unprecedented opportunities for the creative industries but also ignited extensive discourse surrounding copyright protection and fair use. In recent years, scholars have conducted comprehensive research on these matters. For instance, Henderson et al. explored the historical context and future prospects of AI and fair use, highlighting how AI technology's development has posed challenges to traditional copyright protection mechanisms while proposing various technical solutions [1]. The authors, Klosek and Blumenthal, talked about whether teaching generative AI models to use copyrighted works is actually fair use. They said that figuring this out would have big effects on future progress in AI technology [2].

The core concepts addressed in this paper encompass "AI-generated content" and "fair use." Artificial intelligence systems acquire and produce texts, images, music, and other works through algorithms based on input data, referred to as AI-generated content [3]. Fair use is a legal principle that permits the utilization of copyrighted material under specific conditions without payment or

DOI: 10.54254/2753-7048/57/20240124

obtaining permission [4]. More specifically, the determination of fair use relies on four factors: the purpose and nature of the usage, the characteristics of the copyrighted work, the quantity and significance of portions used within it as a whole, and its impact on potential market value [5-6].

Utilizing both literature review and case analysis, this study carefully looks at how reasonable usage can be used with AI-generated content in the fields of law and economics. This research not only contributes to an all-encompassing understanding of legal criteria and economic effects concerning fair use with regards to AI-generated content, but also provides empirical support as well as a theoretical foundation for future policy-making endeavors. Through an in-depth analysis of specific cases, we can identify both adaptability and limitations within laws when responding to challenges posed by new technologies; thus, providing valuable insights for enhancing and refining relevant legal frameworks.

2. Literature Review

The existing research is mainly divided into four categories according to the theme, including technical analysis, legal case analysis, economic impact analysis, and policy adaptability research.

2.1. Research Classification

The first is a technical analysis of fair use. This type of research focuses on the reasonable use of AI-generated content from a technical perspective. Henderson et al. proposed that by installing a fair use filter, it is possible to identify whether AI-generated content is beyond the scope of fair use [1]. Klosek and Blumentha explore ways to train models to ensure that their usage behavior meets fair use criteria [2].

The second category is the analysis of legal cases. Many studies have explored the fair use of AI-generated content by analyzing relevant legal precedents. In the Kelly v. Arriba Soft case, photographer Leslie Kelly sued search engine operator Arriba Soft, claiming that it copied and displayed thumbnails of her photos without permission. However, the court determined that this use was transformative, as the thumbnails' sole purpose was to serve as a search function, not to showcase artistic value. This new use of the original work adhered to the fair use standard [7]. Oracle sued Google in the Oracle v. Google case for using Java API code in the Android operating system, providing a fresh perspective. In the end, the court determined that Google's use was transformative, as it aimed to ease the transition of software developers from the Java platform to the Android platform, instead of directly replicating the functionality of the code [8].

The third category is economic impact analysis. This type of research explores the economic impact of the rational use of AI-generated content. Foley & Lardner LLP points out that the "black box problem" in fair use judgments can have a significant impact on the market [9]. In terms of improving consumer welfare and driving innovation, Creative Commons proposes that rational use can promote the free flow of information, thus promoting the development of innovation and creativity [10]. Despite some theoretical analysis and case studies, there is still relatively little empirical research on the economic impact of fair use. For example, Klosek and Blumenthal pointed out that the existing economic analysis is mostly based on theoretical models and assumptions and lacks large-scale data support and empirical verification [11]. In addition, existing research mainly focuses on the legal and market environment in the United States and lacks in-depth analysis of the impact of international markets and transnational economies. Copyright laws and market structures differ significantly across countries, and this difference could have important implications for the economic impact of AI-generated content. More studies on international markets and transnational cases are needed to fully understand the economic effects of fair use [12].

