Proceedings of the 4th International Conference on Global Politics and Socio-Humanities
DOI: 10.54254/2753-7048/2026.34525

Exclusive Licensing and Antitrust Regulation of AI Music Training Data

Boxi Zhang

Faculty of Law & Intellectual Property, Hubei University of Automotive Technology, Shiyan, China
zhangbx0106@outlook.com

Abstract. The deep integration of the digital music industry and artificial intelligence technology has given rise to exclusive licensing of AI music training data, a new type of copyright transaction model. While playing a positive role in promoting the legalization of music, reducing transaction costs, and stimulating creative vitality, this model also brings monopoly risks such as the centralization of training data and the formation of market barriers due to its exclusive characteristics, posing potential impacts on market competition order and public interests. This paper conducts research from the dual dimensions of copyright law and antitrust law, arguing that the licensing model should not be simply affirmed or denied, and a scenario-based identification approach should be adopted to distinguish the boundary of rationality. The study finds that the technical characteristics of AI training data pose new dilemmas for relevant market definition and market dominance determination, making traditional regulatory methods inapplicable. This paper puts forward targeted improvement plans: improving relevant market definition with quality substitution tests, optimizing differentiated licensing regulation with data substitutability and market impact, reconstructing the concentration review system with the dual-dimensional standard of "data + technology", and promoting the dynamic coordination of copyright law and antitrust law.

Keywords: AI Music Training Data, Exclusive Licensing, Antitrust Regulation

1. Introduction

With the deep integration of artificial intelligence technology and the digital music industry, exclusive licensing of AI music training data has emerged as a new form of copyright transaction. While promoting the legalization of music, reducing transaction costs and stimulating creative vitality, its exclusive feature also brings monopoly risks such as data centralization, solidified market barriers and distorted competition order, raising a new proposition for balancing copyright protection and market competition. At present, relatively rich research results have been formed on digital music copyright licensing and antitrust regulation. Existing studies have demonstrated the institutional function of exclusive licensing of digital music copyright from the perspective of copyright law, proposed a governance path of "statutory license + classified regulation", and adopted differentiated restrictive rules for different types of licenses [1]. Other studies, taking the traditional digital music market as the object, have carried out antitrust analysis from the three dimensions of

DOI: 10.54254/2753-7048/2026.34525

abuse of market dominance, monopoly agreements and concentration of undertakings, pointing out practical dilemmas in relevant market definition and market dominance determination [2].

However, existing theoretical and regulatory studies still have obvious deficiencies and fail to respond to the challenges brought by technological changes: first, in judging the legitimacy and rationality of exclusive licensing of AI music training data, existing studies mostly follow the logic of traditional copyright exclusive licensing or a single competition analysis framework, without conducting scenario-based and technical identification combined with the technical characteristics of AI training data, failing to distinguish between commercial and non-commercial use, and between ordinary training and generative use. The boundary between the exercise of private rights and public interests is not clear enough, making precise regulation impossible. Second, in terms of the adaptability of antitrust regulation, existing studies still rely on the experience of traditional digital music copyright regulation, failing to fully respond to the new monopoly challenges brought by AI music training data. The analysis of core elements such as relevant market definition and market dominance determination cannot adapt to the new characteristics of pan-entertainment, two-sided markets and data-driven competition. The traditional hypothetical monopolist test and market share standards are inapplicable, and there is a lack of effective responses to new issues such as data centralization, algorithmic barriers and covert data alliances. Third, in terms of legal coordination, existing studies are mostly conducted separately from copyright law or antitrust law, lacking a governance idea of dynamic coordination between the two laws, making it difficult to form a regulatory system that balances right protection, technological innovation and market competition.

Based on this, based on the dual dimensions of copyright law and antitrust law, this paper clarifies the rational boundary of exclusive licensing of AI music training data with a scenario-based approach, analyzes the new antitrust challenges it raises, improves the regulatory system from four levels: relevant market definition, differentiated licensing regulation, concentration review and inter-law coordination, and constructs a legal governance framework that balances right protection, technological innovation and market competition, providing theoretical support and regulatory reference for the healthy and orderly development of the AI music industry.

