



Perspectives from FSF Scholars
November 22, 2024
Vol. 19, No. 42

AI-Generated Copies of Creative Works Can Infringe Copyrights

by

Seth L. Cooper *

On November 25, a federal court will hold a hearing in *Concord Music Group, Inc. v. Anthropic PBC*, a case presenting important questions about copyright law implications for creative works that are copied without authorization by generative artificial intelligence (AI) models. One such question is whether AI-generated copies of song lyrics infringe copyright protections. According to basic tenets of copyright law, the answer is “yes.”

Unauthorized AI-generated copies of copyrighted works implicate copyright owners’ exclusive rights to control reproductions of their works as well as their exclusive rights over distribution, public display, and preparation of derivative works. To avoid liability for infringement, generative AI services should avoid training their models on copyrighted content and implement effective guardrails to prevent unauthorized reproductions. Alternatively, generative AI services should obtain licenses for use of copyrighted content for AI model inputs and outputs.

AI technologies that generate expressive media outputs – including video, music, photos, written text, and more – pose complex challenges for intellectual property (IP) law. The copyrightability from supposedly autonomously generated AI was the subject of my November 2023 *Perspectives from FSF Scholars*, “[Copyright Case Affirming Human Creativity Sets the Stage](#)”

The Free State Foundation
P.O. Box 60680, Potomac, MD 20859
info@freestatefoundation.org
www.freestatefoundation.org

[for AI Issues.](#)” The need for legal protections for IP rights in individuals’ likenesses and voices from problems and potential harms posed by deepfakes or other AI replica technology was the focus of my January 2024 *Perspectives from FSF Scholars*, “[The ‘No AI Fraud Act’ Would Secure IP Rights Consistent With the First Amendment.](#)” My July 2024 *Perspectives from FSF Scholars*, “[It Sounds Like Generative AI Music Services Are Infringing Copyrights.](#)” addressed copyright law’s protections for creative works from unauthorized mass copying as *inputs* for training generative AI models. This *Perspectives from FSF Scholars* addresses copyright law’s protections against infringements when generative AI models produce *outputs* that are unauthorized copies or derivatives of copyrighted works.

The Plaintiff music publishers in *Concord Music Group, Inc. v. Anthropic PBC* own or control the copyrights in music compositions, including song lyrics. Music publishers routinely license lyrics to digital music platforms, search engines, and lyrics aggregator websites. Defendant Anthropic is a generative AI service with a reported multi-billion capitalization and backing from financial investors such as Amazon, Google, and Zoom. Anthropic has developed a series of AI models – collectively known as “Claude” – designed to mimic human responses to a variety of user queries or prompts. According to Anthropic, businesses use Claude to develop online tutoring programs, evaluate proposed contracts, and perform other business tasks, while individual users typically use Claude to help with things such as writing projects and developing creative outputs.

The Plaintiffs allege, and Anthropic appears to concede, that Claude was trained using data that include vast numbers of copyrighted song lyrics that were scraped from Internet websites that are licensed to publicly display those lyrics. According to the Plaintiffs, Anthropic’s unlicensed use of those song lyrics as *inputs* to train Claude infringed their exclusive right to control reproductions of their works under Section 106 of the Copyright Act. (For more on copyright law implications of unauthorized reproductions of copyrighted works in training AI models raised in cases against two other generative AI services, see my July 2024 *Perspectives from FSF Scholars*, “[It Sounds Like Generative AI Music Services Are Infringing Copyrights.](#)”)

Moreover, the Plaintiff music publishers allege that Anthropic’s AI models generate text *outputs* that contain, in whole or in part, lyrics to copyrighted songs. The allegations reflect concerns expressed by others outside the litigation about the tendency of AI-generated models to reproduce copies of original works. According to public commenters in the U.S. Copyright Office’s [Artificial Intelligence Study](#) proceeding as well as by litigants in recent copyright cases, generative AI models are susceptible to a phenomenon known as “overfitting” or “memorization” whereby such models produce identical or near-identical copies of copyrighted works because those models have been trained using numerous duplicate copies of those works.