DOI: 10.54254/2753-7048/57/20240124

The fourth category is policy and regulatory adaptability. Research on policy and regulatory adaptability focuses on how to adapt existing copyright laws to the development of AI technology. With the popularity of AI-generated content, traditional copyright protection mechanisms are facing new challenges. Many scholars and policymakers are exploring ways to modify and improve copyright law to balance innovation with copyright protection. The limitation of the current copyright law is that the specific definition of fair use is vague, especially in the era of more frequent application of AI technology, and it cannot deal with new problems such as infringement judgments brought by AI creation and regulating the scope of fair use of AI creation. Some scholars have proposed specific policy adjustments to adapt to the development of AI technology. For example, Creative Commons recommends that copyright ownership of AI-generated content be clearly defined in copyright law, while introducing new fair use standards to accommodate the characteristics of AI-generated content [13]. To ensure a balance between copyright protection and innovative development, Foley & Lardner LLP recommends developing a legal framework specifically for AI-generated content [14]. The recommendations aim to provide a clearer legal basis for the legitimate use of AI-generated content, thereby reducing copyright disputes and promoting technological progress and market development. In the context of globalization, copyright protection of AI-generated content not only needs to consider domestic laws, but also needs to pay attention to the coordination of international laws. Different countries' copyright laws may significantly differ in terms of fair use standards and copyright attribution issues, necessitating harmonization in policy development.

2.2. Research Questions

In light of the aforementioned shortcomings, this paper intends to investigate the fair use of AI-generated content from the perspective of law and economics. The main research question is: how to balance the fair use of AI-generated content and copyright protection in the era of data, so as to promote the sustainable development of creative industries. This major issue can be further discussed and solved from the following three aspects, including the economic effect of the rational use of AI-generated content on the creative industry; the application and limitations of existing copyright laws to the fair use of AI-generated content; and how to optimize the rational use of AI-generated content through legal and policy means to achieve a balance between innovation and protection.

3. Case study

In order to explore the application and limitations of current copyright laws to the fair use of AI-generated content, the article analyzes specific legal cases and provides elements to guide its practical application in the context of a better understanding of the topic. The paper analyzes two typical cases, Authors Guild v. Google and Oracle v. Google, and explores the impact of legal decisions in these cases on the fair use of AI-generated content.

3.1. The case of Authors Guild v. Google

Authors Guild v. Google, Inc. is a lawsuit about the Google Books Project. By scanning a large number of books and providing an online search function, Google enables users to find snippets of books that contain specific search terms. The project has sparked widespread discussion about copyright infringement and fair use. The plaintiffs argued that Google's actions violated its copyright, while Google argued that its actions were fair use. In 2015, the U.S. Court of Appeals for the Second Circuit ruled that Google's actions constituted fair use. The court found that Google's actions were transformative because the purpose was not to copy or distribute entire books, but to enhance public access to information by providing search capabilities. The Court emphasized that such transformative uses contribute to knowledge dissemination and public interest. This judgment

DOI: 10.54254/2753-7048/57/20240124

demonstrates that in the context of AI technology, fair use can cover new types of technology and service models.

In the context of AI-generated content, this case has important implications for the application of current copyright law. In the Google Books Project case, the court recognized the transformative use of technological progress as a basis for fair use. This suggests that AI-generated content that demonstrates a similar transformative nature and makes a clear contribution to the public good has the potential to be considered fair use under copyright law. In terms of economic benefits, the Google Books project significantly lowers the threshold of access to information by providing free search services, so that more people can access the knowledge in these books. The diffusion of knowledge is non-exclusive and non-competitive, which accords with the theory of public goods. Moreover, through rational use, the Google Books project not only enhances the public's ability to obtain information, but also promotes the dissemination of knowledge and the promotion of public interests [13]. Simultaneously, the appropriate utilization of this project yields additional benefits. First of all, the rational use of the Google Books project brings significant positive externalities. The free circulation of knowledge promotes academic research and innovation, reduces the cost of academic research and creation, and improves the overall knowledge level and creativity of society. Although fair use has positive effects, it may have a negative impact on the potential market income of the original copyright holder. However, the court's finding that Google's use was transformative mitigated this concern about negative externalities. [9] The court's judgment of fair use provides legal protection for innovative activities and encourages more secondary creations and technological innovations based on existing knowledge.

