The Impacts of Artificial Intelligence on Music Making

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Aryan Medam

Spring, 2025

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Bryn E. Seabrook, Department of Engineering and Society

Introduction

Generative AI is defined by Nvidia, maker of the industry standard AI training chips, as allowing "users to quickly generate new content based on a variety of inputs. Inputs and outputs to these models can include text, images, sounds, animation, 3D models, or other types of data." (NVIDIA) Today software like Suno and Udio AI can create tracks this way, but still sound stilted and robotic. These models were likely trained on copyrighted music and often mimic an artist's style or voice. Legal disputes over the definition of fair use in this case continue to brew in U.S courts. The STS framework this paper will use is the Social Construction of Technology (SCOT) which argues that technology is shaped by human behavior and values as opposed to technology shaping human behavior. In that way decades of recorded music have come together to train these AI models.

Historically, generative music has been explored in the works of pioneers like Brian Eno, who would use probabilities to trigger notes and sequences to form drawn out ambient pieces. Similarly composers such as Motzart and Charles Ives would use a pair of dice to help them make musical descriptions (a practice known as aleatoric music). Both of these ideas rely on the knowledge of musical principles to determine a framework for composition. This results in questions about AI authorship of music considering the human "composer" is working with a set of weights and probabilities they have no access to and can not directly adjust. (Encyclopædia Britannica)

Previous papers in the subject have argued that imitation of classical pieces function as the highest barometer for AI "creativity" in the field of music. Traditional western art music (Bach, Beethoven, Haydn etc.,) are bound to more strict compositional forms and rules than more contemporary forms of music like jazz, modern classical or even electronic music.

Classical music also has less timbral diversity. As a result, it may be a good strategy for performing a Turing test where participants are asked to differentiate between AI generated classical pieces and actual classical pieces though it may not serve as a valuable tool for understanding how AI music might evolve into more irregular and unstructured forms. AI music making tools have also evolved and are now being trained on a larger set of data. While specialized models for specific styles exist, these general AI music making models allow us to observe a wider range of musical styles (usually with a pop slant).

STS frameworks that discuss the core idea of this paper include the social construction of technology that describes technology as a neutral actor that is imbued with society's values and use cases for it. Under this framework lies the idea of "interpretive flexibility" that suggests that society's values and perceptions affect a technology's design and meaning. (Xue) In this case the consumer's expectation that music is a largely cheap or free commodity may have led to companies like SUNO and UDIO training their models on publicly available music. Historically, methods of music consumption relied on the sale of vinyl records or CDs, but the proliferation of file sharing across the web (namely by napster) allowed the free distribution of music and eventually the rise of streaming. This raises the question of the perceived value of music. Spotify's own statistics reveal that in 2023 only 66,000 of its artists (totalling greater than ten million) made more than 10,000 dollars that year. (Spotify) The amount of artists who could reasonably make a living from the streaming platform alone are a much smaller subset of that. Especially, considering artists signed to labels would receive a smaller percentage of their streaming revenue and would divide that amongst producers, songwriters, and other rights

holders. This articulates the perceived value of music as bands and artists can expect to receive fractions of pennies from any single stream.

Therefore under the SCOT framework it becomes apparent that this lack of perceived value creates the social conditions for AI technology that indiscriminately trains on music across the web. The internet creates the consumer expectation for free art funded by ads or subscriptions and AI music startups extrapolate that idea when training their datasets. The idea that these companies must train their models on the open internet is a crucial one considering (given the limitations of AI as we understand it now) it is the only way they could produce a functional model.

I was unable to find a source that discussed SCOT and AI music generation programs but the "Analysis of Pinball's Purpose Surrounding and During NYC's Pinball Prohibition" discussed ideas around "interpretive flexibility" that is key in the argument about the perceived value of music. (Green) Additionally, the discussion around pinball was informative in determining how legislation could address the issues surrounding infringement when training on copyrighted works. In the case of pinball no distinction was made between machines made for recreation and those made from gambling. New York City's government banned the machines across the board rather than assessing each machine. Likewise it becomes difficult to assess what AI generated musical works may be infringing on another party's license.