2. Rationality of exclusive licensing of AI music training data from the perspective of copyright law

2.1. Controversy and criticism on the rationality of existing exclusive licensing of AI music training data

The legitimacy-and-rationality theory holds that exclusive licensing of AI music training data is an exercise of private rights by copyright owners and complies with provisions on exclusive licensing in the Copyright Law. This view states that "as an important part of private rights, copyright is designed to protect the legitimate rights and interests of right holders. Its essence is to grant copyright owners the freedom to dispose of their works independently, enabling them to maximize economic interests through exclusive licensing and other means" [3]. Some scholars also suggest that data training for generative AI can be justified through the three-step test, and a special fair use clause for AI learning and creation should be added to the Copyright Law to directly recognize the legitimacy of using works for AI training, thereby supporting the justification of the exclusive licensing model [4]. Other studies point out that using works for AI large-model training has limited impact on the legitimate rights of copyright owners and can produce significant public interests, so it should be included in fair use, further affirming the justification and legitimacy of such licensing [5]. Another view explicitly argues that AI data training is transformative and non-expressive, so

DOI: 10.54254/2753-7048/2026.34525

both commercial entities and research institutions should be included in the scope of fair use subjects, directly recognizing the legitimacy of using works for AI training and strengthening the private-right justification of exclusive licensing [6].

Nevertheless, such views obviously ignore the particularity of AI music training data: AI training requires technical reproduction and feature extraction of massive works, and the use is non-specific and process-based, which is fundamentally different from the direct utilization of traditional works [7]. Fully recognizing the legitimacy of exclusive licensing will lead to misalignment of right boundaries, unstable legal basis for licensing effectiveness, long-term monopoly of training data, hindrance to knowledge dissemination, and violation of the legislative purpose of copyright law to promote cultural communication.

The necessity-of-restriction theory argues that exclusive licensing of AI music training data is likely to trigger monopoly risks and should be fully restricted or even prohibited. Based on empirical research on competition risks of exclusive licensing of digital music copyright, this view points out that exclusive licensing will significantly strengthen platform user lock-in, result in long-term lock-in of core music libraries and training data, raise market entry barriers and squeeze the living space of small and medium competitors. Extensive exclusive licensing is inherently exclusive and restrictive of competition and should be fully banned or strictly restricted [8]. Some scholars further note that even if exclusive licensing has certain value in promoting legalization, platforms will still seriously damage competition order and consumer interests if they refuse to distribute or impose unreasonable high prices by virtue of exclusive agency rights. Therefore, the extensive exclusive licensing model should still be strictly restricted or even negated [9]. Citing the National Copyright Administration's suspension of exclusive licensing of digital music copyright in 2022, it is held that "exclusive licensing will lead to long-term lock-in of core music libraries and monopoly of secondary licensing channels", forming substantial market barriers [3].

2.2. Scenario-based identification and rationality of exclusive licensing of AI music training data

The scenario-based identification advocates judging the rationality of exclusive licensing by distinguishing specific scenarios, recognizing its legitimacy in specific situations and preventing unreasonable exclusive licensing through scenario-based restrictions, which is more in line with the actual development of the AI music industry. AI music training data has dual attributes: as a work carrier, it is protected by copyright law, and rights holders enjoy the right of legal disposal; as training material, its use is technical and large-scale, and the needs of technological innovation should be taken into account [10].

From the perspective of the right basis, exclusive licensing of AI music training data must be premised on legal data sources. Technical intermediate reproduction that only extracts abstract rules of works can be included in the exception category of copyright law, but situations that can be reversely restored or generate similar expressions with high frequency require explicit authorization [10]. From the perspective of usage scenarios, exclusive licensing for non-commercial scientific research training can be appropriately relaxed, while exclusive licensing for commercial AI music generation, profit-making services and other scenarios requires the establishment of right checks and balances such as compulsory sublicense and fair pricing [7]. This scenario-based identification approach responds to the technical characteristics of AI training data and is the most feasible theoretical proposition at present.

