Proceedings of CONF-MPCS 2025 Symposium: Mastering Optimization: Strategies for Maximum Efficiency
DOI: 10.54254/2753-8818/2025.CH22295

Fair Use in AI Data Training: Judgment Criteria and Balancing Mechanisms

Yaorunyang Ding

Institute of Foreign Languages, Heilongjiang University, Harbin, China
20220377@s.hlju.edu.cn

Abstract: Against the backdrop of the rapid development of artificial intelligence technology, the issue of fair use in AI data training has sparked widespread attention and discussion. Large model training relies on massive amounts of data, whose technical characteristics differ significantly from how traditional works are used. However, market failures are prevalent, with high licensing costs and difficulties in obtaining rights holders' permissions hindering the effective operation of traditional authorization mechanisms. Therefore, the use of works in large-scale model training should be considered as fair use, since it has a limited impact on the legitimate rights of copyright holders while offering significant social and public benefits. In addition, within the framework of copyright law, it is essential to clarify the rules and criteria for fair use in machine learning and to define the obligations and responsibilities of AI data trainers. This will help balance the interests of copyright holders, society, and data trainers, thereby promoting the healthy and sustainable development of AI technology and the adaptive evolution of copyright law.

Keywords: AI data training, Fair use, Interest balancing, Machine learning types

1. Introduction

In today's digital era, the rapid advancement of AI technology is profoundly transforming various fields of human society. With its powerful innovation and practicality, AI has injected new impetus into social progress and prosperity. However, with the wide application of AI technology, especially the use of works in data training, the question of whether such use constitutes fair use has triggered intense debate in both theoretical and practical circles. This issue not only involves the balance between technological innovation and intellectual property protection but also relates to the adaptability and flexibility of the legal system in the new era.

At present, the issue of fair use in AI data training has been extensively studied internationally. The EU provides a legal basis for data mining activities through “the Text and Data Mining Exception” in the Digital Single Market Directive. The United States, through judicial practice, has gradually explored the boundaries of fair use, from the four-factor test to the expansion of transformative use. China's fair use system, at the same time, is closer to the civil law tradition of " Limitations and Exceptions to Copyright" which exhaustively lists the circumstances under which fair use can be recognized, reflecting a restriction and exception to copyright. In China, there have already been cases where copyright holders have sued AI painting software companies for using their works to train models without permission. These cases indicate that there is an urgent need for clear legal norms regarding fair use in data training.

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.CH22295

The foundation of the fair use system lies in fairness and justice, ensuring that while protecting the rights of creators, it does not harm the public interest. Whether an AI data training behavior constitutes fair use needs to be comprehensively considered in terms of the principles of fair use, transformative use, and legal regulations. Therefore, this paper will delve into the issue of fair use in AI data training, clarify the rules for fair use in AI data training, and seek a balance between technological innovation and intellectual property protection.

2. The Disputes Over Fair Use in Artificial Data Training

With the rapid development of AI technology, Artificial Intelligence Generated Content (AIGC) has become indispensable. However, as AI flourishes, the question of whether the use of works in AI data training constitutes fair use has sparked intense debate in both theoretical and practical circles. While AI-generated content often possesses innovation and practicality, contributing significantly to social progress and prosperity, its application also carries potential risks—namely, the possibility of impacting the market value of original works, thus triggering a series of disputes.

2.1. AI data training behavior constitutes reasonable use

As an important part of the legal system, the fair use system can effectively reduce transaction costs. It aims to balance the relationship between copyright protection and technological innovation to promote scientific and technological progress and safeguard social public interests.

From a technical perspective, AI training data plays a crucial role in driving social innovation and technological breakthroughs, and the fair use system provides a legal basis for this technology. For example, the "text and data mining exception" proposed in the Copyright Directive on the Digital Single Market Directive of the European Union clearly states that fair use provisions can provide legitimate exemptions for data mining activities in specific scenarios, highlighting the positive role of this mechanism in the field of technological development. Most members of the American academic community have shown openness and support for the rational application of machine learning [1]. indicating the positive role of fair use in promoting scientific research and technological innovation. For instance, Thomas Dietterich in the U.S. has publicly supported the fair use of machine learning. Domestic scholars argue for the establishment of a more flexible system to encourage the application of fair use in AI training data behavior. Some scholars have proposed a copyright theory based on the reader's perspective, offering two strategies for addressing the issue of whether AI can be a rights holder. This approach aims to balance technological innovation with intellectual property protection.[2].