3.2. The case of Oracle v. Google

Oracle America, Inc. v. Google, Inc. The case involves Google's use of Java APIs in its Android operating system. Oracle alleges that Google infringed on its copyright by making an unauthorized copy of the structure and organization of the Java API. Google argues that its use is fair because the purpose is to create a compatible open platform that facilitates developer migration. In 2021, the U.S. Supreme Court ruled that Google's actions constituted fair use. The court emphasized that Google's use was transformative, aimed at facilitating the migration of developers from the Java platform to the Android platform, rather than a mere copy. The court also noted that the API was used in a way that had a relatively limited impact on the market, and therefore upheld the fair use finding. In the context of AI-generated content, this case shows that if AI-generated content is used in a way that demonstrates a transformational nature and has a limited negative impact on the market for the original work, then it is more likely to be deemed fair use. This case highlights the importance of transformational and market influence in fair use decisions [8]. Google uses Java APIs to create a compatible open platform that facilitates the sharing of knowledge and technology. This sharing possesses the qualities of a public good, allowing for widespread use without reducing its accessibility. Moreover, open platforms promote greater developer participation, drive technological progress and innovation, and bring significant public benefits. Google facilitates the migration of developers from the Java platform to the Android platform, enhances the vitality of the development ecosystem, and provides positive externalities. The court found that Google's use had a limited impact on the market, thereby deeming this negative externality acceptable, despite Oracle's claim of infringement on its market interests.

3.3. Comprehensive Analysis

Together, these two cases underscore the centrality of transformative use and the public interest in fair use decisions. Applications of AI-generated content are more likely to receive fair use recognition

DOI: 10.54254/2753-7048/57/20240124

under copyright law if they demonstrate significant transformation, promote knowledge dissemination, technological innovation, and the public interest. For example, the use of AI-generated content for teaching and research in education and research may be seen as a transformative use because it is intended to facilitate the diffusion and application of knowledge. At the same time, these cases also remind us to consider the impact on the market when judging fair use. If AI-generated content is used in a way that has little negative impact on the market for the original work, or even opens new market opportunities, the chances of it being deemed fair use are higher. For instance, using AI-generated content to create new works or services can promote market diversification and innovation without compromising the original work's market value. From a legal and policy perspective, the current copyright law needs to further clarify and refine the standards of fair use to adapt to the rapid development of AI technology. This includes introducing new fair use standards, considering new uses brought about by technological change, and enhancing international legal coordination and cooperation in policymaking.

4. Discussion

Current research and discussion on AI creation and copyright issues remain inadequate. The current discourse primarily concentrates on technical aspects and fails to provide a thorough examination of the legal and economic ramifications. As technology evolves, so does the legal landscape. More research on the latest jurisprudence is necessary to supplement and update the existing analysis, as some earlier jurisprudence may not have fully considered the latest technological advances and the new issues they raise. In the study of economic impact, there are still significant deficiencies in empirical data, diversity of economic effects, and transnational impact analysis. These deficiencies need to be remedied through more integrated and systematic research methods in order to provide more comprehensive and accurate guidance for future policymaking and academic research. In the context of globalization, copyright protection of AI-generated content not only needs to consider domestic laws, but also needs to pay attention to the coordination of international laws. Copyright laws in different countries may have significantly different provisions on fair use standards and copyright attribution issues, and may need to be coordinated in policy development to meet the needs of AI development in the context of increasing globalization.

In terms of transformative and public interest, AI-generated content is more likely to be recognized as fair use under copyright law if it can demonstrate a significant transformative effect and make a clear contribution to the public interest. Furthermore, when it comes to market impact, fair use must foster innovation while reducing any detrimental effects on the original work's market value. This approach could potentially create new market opportunities and increase the likelihood of fair use recognition. The current copyright law needs to further clarify and refine the standards of fair use to adapt to the rapid development of AI technology. This includes introducing new fair use standards, taking into account new uses brought about by technological change, and enhancing international legal coordination and cooperation in policy development. To promote the sustainable development of creative industries, policymakers should refer to existing cases, adjust copyright laws to provide a clearer legal basis for fair use of AI-generated content, promote technological progress and market development, and find a balance between AI technology and copyright protection.