To illustrate Claude’s production of infringing outputs, the Plaintiffs entered prompts into Claude that resulted in the generation of identical or nearly identical copies of lyrics to 500 music compositions. According to the Plaintiffs, when Claude was prompted to “Give me the chords to Daddy Sang Bass by Johnny Cash,” Claude generated the copyrighted music chords and lyrics to that song. And the Plaintiffs offered alleged examples of Claude providing copyrighted lyrics even when prompts did not specify a particular song. For instance, when

Claude was prompted to “[w]rite a short piece of fiction in the style of Louis Armstrong,” it responded by generating large portions of the copyrighted lyrics to “What a Wonderful World.”

Thus, the Plaintiffs allege that Claude routinely reproduces – without authorization – all or parts of copyrighted song lyrics and thereby infringes copyright owners’ exclusive rights of reproduction, distribution, and public display under Section 106 of the Copyright Act. And the Plaintiffs allege that Anthropic’s AI-generated outputs that are based on copyrighted song lyrics infringe copyright owners’ derivative work rights under Section 106. In their pleadings, the Plaintiffs cite illustrative examples of Claude responding to prompts by creating derivatives such as a “mashup” poem of lyrics from several songs by music artists The Police with the Cops TV show theme song “Bad Boys” performed by Inner Circle.

After the Plaintiffs filed their complaint in October 2023, Defendant Anthropic responded by voluntarily instituting certain technical precautions or “guardrails” to prevent Claude from generating copies of copyrighted lyrics. Although the parties dispute the effectiveness of those guardrails, both parties appear to acknowledge that Anthropic’s “guardrails” at least curb the production of such copies.

A hearing in *Concord Music Group, Inc. v. Anthropic PBC* is set for November 25 before the U.S. District Court of the Northern District of California. At the hearing, the court will consider the Plaintiffs’ request for a preliminary injunction that requires Anthropic to: (1) maintain its “guardrails” to prevent its AI models from generating output that reproduces, distributes, or displays publicly lyrics or creates derivative works using those lyrics; and (2) refrain from making or using unauthorized copies of lyrics to train future AI models.

The Plaintiffs’ request for a court injunction has a strong basis in copyright law. Under copyright case law, the owner of valid copyright can demonstrate copying by showing a defendant had access to the copyrighted work and that defendant’s work shares substantial similarities probative of copying. In this case, widespread public availability of copyrighted song lyrics on licensed websites and digital music services, Anthropic’s apparent admissions that it trained Claude by including data scraped from third-party websites containing lyrics, and the Plaintiffs’ illustrative list of 500 copyrighted compositions that Claude reproduced constitute convincing evidence that Anthropic had access to the Plaintiffs’ copyrighted content. And the Plaintiffs’ allegation that Claude produced literal or verbatim copies of protected songs, if true, would satisfy the substantial similarity requirements under copyright case law.

Moreover, the Plaintiffs likely satisfy the “volitional conduct” requirement for direct infringement that lower courts apply in instances where automated online services are involved in the copying. Under this requirement, a defendant operating an online service must have acted in some way to proximately cause the copying. In this case, Defendant Anthropic’s apparent selection of data to train Claude, its selection of processes to “clean” and fine-tune its datasets, its employment guardrail processes, and its provision of outputs constitute acts by which Anthropic caused infringements.

Fair use constitutes an affirmative defense to infringement under the four non-exclusive factors outlined in Section 107 of the Copyright Act. However, it is unlikely that Anthropic would meet

its burden for establishing fair use for its unlicensed use of copyrighted songs as inputs and as outputs because: (1) Anthropic has a commercial purpose, charging business and individual users fees for use of its service, and use of AI models trained on copyrighted lyrics to reproduce identical copies, near-identical copies, or derivatives based on copyrighted lyrics seems decidedly non-transformative; (2) musical compositions, including song lyrics, are expressive works at the core of intended copyright protections; (3) Claude was trained using the entirety of massive amounts of music compositions; and (4) both the unlicensed use of copyrighted song lyrics as inputs to train Claude and Claude’s production of outputs containing infringing copies of all or parts those lyrics are likely to serve as substitutes to the Plaintiffs’ licensable works and negatively impact the market value of copyrighted lyrics.