When data training activities are carried out within strict limits and under supervision, ensuring that they do not significantly interfere with the normal use of original works or unreasonably harm the legitimate rights of copyright holders, they can be considered as fair use, aligning with the fairness and public interest goals pursued by copyright law [3].

2.2. AI data training behavior does not constitute fair use

The fair use system is one of the statutory systems that limit the scope of copyright. According to the fair use system, under specific circumstances, others can use works without the permission of the rights holder or payment of compensation [4]. However, due to the complexity of data sources, this process may involve infringing on the original author's rights of reproduction, adaptation, and the right of network dissemination. Moreover, the conflict between the high cost of the traditional authorization model and the limitations of technical means will have a certain impact on the industrial efficiency of artificial intelligence and the fairness of the copyright market.

DOI: 10.54254/2753-8818/2025.CH22295

Some scholars question whether AI data training should be considered fair use. First, AIGC may not meet the second or third conditions of the "three-step test," thus failing to comply with the principles of fair use. Professor Wu Handong analyzed that the demand for large amounts of data in AI technology has increased, and the traditional method of obtaining individual authorizations is not only technically and practically difficult but also prohibitively expensive. Scholar Eric Sunray similarly states that defining the use of works in AI data training as fair use overlooks the importance of these foundational works in generating outputs [5]. Additionally, the charging pressures under the statutory licensing system pose obstacles to the growth of the AI industry, distinctly impacting the existing copyright legal framework [6].

In summary, the application of AIGC should ensure that the copyright of original works is respected and prevent undue impacts on their market value. At the same time, some scholars have examined the relationship between the exercise and limitation of copyright through the equity theory and advocated that compared with the fair use system, the statutory license is a more ideal strategy for promoting the enhancement of knowledge value. It ensures the effective circulation and use of works while giving due respect to copyright holders. Such an institutional design aims to balance the benefits of work dissemination with the protection of the legitimate rights of copyright holders.

3. Necessity analysis of machine learning behavior constituting fair use

Copyright law need to balance the interests of copyright holders and the public, incentivizing the creation of works while preventing knowledge monopolies to promote cultural and technological innovation. Machine learning, as a core artificial intelligence technology, necessitates fair use not only for the healthy development of the technology but also for the application and balance of copyright law. In this context, the use of AI is non-traditional, requiring the protection of the original work's market while promoting the exchange, dissemination, and development of technological and cultural innovations. Achieving these goals requires a well-functioning copyright system to coordinate—this involves the fairness of transaction costs and the transparency of information. The existing copyright system has exposed many problems in practice, such as the lack of effective market regulation, difficulties in the licensing process, and high transaction costs, even leading to "market failures." Faced with these challenges, the introduction of the fair use system becomes a necessary solution, balancing the interests of copyright holders, businesses, and other stakeholders. In terms of enterprises and the market, this system reduces the transaction costs of works and realizes the optimal allocation of resources, better expanding the public interest.

3.1. Consideration of the balance of interests

The spirit of civil law, the requirements of social morality, and the principles of human rights and public interest all emphasize the importance of interest balancing [7]. This concept is equally reflected in machine learning technology. The impact of machine learning technology on the current copyright system lies in its redistribution of existing interest models, creating tension between the strong demand for copyright protection and the driving force for development. In terms of interest balancing, it involves weighing the interests of copyright holders, machine learning technology developers, and the public interest. In the relationship between technology and copyright protection, it is necessary to consider the tolerance of the copyright legal system for technological innovation. Excessively strict or lenient copyright protection policies are insufficient to address current challenges, and a balance must be sought that maintains strict copyright protection while avoiding excessive restrictions on the technological environment. Achieving this goal requires moderate adjustments to copyright laws to ease the tension between innovation and copyright protection, ensuring the balance of interests among all stakeholders.