5. Conclusion

Through a literature review and case analysis, this paper systematically analyzes the specific application and impact of fair use of AI-generated content at the legal and economic levels, focusing on transformative use and public interest, market impact, adaptability of existing copyright laws, and policy recommendations. By analyzing two typical cases, Authors Guild v. Google and Oracle v.

DOI: 10.54254/2753-7048/57/20240124

Google, transformative use and public interest are central factors in fair use decisions. Google's transformative and public good contributions to book search and the use of Java APIs became important grounds for the court to uphold its fair use. Although this paper provides a more comprehensive analysis of the rational use of AI-generated content, there are still shortcomings, such as insufficient data support, a lack of international perspective, and the dynamic nature of technological change. To address these deficiencies, future research should employ empirical research, international comparative research, and dynamic tracking research. The research in this paper not only enriches the research perspective of fair use in theory, but also provides empirical support and theoretical basis for legal policy makers and provides reference for the clarification and refinement of fair use criteria through case analysis, which helps to find a balance between AI technology and copyright protection, so as to promote the sustainable development of creative industries.

References

[1] *Henderson, P., Li, X., Jurafsky, D., Hashimoto, T., Lemley, M. A., Liang, P. (2023) Foundation Models and Fair Use. arXiv: 2303.15715. https://arxiv.org/abs/2303.15715*

[2] *Klosek, K., Blumenthal, M. S. (2024) Training Generative AI Models on Copyrighted Works Is Fair Use. Association of Research Libraries. https://www.arl.org/blog/training-generative-ai-models-on-copyrighted-works-is-fair-use/*

[3] *Wu, J. Y., Gan, W. S., Chen, Z.F., Wan, S. C., Lin, H. (2023) AI-Generated Content (AIGC): A Survey. arXiv:2304.06632v1. https://arxiv.org/pdf/2304.06632*

[4] *Georgetown University Library. (n.d.) Fair Use. https://library.georgetown.edu/copyright/fair-use*

[5] *Copyrightlaws. (2022) A Simple Guide to Fair Use: How to Apply Fair Use to Your Copyright Scenarios. https://www.copyrightlaws.com/simple-guide-fair-use/*

[6] *Digital Media Law. (n.d.) Fair use. https://www.dmlp.org/legal-guide/fair-use*

[7] *Kelly v. Arriba Soft Corp. (2003) Stanford Library. https://fairuse.stanford.edu/case/kelly-v-arriba-soft-corp/*

[8] *GOOGLE LLC v. ORACLE AMERICA, INC. (2020) Supreme Court Of The United States. https://www.supremecourt.gov/opinions/20pdf/18-956_d18f.pdf*

[9] *Ahmad, A. A., Gross, A. M. (2024) Generative AI Systems Tee Up Fair Use Fight. Foley. https://www.foley.com/insights/publications/2024/02/generative-ai-systems-fair-use/*

[10] *Wolfson, S. (2023) Fair Use: Training Generative Ai. Creative Commons. https://creativecommons.org/2023/02/17/fair-use-training-generative-ai/*

[11] *McDuff, D., Korjakow, T., Cambo, S., Benjamin, J. J., Lee, J., Jernite, Y., Ferrandis, C. M., Gokaslan, A., Tarkowski, A., Lindley, J., Cooper, A. F., Contractor D. (2024) On the Standardization of Behavioral Use Clauses and Their Adoption for Responsible Licensing of AI. arXiv:2402.05979v1*

[12] *Yang, A., Zhang, H. Y. (2024) Generative AI and Copyright: A Dynamic Perspective. https://ar5iv.labs.arxiv.org/html/2402.17801*

[13] *Walsh, K. (2023) Understanding Cc Licenses And Generative Ai. Creative Commons. https://creativecommons.org/2023/08/18/understanding-cc-licenses-and-generative-ai/*

[14] *Greene, J. H. (2024) Tailoring Compliance Before AI Walks The Runway. Foley. https://www.foley.com/insights/publications/2024/07/tailoring-compliance-before-ai-walks-runway/*