The following research includes: querying programs to create songs in genres I am familiar with, sentiment analysis on internet forums, and discussing these ideas with music department faculty. Doing so will allow this paper to build a broad range of understanding in

terms of the ethics and future of AI generated music. There are some concerns about the timeframe in which this would need to be accomplished in order to complete the paper according to schedule.

Methodology and Frameworks

This paper analyzes various aspects of how AI could be involved in music by collecting information from various sources about how it could be used to create music and the various impacts that it has on different spheres of the musical landscape. This paper also analyzes specific use cases in which AI is used and the results it producers from different prompts. Additional discussion of the legal copyright frameworks and technical details of AI implementation are included to supplement the arguments about creativity, music and the broader philosophical implications of the subject. The keywords used to search for this information: AI, AI in music, AI and copyright law, AI and the music industry, and voice copyright protection.

Analysis and Results

Artificial Intelligence is transforming the music industry in profound and multifaceted ways—reshaping how music is composed, performed, distributed, and experienced. From the use of advanced machine learning models like GANs and Transformers to generate original compositions, to ethical debates around authorship, copyright, and voice cloning, AI challenges longstanding cultural, philosophical, and legal frameworks. While it democratizes access to musical tools and revives lost artistic voices, it also raises concerns about authenticity, creative labor, and algorithmic homogenization. All of which are ideas this paper explores.

On a technical level modern AI music systems are powered by machine learning techniques including Generative Adversarial Networks (GANs), Long Short-Term Memory networks (LSTMs), Transformers, Variational Autoencoders (VAEs), and diffusion models. According to Amazon Web Services (n.d.), GANs function through the interplay of a generator and a discriminator, enabling realistic synthetic audio production. Chen et al. (2024) detail how AI music generation can follow symbolic paths (like MIDI) or audio-based paths (waveforms), with hybrid models showing promise in capturing structural and expressive aspects. Such representations help models emulate various musical forms and human-like creativity.

Transformers, especially models like Music Transformer and MuseNet, are notable for their ability to model long-range dependencies in music, allowing AI to generate more coherent and structured compositions. Unlike LSTMs that struggle with longer sequences, Transformers maintain memory across extensive passages of music, making them particularly suited to classical and progressive styles. VAEs, on the other hand, are powerful in capturing the latent space of musical attributes, enabling interpolation between styles and genres.

Beyond architecture, training data and preprocessing are crucial. The bias embedded in training corpora—such as overrepresentation of Western tonal music or commercially successful pop songs—can skew AI-generated output toward normative structures, reducing diversity. Techniques like data augmentation, transfer learning, and reinforcement learning are increasingly employed to refine models, making them more robust and contextually aware. This technical foundation reveals that the artistry of AI music is often grounded in engineering decisions, underscoring the interplay between creativity and computation. Tools like SUNO and UDIO currently seem incapable of developing ideas that exist outside of the pop or commercial music landscape. As aesthetic criteria become more specific the models tend to revert to creating either

inoffensive pop pastiche or music that is unnatural androbotic. This effect is amplified by these platforms banning users from inputting artists' or bands' names into their prompt as a way to shield them from copyright violations.

Lee (2024) explores the rapidly evolving legal frameworks surrounding AI in music. With AI now capable of replicating an artist's voice and musical style, issues of copyright, identity theft, and fair use become paramount. Record labels are lobbying for protections against AI-generated impersonations of real artists, framing these efforts as essential for maintaining artistic integrity. Universal Music Group, as Variety (2023) reports, has taken steps to block AI models from accessing copyrighted material. The legality of training data usage remains unresolved and may shape the future of digital content rights. Current copyright laws, rooted in human authorship, struggle to accommodate works created by or with the assistance of non-human agents. The distinction between derivative and original works is blurred when AI generates content in the style of a specific artist. Moreover, the legal status of AI as an author or tool remains unclear. Should ownership be assigned to the user, the developer, or the AI itself? This ambiguity creates a legal gray zone that courts and policymakers are only beginning to navigate. This reflects on the perceived value of music which has tanked due to the rise of streaming services. These services (Spotify, Youtube, Apple Music, etc.,) allow users unlimited access to nearly any piece of recorded music for a monthly fee. Copyright infringement is a defense mechanism designed to protect the monetary value of original lyrics and melodies that are featured in a recording. They do not protect an artist's "musical style" which is often a product of an artist's influence. In this way this reflects the social construction of technology in the creation of AI systems that generate music. The lack of perceived value combined with the free access to training data contributes to the advancement of this application of generative AI.