DOI: 10.54254/2753-8818/2025.CH22295

From the perspective of copyright holders, the fundamental reason for opposing the use of their works in AI training without permission lies in the negative economic forecast for the future. For authors, the key to motivating continued creation is ensuring they receive appropriate material and spiritual rewards. AI training may reduce their expected benefits. If their works are used for large-scale data training without their consent, this could directly compress and reduce the economic returns they should obtain through licensing. However, rights protection should remain within reasonable limits. Overly broad protection measures may hinder the generation of more high-quality works, harming cultural diversity and the welfare of society as a whole. At the same time, overly strict copyright protection policies may increase the technical threshold and cost of machine learning, limiting the pace of innovation. Therefore, finding a mechanism for interest balancing becomes crucial.

Neither overly strict nor overly lenient copyright protection strategies can effectively solve the problem, so we need to explore a new path, seeking a balance between strict legal restrictions and lenient policies. Fundamentally, the solution involves adjusting and improving existing copyright laws to ease the conflict between technological innovation and copyright protection, actively responding to the challenges of the new technological era, and ensuring the harmonious coexistence of multiple interests.

3.2. The Requirement for Maximizing Interests

Machine learning typically relies on large amounts of data for model training, which can easily lead to potential copyright disputes and may also result in problems such as biased and unfair algorithmic decisions. Therefore, when using machine learning technology, it is essential to balance the dual needs of copyright protection and technological development. The fair use system provides a theoretical basis and practical path for resolving these conflicts. This system allows the use of works in certain ways without the permission of the copyright holder or payment of compensation, aiming to promote activities such as knowledge dissemination, education, and research that serve the public interest. The following we will explore the actual impact of AI training using works on creators' rights and the rationality of companies' expected benefits from the perspective of maximizing interests.

From the perspective of economic benefits and social utility, examining the relationship between AI training and the use of works reveals that AI data training does not substantially harm the original market of works or the rights of creators at the individual level. The specific manifestations of emerging application fields and their impact on works are often unforeseen at the time of creation. Copyright holders can hardly predict how AI will form new application scenarios through training, thereby expanding the potential market space for their works. This use of machine training goes beyond the capabilities and original intentions of ordinary authors in creating and disseminating works, exceeding the scope of what could be anticipated during the production and publication process. Therefore, it does not affect the normal use of the works by the authors.

The development of AI technology has brought new growth points to the cultural industry, such as personalized recommendation services and intelligent creation tools, which have actually expanded the audience for works and created new business opportunities. In the current context of rapid technological development represented by machine learning, companies using large amounts of works for technological research and development is an important means of driving technological progress and has become a key strategy for advancing technology and promoting economic development. However, according to existing regulations, unauthorized use of copyrighted works may lead to liability for damages [8]. The vast library of works and the high cost of compensation place a heavy burden on companies and may even lead to lengthy litigation processes and significant social costs. Given that machine learning technology typically relies on large amounts of data resources, simplifying the licensing process and payment mechanisms is crucial. This can be

DOI: 10.54254/2753-8818/2025.CH22295

achieved through exploring licensing agreements and collective management organizations to build a more flexible and efficient copyright management system, ensuring that machine learning developers can access and use works in an efficient and economical manner. This approach not only motivates creators and ensures the legitimate income of copyright holders but also promotes technological progress and development, enhancing corporate economic benefits and driving the comprehensive prosperity of technology and culture.

3.3. The Lack of Rationality in Copyright Holders' Suppression of Data Training

Although individual creators' economic rights through copyright law and the public's broad right to use cultural works are the core values of this law, when constructing a legal framework, the focus should be on higher-level goals—promoting cultural prosperity and social progress and safeguarding the public interest. For AI that relies on large amounts of works for training to optimize algorithms, high costs greatly limit further technological development and improvement. Without sufficient data resources, the possibility of AI producing high-quality content diminishes. In the current context, continuing to insist on strict copyright control over works may instead inhibit the development of public knowledge.

In fact, regarding the potential market of works, the market driven by AI deep learning is different from the traditional market. Authors cannot foresee these applications or prompt changes in the usage patterns of their works, so AI data training does not actually hinder the basic uses of works. In the Hathi Trust and Google Books cases [9], judges used the concept of "transformative use" to determine whether the use constituted fair use, broadening the scope of fair use beyond non-commercial entities. This approach indicates that AI data training aimed at promoting technological development does not unduly infringe on the legitimate rights of copyright holders.

In the process of AI data training, some scholars believe that the use of works in generative AI data training should be defined as "non-expressive use" and therefore should not fall within the scope of copyright. They believe that adopting a strategy of explicit exemption protection for training data in copyright law, compared to the approach of determining fair use through post-use review, is more conducive to maintaining overall interest balance and stimulating innovation.

