Philip Bell

2 August 2014

ENGL 161

WP4 - Research Paper Final Draft

# Using Copyright Material In AI Machine Learning; You Wouldn't Download A Car

## Introduction

When looking at how copyright is applied to music and creators, there is a longstanding combination of litigation and already-set-precedent that has determined who the owner of said music is, how, and where that music can be used or distributed. With this ownership comes the ability to protect your assets from infringement and in the music industry, this applies to the material one has created as well as its redistribution. When it comes to generative artificial intelligence (AI), the law is less clear and there are a lot of legal questions still unanswered. More complete research needs to be done about whether or not people care about listening to AI-generated music vs. human-created music and even though copyright infringement is an extremely litigated topic, it's important that more precedent be set about using copyrighted material when training artificial intelligence.

As copyright law adapts, AI and its abilities blow past it at light speed. It is easy for anyone to log into a website today and create musical or artistic content simply by supplying an AI module with some information to get its resources going. This research looks at how, when used ethically, AI can be utilized for making improvements for individual artists and bands,

companies and organizations in fields outside of the music industry, and how the public perceives said music.

Sprinting from the start line as a creative artist, research was started on this topic with an inherent bias against using AI to generate music. As my research progressed, I found myself flipping the switch on my bias and beliefs about AI in the music industry were challenged at every turn I took. Some of this research has uncovered that there appears to be a persistent line of thought from other researchers that the public should care whether or not music is created with AI, but multiple surveys and tests have proven these theses to be incomplete or incorrect (Zlatkov). A more complete look at two previous studies by Moffat and Kelly (2006) and Pasquier et al. (2016), as well as what Zlatkov and his team uncovered are outlined in a later section of this paper.

**Generative AI in the Music Industry**

AI is no stranger to music. The first use of AI to help generate a musical composition was completed in 1956 when Lejaren Hiller and Leonard Isaacson conducted research and development at the University of Illinois Urbana-Champaign campus. The purpose of the AI machine back then was sole: to create a four-movement string quartet. Students from the university performed three of the four movements later in the year to demonstrate how applicable and realistic it was (Zipris).

In the 1960s Iannis Xenakis continued down the path of using computers, algorithms, and mathematics to aid in computer-generated music. Where he differed from his predecessors was that his work illustrated methods for the creation of new sounds. Xenakis applied a much more

scientific method and mathematical approach to his creations. Instead of simply applying existing knowledge and algorithms to his musical pieces, he took a more stochastic approach, creating music that had randomness intertwined with his science experiments. Then he analyzed the pieces statistically, drew data from them, and created more. His approach shined the light on using AI as a supplementary analysis tool, rather than only a composition tool (Ranwala).

Composer and professor of music, David Cope, started to rely on the concept of recombinancy in the 1980s, which is a concept we still see prevalent in today's AI learning modules. Recombinancy is the process by which AI does its learning; it reuses material from previous works, modifies and combines them to create new work. Cope also combined the same thought processes from his predecessors of the 50s and 60s and his works hinged on three main principles which were:

- **Deconstruction** whereby the machine analyzes and separates data.
- **Signatures** which showcase commonalities allowing the machine to understand different musical genres.
- **Compatibility** which he uses synonymously with recombinancy, defined above (Ranwala).

His process was six fold and quite detailed and laid a lot of the groundwork for what we see in the way AI generates material today.

Since the 2000s, AI is becoming more and more mainstream in the arts. AI is used in classrooms when teachers are picking music. A suggestion will come up on their computer screen based on what the in-room microphone has heard and what the AI analyzed (Yu). AI is used in recording studios to help producers even out audio levels and to help artists overcome writer's block (Hight). Another use of AI that has been positively received by the general public

and the music industry was when the late John Lennon's last demo was finalized by fixing the vocals for the song "Now and Then" with the remaining Beatles (Welch).

**In re: Human vs. Machine - The Copyright Law Depositions**

The United States Copyright Office has confirmed that copyright law will only protect works made by humans (Office). This means that output from AI modules can not be copyrighted. Mark Lemley, who is director of the Stanford Program in Law, Science and Technology, and co-author Bryan Casey, explore several studies that outline some other possibilities beyond what the US Copyright Office has claimed in their article, "Fair Learning." They cite one claim that AI should be the owner of its output, or more specifically that the company who developed, owns, or deploys the AI is the author and owner of its output. After all, it is perfectly legal for a record label to own copyright of master recordings or a band's full discography (GRAM) - it is not a stretch of the imagination to look at the developers of AI modules as the owners of the output.

Fortunately there is another avenue explored by Lemley and Casey which winds us down the path of opportunity and aligns The Copyright Office's stance with the use of AI to create music. What if the human's input to the AI module was protected by copyright? The more detailed the prompt given to the AI module, the easier it would be to protect under copyright law. The way copyright law currently reads, it protects an artist who can prove that their idea and their representation of their idea is unique enough to warrant copyright protection. Furthermore, this artist has to prove that they suffered financially by someone copying their work in order to file suit for copyright infringement (Lemley and Casey). Current copyright law also protects

creative works during the author's life plus an additional 70 years after their death so a lot of leeway is given to the human in this regard. There needs to be some precedent set here from the court system about ownership so that this topic doesn't devolve into an anything-goes area of law.