DOI: 10.54254/2753-7048/2026.34525

3. New antitrust challenges caused by exclusive licensing of AI music training data

On the basis of clarifying the rationality of exclusive licensing of AI music training data, it is necessary to further examine the antitrust issues it causes in market practice. The combination of the technical characteristics of AI music training data and the exclusive licensing model has broken through the application scenarios of traditional antitrust regulation, bringing a series of new challenges, and traditional rules show obvious insufficient adaptability in responding to these challenges.

The technical characteristics of AI music training data significantly increase the difficulty of defining relevant markets. Traditional antitrust relevant market definition relies on demand substitution and supply substitution analysis, but the use of AI music training data is concealed and derivative, and its market boundary goes beyond the scope of traditional digital music services. On the one hand, the demanders of AI music training data include not only professional music platforms but also pan-entertainment entities such as short video platforms and AI tool developers. "The pan-entertainment trend increases the impact on relevant market definition", and the demand substitutability of these entities is difficult to evaluate through the traditional SSNIP test method [2]. On the other hand, the supply of AI music training data involves multiple links, such as work reproduction, feature extraction and model training, and technical use in some links is not directly for consumers, leading to vague supply substitution relations [10]. The interweaving of this pan-entertainment trend and technical complexity makes "the traditional hypothetical monopolist test inapplicable to digital music platforms", and relevant market definition falls into a dilemma [2].

4. Developmental suggestions on antitrust regulation of exclusive licensing of AI music training data

In response to the new antitrust challenges caused by exclusive licensing of AI music training data, it is not advisable to simply apply traditional regulatory logic. Instead, based on its technical characteristics and market rules, a developmental regulatory system integrating copyright law and antitrust law, combining ex ante prevention and ex post regulation, should be constructed to achieve a dynamic balance among right protection, technological innovation, and market competition.

4.1. Improving relevant market definition methods and market dominance determination standards

In view of the technical characteristics of AI music training data, innovative ideas for relevant market definition should be adopted. Due to the large number and pan-entertainment trend of demand subjects for AI music training data, the traditional demand substitution SSNIP evaluation is difficult to determine its market. Some scholars propose introducing the "quality substitution test" into the evaluation method, using service quality changes as indicators of demand elasticity testing, but do not elaborate on the specific operation of the "service quality" dimension. This paper supplements here: service quality should include the richness and integrity of training data, sound quality clarity of tracks generated by the platform, and multi-terminal adaptability. When any of the above service qualities decreases slightly, observe the changes in user churn, and determine the market according to the SSNIP principle. Market dominance determination should break through the traditional market share standard, focusing on "the substantial share of training data (including quantity and core quality), the term of exclusive rights, platform technical strength and data

DOI: 10.54254/2753-7048/2026.34525

processing capabilities". Stricter review standards shall apply to leading platforms that master core training data and have algorithmic advantages.

Existing differentiated licensing and antitrust regulation programs for AI music training data fail to fully adapt to its characteristics of diverse data sources, high value dependence on derivative processing, and market competition centered on the "data-algorithm-computing power" closed loop, with multiple structural defects: first, the classified regulation logic centered on "training set/test set" is rigid, only setting term and quantity thresholds for copyrighted works, not covering non-copyright training data such as UGC content, public domain materials and derivative data, nor regulating the "data value monopoly" formed by platforms through data cleaning, annotation and algorithm optimization, let alone restricting the implicit exclusive behavior of cross-platform data pool alliances; second, the traditional framework of vertical monopoly agreements and concentration review fails, which can neither identify new exclusive behaviors such as data capture restrictions, derivative data control and data service tying, nor incorporate core competition indicators such as data scale, diversity, uniqueness and algorithmic computing power barriers into concentration review evaluation factors, lacking penetrating supervision over concealed data concentration forms such as non-equity cooperation and data sharing alliances.

4.2. Optimizing differentiated licensing regulation

Adjust the classification standard to "data substitutability + market impact" to replace the original "training set/test set" division. For core irreplaceable training data, the FRAND (Fair, Reasonable and Non-Discriminatory) principle shall be compulsorily applied, requiring platforms to open non-exclusive data interfaces. For general data and public domain data, a standardized bulk licensing channel shall be established through copyright collective management organizations to simplify licensing procedures and reduce compliance costs for small and medium-sized entities.