4. Judgments Criteria for Fair Use in AI Data Training

Determining whether AI data training constitutes fair use requires a comprehensive consideration of the nature and purpose of the use, the potential impact on the original work, and an in-depth analysis of the purpose and nature of the use. China has been continuously adjusting and refining relevant criteria in legal formulation and judicial practice, such as the expanded interpretation of the "three-step test" and the combination of the three-step test with the four factors of fair use in the third revision of the Copyright Law. These measures aim to build a more open and flexible general clause for rights limitations, ensuring the advancement of AI technology while effectively protecting the legitimate rights of copyright holders, achieving harmonious coexistence and development between the two.

4.1. The Degree of Threat to the Normal Use of the Original Work

The degree of threat posed by AI training data to the normal use of the original work can be analyzed from multiple perspectives, including the risk of infringing on the exclusive rights of copyright holders and the potential impact on the original work's market.

First, from the perspective of work utilization, AI poses risks of infringing on the reproduction and adaptation rights of copyright holders during the data input, machine learning, and content output stages. AI, by crawling vast amounts of works, may exceed the boundaries of fair use, potentially

DOI: 10.54254/2753-8818/2025.CH22295

leading to works becoming unprotected or overprotected. This process may infringe on the reproduction and adaptation rights of copyright holders, posing a significant threat to the normal use of the original work.

Second, regarding the potential impact on the original work's market, when determining whether AI data training constitutes fair use, it is necessary to consider whether it has a transformative nature and whether it will negatively affect the sales or potential market of the original work. The rights of copyright holders include the benefits obtained from exercising their rights in the existing market and the potential market benefits they are entitled to. This means that AI's use of works should not conflict with or compete with the normal use of the work. If AI-generated content achieves a transformation in the content or purpose of the original work and does not diminish the market value of the original, it may be considered fair use. If AI-generated content neither hinders the normal use of the original work nor excessively harms the legitimate rights of the copyright holder, it can be considered compliant, posing a minimal threat to the original work.

4.2. Defining the Purpose and Nature of Work Use

Defining the nature of fair use in AI training data requires a comprehensive consideration of the provisions on fair use in copyright law, the input and training process of non-expressive machine learning, and the output process of expressive machine learning.

According to the current Copyright Law, determining whether AI data training constitutes fair use under "special circumstances" is a significant issue. AI models are considered fundamental technical resources in technical systems, and their technical effects have a universally beneficial nature. The use of large amounts of works in model training serves legitimate purposes under copyright law, such as personal use, appropriate citation, or educational and scientific research use. Therefore, to meet the needs of AI technology development, it is necessary to expand the interpretation of "special circumstances." The U.S. *Campbell* case [10]. first recognized that transformative use of works promotes the goals of copyright law in advancing technology and art, arguing that "the more transformative use, the more likely it is to constitute fair use." Both domestically and internationally, the introduction of fair use exception clauses and the application of the three-step test have added a degree of flexibility. If AI-generated content neither hinders the normal use of the original work nor excessively harms the legitimate rights of the copyright holder, and the transformative use of the work is significant, it has a reasonable nature.

Non-expressive machine learning, during the data input and model training stages, is characterized by "non-expressive use" rather than direct use of the expressive content of works. Machine learning does not directly reproduce or express works, thus avoiding potential infringement issues. During the input stage, machine learning systems typically process and analyze large amounts of data, which may include copyrighted content, but the model itself does not directly reproduce or express this content. During the model training stage, the goal of AI is to statistically analyze data, extracting features and patterns through the learning of large amounts of data, rather than generating new works or expressions. The output stage of expressive machines is typically considered "expressive use," meaning the output content has a certain degree of creativity or originality and can be seen as a new form of expression.

5. Suggestions for Improving the Rules of Fair Use in AI Data Training

The current copyright law in China does not clearly stipulate the rationality of generative artificial intelligence in aspects such as creative use of works and data training. This leads to a lack of clarity and predictability in the practical operation of the fair - use system, and also reduces its adaptability and flexibility. From an international perspective, the United States has evolved from the four - factor

DOI: 10.54254/2753-8818/2025.CH22295

test, to an increasingly expansive interpretation of transformative use, and then to the "transformative + commercial" standard. This has, to a certain extent, influenced the connotations and denotations of relevant legal concepts. Therefore, the following specific suggestions are put forward to improve the rules of fair use in AI data training.