The SCOT (Social Construction of Technology) when applied to this problem indicates that this technology exists because of social changes to the way music is consumed. In passive listening environments that well-streamed music encourages consumers may be less attentive to the quality of the work and therefore more adjusted to the possibility of using AI in these situations. Furthermore, the cost of music equipment and technical proficiency required to create a song has fallen with the rise of digital recording. This is another social condition that has implications for the lower perceived value of music. Well produced music can now be recorded from anywhere onto a laptop instead of the previous generation's dependence on studios, heavy amps, tape machines, and acoustic drum kits.

Despite advances, AI-generated music still grapples with limitations in emotional depth, long-form structure, and stylistic nuance. Chen et al. (2024) identify the need for better evaluation metrics, bias mitigation, and richer datasets. Future research may integrate interdisciplinary methods—blending music theory, cognitive science, and ethics—to refine AI creativity. The development of real-time, interactive systems holds promise for live performance and education, potentially reshaping how music is taught and experienced. One major challenge is the lack of interpretive intelligence in current models. While AI can generate technically competent music, it often lacks the narrative arc and emotional variability that characterize human compositions. The reliance on large datasets also introduces biases—models trained predominantly on Western or mainstream genres may fail to capture the intricacies of regional or experimental music. An example of this would be my tests with UDIO's music generation platform. When prompted to create a classical piano piece in the key of D minor it returned a robotic sounding pop song with stale lyrics about the piano in the G major. This showcases a lack of flexibility in these models to interpret musical context. The likely reason that the piece was in

the key of G major is that most pop music is written in the keys of G major, and C major (or their relative minor keys of E minor and A minor). These keys are relatively simple to perform across both piano and guitars in E standard tuning. Hooktheory, a service with a database of 30,000 popular songs for musicians to learn on their instruments, notes that D minor is the 10th most popular key with about 7 percent of all the songs in the platform being in this key. The keys of C major and G major, ranked number one and two by popularity, are represented at 16 and 13 percent respectively. This is also a literal representation of the SCOT framework where commonly accepted keys and musical structures that grew to popularity based on their ease of use for guitar and keyboard players are now being co-opted by software that strives to replicate the human experience of making music. AI programs have no understanding of what a key signature is (or any other musical concept) but through their training processes imitate the tendencies of human players.

Additionally, the choice to produce vocal music as opposed to focused instrumental music, even when prompted multiple times, provides further insight into the training data that was likely used and the use cases of those who utilize the platform. This bias toward vocal music indicates that much of this training data is centered toward vocal music to the point where parts of the prompt focused on instrumentation and timbre are interpreted lyrically. Essentially, an ideal prompt for these models would be: “Write an upbeat synth-pop song about the beach.” A prompt like this would tokenize into “tone adjectives” about the genre and a lyrically subject matter. In cases where genres are specific and have a large corpus: synthpop, modern-country, or LoFi. These tracks are often posted on platforms such as Youtube to engage some audience's desire for inoffensive background music. While this market was perceived as saturated, these AI platforms expedite this process and offer increasingly specific choices of musical wallpaper. This

illustrates a shift away from active listening into a space of "soundtracking" that has become more common with the shift away from physical media.

From an economic perspective AI may have an impact on the paid opportunities available for commercial musicians. With platforms like Spotify showcase the duality of democratization and centralization. More artists can upload music, yet only a small fraction benefit financially (Loud and Clear, 2024). Xue (2024) critiques Spotify's freemium model for prioritizing investor returns over artist compensation. As AI becomes more integrated into music generation, questions arise about who profits from AI-generated content—developers, platforms, or traditional artists? The last of which would be difficult to structure without explainable transformer models (not currently available). AI dramatically reduces the cost and time required to create music, which can empower independent creators but also threatens to devalue musical labor. Production libraries and content farms already deploy AI to mass-produce generic background music for ads, videos, and podcasts. This glut of supply may lead to oversaturation, diminishing the market value of original compositions and eroding income streams for human musicians. Live music, the lifeblood of the modern music industry in the streaming era, seems more resistant to change. As musicians continue to depend on their live shows for a greater portion of their income we may begin to see the actual recordings become vehicles for the live experience. Using generative AI to save money on the production of their music could allow artists more time and money on their live show.