Furthermore, anything-goes is what might happen if AI isn't applied to music ethically. Just last year, in 2023, a song was uploaded to popular streaming platforms mimicking popular artists Drake and The Weeknd. "Heart on my Sleeve" was taken off streaming platforms almost immediately when Universal Music Group (UMG) claimed copyright infringement. Streaming platforms obliged, but not before the song received multi-million listens, including 15 million hits alone on TikTok. UMG argued that these platforms have a "legal and ethical responsibility to prevent the use of their services in ways that harm artists." UMG further claims that using copyright material of their artists for machine learning is a breach of contract and violates copyright law, but that does not appear to be the case according to The Copyright Office, nor the consensus on the matter (Snapes).

When looking at what can be used for teaching AI modules, UMG brings up an interesting point, but not one without some dissent. Lemley and Casey argue that copyright material should be used to teach AI modules. Jenny Quang agrees and goes one step further claiming that there need to be safe harbor rules for data mining practices because current laws are vague. There need to be precedents set by court cases that lay out what is legal and what would be allowed in the eyes of the court of copyright law when it comes to the legality of using copyright material for machine learning (Quang).

Data mining allows AI to find patterns in data collected. This data can include millions of pictures, videos, songs, lyrics, poems, etc. A lot of companies employing this technology are using the fair use doctrine as their legal protection (Fair Use), but Quang argues for more specificity and claims that "establishing a clear right to use copyrighted materials for data mining is consistent with the goals of copyright law and would remove the barriers to innovation." Furthermore, the transformative vs. non-transformative argument is still being litigated in court (Samuelson) which outlines whether or not a new piece of work adds enough new qualities to be considered different from the original. Lemley and Casey agree and claim that the outputs created by AI that was trained with copyrighted material is more inclusive, and the algorithms produced are better, safer, and more fair than modules using only public domain (Lemley and Casey).

But as we propel into these uses of AI, we hope that things are done ethically, but can not assume they are. Fabio Morreale outlines some very basic, but very important ethical issues. The senior music lecturer at the University of Aukland in New Zealand claims that because the field of AI is advancing so much faster than its corresponding litigation, it is important that those who are in the field remain proactive in policing the use of AI themselves (Morreale). Morreale specializes in the field of Music Information Retrieval which is an interdisciplinary science that looks at extracting information from music. This field looks closely at and combines such fields as musicology, psychology, machine learning, optical musical recognition, and other disciplines.

200 creative artists in the music industry, known as the Artist Right's Alliance, signed an open letter in April of 2024 recognizing that AI can have a positive impact on musicians, their creative content, and the music industry, but they argue against its widespread use (Artist

Right's). They urge companies to use AI ethically and want to stop irresponsible business practices, like the example outlined above mentioning Drake, The Weeknd, and UMG. They claim that smaller artists are the ones who suffer, not large creators with millions of followers, millions of records sold, and millions of dollars tied up in litigation to protect their content. The Artist Right's Alliance also argues for fair payment to any artist who's work is used to train AI (Hammond). Right now this is an uphill battle that is fresh, novel, and still in limbo while litigation continues.

> "When used irresponsibly, Al poses enormous threats to our ability to protect our privacy, our identities, our music and our livelihoods. Some of the biggest and most powerful companies are, without permission, using our work to train Al models. These efforts are directly aimed at replacing the work of human artists with massive quantities of Al-created "sounds" and "images" that substantially dilute the royalty pools that are paid out to artists. For many working musicians, artists and songwriters who are just trying to make ends meet, this would be catastrophic."

**Who Cares?**

This research has uncovered several studies that were previously conducted in an attempt to learn whether or not people think AI-generated music is creative and whether or not there is a bias against content created by AI when compared to human-created art. Outlining a study conducted by Dimiter Zlatkov et al. at the University of British Columbia in 2023, the authors gave several examples of field studies that have already been conducted. For the sake of this paper, a few of the studies have been summarized below.

Moffat and Kelly from 2006 attempted to explore whether or not there was a bias against computer-composed music, but there were only 20 participants in their study. They found that their stats appeared to show prejudice against AI-created works, but the small sample size made it impossible to draw any real conclusions. Three styles of music were displayed during this test.

Pasquier et al. expanded on Moffat and Kelley's previous work and asked again if listeners hold a bias against computer-composed music. They believe that with an improved test, adjusted variables, and a limitation to a single style of music produced, they will be able to streamline results. However, their results did not find any significant data one way or the other, that showed a bias for or against musical creation by AI. But Pasquier, himself, was persistent and later tested in 2023 with Zlatkov et al.