5.1. Clearly Define the Types of Machine Learning Eligible for Fair Use

Artificial intelligence machine learning can be divided into two types: "non - expressive machine learning" and "expressive machine learning". [11] Each type has specific conditions and limitations in the determination criteria for fair use. "Non - expressive machine learning" refers to machine learning that does not output expressive content and uses the factual information of works. This type of machine learning does not involve the expressive use of specific works, so it generally does not trigger copyright infringement issues. For example, during the data training process, it does not directly utilize the copyrighted content itself and does not output expressive content. Such cases can be regarded as fair use. "Expressive machine learning", on the other hand, involves the utilization of the expressiveness of specific works. Theoretically, it can be considered a form of fair use, especially when the output results are mainly used to test the effectiveness of algorithms. However, if the processed results are substantially similar to the original works and involve the protected expressions in these works, it may constitute an infringement. Additionally, if learning is carried out on the works of specific creators without permission, there may be a risk of infringement. This is because the purpose of using works by such expressive AI lacks transformative nature, and the generated content may affect the potential market of the author.

5.2. Clarify the Judgment Criteria for Fair Use in AI Data Training

China's Copyright Law has introduced the "three-step test" as the judgment standard for fair use of works. However, this standard still has some ambiguous aspects when determining whether AI data training constitutes fair use. This situation can be addressed through a more open - ended interpretation of the restrictive conditions or by drawing on judicial practice experience, thus providing some leeway for justifying exemptible reproduction in machine learning. According to the "three-step test", the first step is to determine whether it falls under "special circumstances". The United States determines whether a use falls within the scope of fair use by listing the purposes of use in the Copyright Act. Although these legal provisions adopt a closed - list approach, in practice, if the use behavior of a certain type of work aligns with the listed legislative purposes, a new type of fair use can be defined.[12]. China's current fair - use system in the Copyright Law has similarities. When it is difficult to judge whether the behavior of AI using works can be considered fair use based on the relevant provisions of the current Copyright Law, an extended interpretation can be made through the legislative purposes of the fair - use types listed in Article 24, but it must be ensured that the scope of the purposes does not exceed the limits set by the legal list. For example, Articles 3 to 6 respectively define some types of fair use, including "personal learning", "school classroom teaching", and "reproduction or quotation in news reporting". These provisions reflect the goals of protecting citizens' rights to participate in cultural activities and promoting the continuous inheritance and development of culture. Whether in the fair - use norms of Anglo - American countries or in China's fair - use regulations, despite differences in specific forms, they all embody the core requirement of the Copyright Law to promote the progress of culture and science and technology. The realization of this goal often depends on specific fair - use rules to appropriately limit the rights of copyright holders. Therefore, when the use behavior of a work meets the above - mentioned legislative purposes, it can be regarded as a fair - use behavior.

DOI: 10.54254/2753-8818/2025.CH22295

According to the second criterion, "shall not conflict with the normal use of the work", the behavior of using a work should not have an actual or potential market substitution effect on the author and should not prevent the author from extracting economic value from the work. In terms of the actual and potential markets, the market driven by AI deep learning differs from the original work market. Authors can neither anticipate these applications nor change the use patterns of their works. Therefore, AI data training actually does not interfere with the basic uses of works. If the content generated by AI contains the substantial expressions of the original work, it cannot be recognized as fair use. The behavior of AI learning can pass the second step of the "three - step test". It can be seen that during the process of AI creation, the way of using works neither interferes with the author's ability to obtain economic benefits from the work and the actual market nor conflicts with its normal use.

Regarding the third criterion, "shall not unreasonably prejudice the legitimate rights and interests of the copyright owner", with the continuous innovation of work - use methods, many new business models have emerged. This makes the criterion of "non - commercial use" as a measure of fair - use behavior no longer suitable for the digital age. In China's Copyright Law, there is no clear stipulation that non - commercial use is a necessary condition for fair use. In academic circles and judicial practice, most people believe that only when works are used for non - commercial purposes can it be considered fair use. In fact, in the Google Books case [13]. in the United States, a weakening of the emphasis on commercial - purpose use has been demonstrated, and the "transformative use" theory has been used as the judgment standard. In addition, AI does not use the substantial expressions in works but extracts data from them. The information such as facts and language rules contained in these metadata belongs to the public domain and should not be regarded as the "legitimate rights and interests" of copyright holders. Therefore, the behavior of AI data training does not unreasonably infringe on the legitimate rights and interests of copyright holders.