4.3. Reconstructing the concentration review system

Construct a dual-dimensional evaluation standard of "data + technology" to replace the single standard of copyright share or turnover. The evaluation elements include core competition indicators such as training data coverage, data exclusivity period, data barrier degree, as well as platform computing power and algorithm model advantages; implement penetrating review of mergers and acquisitions, non-equity cooperation, data alliances and other acts involving core training data, focusing on assessing whether data concentration will form a "data-model" closed-loop monopoly, and strictly preventing concealed data concentration risks.

4.4. Promoting dynamic coordination of copyright law and antitrust law

Coordination between copyright law and antitrust law is the key to solving regulatory dilemmas of AI music training data. Improve the statutory license system in the Copyright Law, "include interactive digital audio streaming services dominated by online music on-demand playback into the scope of statutory licenses", and set restrictions on duration, quantity and sub-licensing proportion to balance right income and public access [3]. Meanwhile, establish a "text and data mining exception" rule, allowing "fair use of copyrighted data for AI training for non-commercial scientific research purposes on the premise of ensuring legitimate data sources" [10]. Antitrust enforcement should adhere to a behavioral approach, "distinguish between legitimate exclusive licensing and abusive

DOI: 10.54254/2753-7048/2026.34525

acts, and avoid excessive regulation that inhibits technological innovation", so as to achieve a dynamic balance among copyright protection, technological innovation and market competition [2].

5. Conclusion

Exclusive licensing and antitrust regulation of AI music training data are common issues faced by copyright law and antitrust law in the digital economy era, whose core lies in balancing the private right protection of copyright owners, the technological innovation incentives of platforms and the fair competition order of the market. The technical and scenario-based characteristics of AI music training data determine that the rationality of its exclusive licensing cannot be simply affirmed or denied, but must be distinguished scenario-based based on the eclectic theory. The new antitrust challenges it brings, such as difficulties in relevant market definition and concealed monopoly acts, require adaptive innovation of traditional regulatory rules. The future regulatory system should achieve precise regulation of exclusive licensing through scenario-based distinction, and form a governance force through the coordination of copyright law and antitrust law. With the continuous development of AI technology in the music field, the licensing model and competition form of AI music training data will continue to evolve. Future research needs to further strengthen the combination of theory and practice, refine rule design under specific scenarios, pay attention to international rule coordination and cross-border regulatory cooperation, so as to provide a more solid legal support for building a healthy and orderly digital music industry ecosystem.

References

[1] Zhao, L., Li, J. (2026) Research on the Regulatory Path of Exclusive Licensing of Digital Music Copyright from the Perspective of Copyright Law. Journal of Chongqing Three Gorges University, 4-6.
[2] Zhang, L. (2025) Research on Antitrust Regulation of Digital Music Copyright Exclusive Licensing Model. Yantai University, 18-37.
[3] Yang, D. (2026) On the Legitimacy of Generative AI Training Data and Copyright Infringement Determination. Journal of Guangxi Administrative Cadre Institute of Politics and Law, (1), 3-10.
[4] Xia, W. (2025) Fair Use of Training Data in Generative Artificial Intelligence. Proceedings of CONF-MPCS.
[5] Ding, Y. (2025) Fair Use in AI Data Training: Judgment Criteria and Balancing Mechanisms. Proceedings of CONF-MPCS 2025 Symposium: Mastering Optimization: Strategies for Maximum Efficiency.
[6] Zhu, W., Chen, L. (2026) Construction of Copyright Fair Use Rules for Artificial Intelligence Data Training. Journal of Taiyuan Normal University (Social Science Edition), 25(01), 46-53.
[7] Niu, B. (2026) Research on Copyright Fair Use Rules in Generative AI Data Training. Hebei Law Science, (4), 90.
[8] Zheng, Y., Lei, Y. (2023) Antitrust Analysis of Exclusive Copyright Model of Digital Music Service Platforms. New Media and Society, 3-18.
[9] Huang, J. (2025) Copyright Governance of Digital Music Platforms in the AI Era: Technical Challenges and Institutional Reconstruction. Social Music Life, 87.
[10] Huang, Y. (2026) Copyright Regulation of Exclusive Licensing of Digital Music Copyright. Communicate & Copyright, (3), 146.