The authors of the new study in 2023, Zlatkov et al. expanded on these two previous studies and asked again if there was a general listener bias against AI generated music. This time they included a ranked survey for 30 musical pieces, allowing participants to put their answers in the order in which they liked. Unfortunately, the authors recognized the interface for doing this might have had flaws and was difficult to use. While larger than the previous two studies, only 163 participants were included. The results showed no significant bias for or against AI.

**Bell 2024 - Blueprint For Better Testing**

Previous research to determine exactly what public opinion is of music created by AI has left a lot to be desired without any concrete results. Much larger sample sizes are needed in order to get the public's opinion on such a hot-button issue and I suggest a series of adaptations to previous studies. First I would like to update the prompts given to AI modules to create AI

samples. Instead of asking for a “string quartet” like Pasquier et al. did in their research, I believe AI modules should be able to handle a more advanced prompt such as, “provide a 32-bar composition in the key of G minor with a 3/4 waltz feel” (Waltz). Give this prompt and others like it to both the AI module and the human hired to create their sample for the best comparison between the two. Another example of a more specific prompt that can be used for making samples in a new study: “provide a death metal EDM mashup song about unicorns” (Unicorn).

Secondly, better questions need to be asked of the participants during the study to obtain a more complete picture from their perspective. Instead of asking an audience if they perceive a piece to be creative, ask them where they think the music should be performed, or how many monthly listeners they think the artist has on streaming platforms. How many views did their most recent YouTube video receive? Asking more effective questions will uncover the true opinion. With more open-ended questions, answers will need to be parsed by AI for the fastes results, but human intervention is encouraged.

**Conclusion**

The position behind this research has shifted from a position of major bias against using generative artificial intelligence within the music industry to a much more open-minded and less-restrictive viewpoint. As this paper has outlined, there have been many valuable additions within the industry provided by AI and fighting AI’s use when implemented ethically is a difficult battle to win. Whether artists are using AI to help them with a chord progression or if AI is fixing a late artist’s last song, it’s clear that AI is here to stay.

Works Cited

Artist Right's Alliance. "200+ Artists Urge Tech Platforms: Stop Devaluing Music." *Medium*, 1 Apr. 2024, artistrightsnow.medium.com/200-artists-urge-tech-platforms-stop-devaluing-music-559fb109bbac.

Fair Use. *U.S. Copyright Office Fair Use Index*, 2024, copyright.gov/fair-use

GRAM, *Group Registration of Sound Recordings on an Album,* 2024, copyright.gov/eco/gram-sr/groups.html

Hammond, Kristian. "Musical Artists' Warning about AI Use," *Center for Advancing Safety of Machine Intelligence*, Northwestern University, 3 Apr. 2024, casmi.northwestern.edu/news/statement-on-musical-artists-open-letter-about-ai.html

Hight, Jewly. "AI Music Isn't Going Away. Here Are 4 Big Questions About What's Next." *National Public Radio*, 25 Apr. 2024, npr.org/2024/04/25/1246928162/generative-ai-music-law-technology.

Lemley, Mark A., and Bryan Casey. "Fair Learning." *Texas Law Review*, vol. 99, no. 4, 2021, pp. 743–85.

Office, *Compendium Of U.S. Copyright Office Practices* § 101, 3rd ed, 2017.

Ranwala, Dillon. "The Evolution of Music and AI Technology." *Watt AI*, 23 Jul. 2023, watt-ai.github.io/blog/music_ai_evolution.

Snapes, Laura. "AI Song Featuring Fake Drake and Weeknd Vocals Pulled From Streaming Services." *The Guardian*, 18, Apr. 2023, theguardian.com/music/2023/apr/18/ai-song-featuring-fake-drake-and-weeknd-vocals-pulled-from-streaming-services.

Quang, Jenny. “Does Training AI Violate Copyright Law?” *Berkeley Technology Law Journal*, vol. 36, no. 4, 2021, pp. 1407-1435, doi.org/10.15779/Z38XW47X3K.

Unicorn, “Unicorn Slaughterhouse,” *Suno,* suno.com/song/2470b68d-f335-4b87-94f5-99acb5c7c35e.

Waltz, “Moonlit Dance,” *Suno,* suno.com/song/ccd6abb9-c205-4e2a-a12c-a1ddc8b5c4b1.

Welch, Chris. “The Beatles’ final song is now streaming thanks to AI.” *The Verge*, 2 Nov. 2023, theverge.com/2023/11/2/23943290/now-and-then-the-beatles-new-song-ai.

Zipris, Ayal. “Illiac Suite.” *Illinois Distributed Museum*, distributedmuseum.illinois.edu/exhibit/illiac-suite.

Zlatkov, Dimiter, et al. “Searching for Human Bias Against AI-Composed Music.” *Artificial Intelligence in Music, Sound, Art and Design*, Springer Nature Switzerland, 2023, pp. 308–23, doi.org/10.1007/978-3-031-29956-8_20.