5.3. Clarify the Obligations and Responsibilities of AI Data Trainers

Improving the rules of fair use in AI data training requires not only clarifying the types of machine learning eligible for fair use and the judgment criteria but also defining the obligations of AI data trainers. The obligations of AI data trainers can be divided into three categories: the obligation of legal data sources, the obligation of data quality management, and the obligation of data transparency. First, AI data trainers need to fulfill the obligation of legal data sources, that is, they need to ensure that all data used are legally sourced. For publicly available data, it should be processed in a reasonable manner on the premise of meeting its open - use purpose; for data containing intellectual property rights, it is strictly prohibited to infringe on the legitimate rights and interests of others; if personal information is involved, personal consent must be obtained or it must meet other legal requirements. Second, there is the obligation of data quality management. Data trainers should implement effective measures, such as formulating detailed data - annotation standards, conducting data - annotation quality assessments, and performing random checks, to improve the quality of training data, enhance the accuracy and stability of data, and thus improve the maturity of the model and the quality of the generated content. In addition, AI data trainers should also fulfill the obligation of information disclosure and transparency. Data trainers should establish a transparent data - use mechanism and disclose information such as the source, purpose, and method of data use to relevant stakeholders to increase public trust in AI data training process and facilitate effective supervision and management by regulatory authorities.

DOI: 10.54254/2753-8818/2025.CH22295

6. Conclusion

The issue of fair use in AI data training encompasses not only legal dimensions but also, more importantly, the equilibrium between technological advancement and intellectual property protection. From the standpoint of balancing interests, China's intellectual property legal framework underscores the harmony between private and public interests. Given the challenge in precisely delineating the specific contributions of copyright holders, data trainers, users, and other stakeholders to resources and value, the absence of guidance from the principle of interest balance could lead to disputes and imbalances in the distribution of interests. Consequently, to foster the advancement of AI technology and protect the legitimate rights of copyright holders, it is imperative to establish clear criteria for judging fair use, delineate the types of machine learning, and specify the responsibilities of data trainers. This approach can ensure the harmonious coexistence and development of both technology and intellectual property, rectify market failures within the work-licensing system, stimulate technological innovation and progress, and harmonize the interests of all relevant parties.

References

[1] *Washington Law Review, [M]. How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem, 2018 ,93(2), pp. 583-584. the Journal of Legal studies, [J]. An Economic Analysis of Copyright Law, ,19 89,18(2), pp.326-361*

[2] *Liang Zhiwen. On Legal Protection of Artificial Intelligence Creations [J]. Legal Science (Journal of Northwestern University of Politics and Law), 2017, 35 (05):156-165*

[3] *L.Ray Patterson, Stanley W. Lindberg, The Nature of Copyright: A Law of Users' Right, 1991*

[4] *Liu Chuntian: Intellectual Property Law, [M] Higher Education Press, 2015 5th Edition, p. 122*

[5] *Sounds of Science: Copyright Infringement in AI Music Generator Outputs, [J] Cath. U. J. L. & Tech, 2021,29(2), pp.185-218.*

[6] *Wu Han Dong. A Question on Copyright Law of Artificially Generated Works [J]. Chinese and Foreign Jurisprudence, 2020, 32 (03): 653-673.*

[7] *Du Jialu. Behavioral Interpretation and Benefit Balance: Fair Use of Artificial Intelligence Training Copyright [J]. Electronic intellectual property, 2024, (06): 27-39.*

[8] *Article 47 of the Copyright Law of the People's Republic of China*

[9] *Google Books Library Project,Books[EB/OL]. [2014-05-06].*

[10] *Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 577 (1994).*

[11] *Artificial Intelligence's Fair Use Crisis, COLUM.J.L. & ARTS ,2017,41, pp.45-97.*

[12] *United States Copyright Act, Article 17, Section 107*

[13] *Authors Guild,Inc. v.Google,Inc. 804 F.3d 202(2nd Cir. 2015).*