Discussion

Lee (2024) and Berkowitz (2024) suggest the future of music must be guided by inclusive legal, ethical, and philosophical frameworks. This leads to questions about if the general public will develop deep emotional connections with AI generated music. Perhaps, reducing the perceived value of music by using AI to generate it will reduce the cultural relevance of the form. Choosing to explore the ramifications of technology from this perspective could lay the groundwork for a modern society's view on art in general. Limitations of this research are a result of the fact that it is difficult to understand the adoption of this technology at an early stage in its development and what the public's reception of it will be.

This leads to the concept of co-creativity, an interesting area for future research, in which AI acts as a collaborator rather than a competitor. In this model, the artist defines high-level goals, emotional direction, or stylistic parameters, while the AI generates raw material that may be edited, arranged, and refined. This dynamic represents a shift in creative agency: artists are no longer sole composers but directors of a computational orchestra. This democratizes music-making by lowering barriers to entry, enabling those without formal training to experiment with musical ideas through intuitive interfaces. However, this democratization is not without risk. As AI-generated content becomes more prevalent, the potential for creative homogenization increases. Recommendation algorithms already nudge listeners toward popular or formulaic content. When combined with AI models trained on these same datasets, there is a danger that the music ecosystem will grow more insular, reinforcing dominant trends while marginalizing experimental, local, or culturally specific expressions. Efforts to diversify training data and promote algorithmic transparency are therefore essential to ensuring that AI serves as a tool for expansion rather than conformity.

Bibliography

Amazon Web Services. (n.d.). *What is a GAN?* Retrieved from https://aws.amazon.com

Berkowitz, A. E. (2024). Artificial intelligence and musicking: A philosophical inquiry. Music Perception, 41(5), 393-412. https://doi.org/10.1525/MP.2024.41.5.393

Chen, Y., Huang, L., & Guo, T. (2024, September 3). Applications and Advances of Artificial Intelligence in Music Generation:A Review. arxiv. https://arxiv.org/html/2409.03715v1

Encyclopædia Britannica, inc. (n.d.). Aleatory music. Encyclopædia Britannica. https://www.britannica.com/art/aleatory-music

Green, E. (2024). Analysis of pinball's purpose surrounding and during NYC's pinball prohibition using the social construction of technology framework (STS research paper). University of Virginia.

Hooktheory. (n.d.). Song keys ranked by popularity. Hooktheory. https://www.hooktheory.com/cheat-sheet/key-popularity

Lange, C. (2020). *A study of artificial intelligence for creative uses in music* (Bachelor's thesis). University of Virginia.

Lee, E. (2024). AI and the sound of music. The Yale Law Journal Forum, 134. Retrieved from https://www.yalelawjournal.org/forum/ai-and-the-sound-of-music

Mycka, J., & Mańdziuk, J. (2025). AI in music recommendation. Neural Computing and Applications, 37(1), 801-839.

NVIDIA. (n.d.). What is Generative Ai?. NVIDIA. https://www.nvidia.com/en-us/glossary/generative-ai/

There are millions of artist profiles on Spotify, yet only a small fraction are generating money. shouldn't a higher percentage of all artists on Spotify be making money? aren't there 10 million of them? Loud and Clear. (2024, March 18).

Sherman, M. (2024, May 6). *With help from AI, Randy Travis got his voice back. Here's how his first song post-stroke came to be.* The Associated Press.

Loud and clear by Spotify. Loud and Clear. (2025, January 24). https://loudandclear.byspotify.com/#money-flow

Variety. (2023, April 13). *Universal Music asks streaming services to block AI access to its songs*. Variety.

Xue, D. (2024). *Free lunch? Spotify and the cost of "freemium" music* (Bachelor's thesis). University of Virginia.