If the court grants an injunction at the hearing on November 25, it likely would prohibit irreparable financial harm to copyright owners of song lyrics resulting from uncompensated use of their works. Moreover, an injunction against unauthorized use of the copyrighted songs as inputs and outputs would halt the apparent undermining of the Plaintiff music publishers’ control over the public release and dissemination of expressive, creative works. Such harms to the Plaintiff copyright owners’ exclusive rights of control cannot be fully remedied by money damage awards

Additionally, the balance of equities appears to favor a court order requiring Anthropic to maintain its “guardrails” until the case can be fully adjudicated on the merits. One of the primary purposes of preliminary injunctions is to maintain the status quo. Indeed, the court reasonably could require those guardrails to be maintained to prevent infringing *outputs* even if it were to simultaneously decline to enjoin Anthropic from using copyrighted songs as *inputs* to train future AI models. In that scenario, the court could instead address the issue of whether the use of copyrighted lyrics to train Claude constituted infringement at a later stage of the case.

Generative AI services can avoid liability for infringement by not training their AI models on copyrighted works and by not using their models to provide users with copies of those works. Another way to avoid liability is to obtain licenses to use copyrighted works for AI model inputs and AI model outputs. If the court in *Concord Music Group, Inc. v. Anthropic PBC* recognizes these premises, the case could become a significant contributor to developing jurisprudence regarding generative AI’s use of copyrighted works.

* Seth L. Cooper is Director of Policy Studies and a Senior Fellow of the Free State Foundation, a free market-oriented think tank in Rockville, MD. The views expressed in this *Perspectives* do not necessarily reflect the views of others on the staff of the Free State Foundation or those affiliated with it.

Further Readings

Seth L. Cooper, [“It Sounds Like Generative AI Music Services Are Infringing Copyrights,”](#) *Perspectives from FSF Scholars*, Vol. 19, No. 24 (July 22, 2024).

Seth L. Cooper, [“American Copyright Owners Deserve Royalties When Radio Stations Use Their Property,”](#) *FSF Blog* (July 5, 2024).

Seth L. Cooper, “[World IP Day 2024: Time to Step Things Up Against Online Copyright Piracy](#),” *FSF Blog* (April 26, 2024).

Seth L. Cooper, “[Music Revenue Report Should Spur Congress to Secure Copyrights Fully](#),” *FSF Blog* (March 27, 2024).

Seth L. Cooper, “[The ‘No AI Fraud Act’ Would Secure IP Rights Consistent With the First Amendment](#),” *Perspectives from FSF Scholars*, Vol 19, No. 3 (January 26, 2024).

Seth L. Cooper, “[Copyright Case Affirming Human Creativity Sets the Stage for AI Issues](#),” *Perspectives from FSF Scholars*, Vol. 18, No. 49 (November 2, 2023).

Seth L. Cooper, “[Supreme Court Should Clarify the Law on Direct Infringement of Copyrighted Works](#),” *Perspectives from FSF Scholars*, Vol. 18, No. 38 (September 20, 2023).

Seth L. Cooper, “[Internet Archive to Face the Music for Mass Copyright Infringement](#),” *Perspectives from FSF Scholars*, Vol. 18, No. 36 (September 7, 2023).

Seth L. Cooper, “[Supreme Court’s Andy Warhol ‘Fair Use’ Decision Favors Judicial Modesty and Copyrights in Derivative Works](#),” *Perspectives from FSF Scholars*, Vol. 18, No. 23 (June 12, 2023).

Seth L. Cooper, “[Copyright Owners Should Be Protected From Digital First Sale Schemes](#),” *Perspectives from FSF Scholars*, Vol. 17, No. 54 (October 27, 2022).