EXHIBIT 135

Characterizing Detectability and Impact of AI-generated Music on Streaming Platforms

Stanley Wu
University of Chicago

Josephine Passananti
University of Chicago

Viresh Mittal
University of Chicago

Wenxin Ding
University of Chicago

Haitao Zheng
University of Chicago

Ben Y. Zhao
University of Chicago

Abstract

Since 2022, generative AI has produced large volumes of lower quality content, from websites to self-published ebooks to art. Early signs suggest the same AI slop is coming for the music industry. We perform a large scale measurement and experimental study to understand the growing phenomenon of AI-generated music on the Spotify streaming platform. We analyze two Spotify datasets, a publicly available metadata dataset representing 99% of Spotify tracks, and a crawled recommendation graph.

By identifying AI music using labels from Deezer, we study AI music on Spotify at scale. We find that the release of new AI-generated music will soon match and even surpass the volume of new music created by human. Yet the large majority (94%) of AI tracks is "slop" – they receive few, if any, listener plays and are rarely recommended. Unlike traditional human artists, AI musicians release much higher volumes of music across many genres. Next, we explore the music publishing pipeline by uploading our own AI music tracks through 11 indie music distributors, who then release them on streaming platforms. We find that distributors have inconsistent and largely unenforced policies on AI music, making it surprisingly easy for AI songs to appear on streaming platforms. Finally, we evaluate current approaches for detecting AI music using both automated tools and human judgement. Deezer's industry-leading detector shows high accuracy under standard conditions, but can be easily bypassed through simple modifications like compression. Using a large-scale user study (900+ participants), we find that typical consumers perform no better than random chance, while professional musicians are only slightly better. Together, these findings suggest that more work is needed on AI music detection to manage the growth of AI music slop. We propose an alternative approach driven by economic incentives.

1 Introduction

Recent analysis has shown that generative AI models have the ability to flood a medium with large volumes of arguably lower quality content, also referred to as "slop[1]" [6,31,43,61]. Recent progress in generative models for music suggest that a similar process might be occurring in the music space, where AI-generated music has already been linked to scams to generate royalties from fake accounts [33], mimicry of musicians who have passed away [38], and taking "inspiration" from existing musicians without credit, consent or compensation [30]. AI slop has also become a major complaint among customers on streaming services [14] and a topic of concern for major record labels [20]. These concerns have led to promises from Spotify and 15 music distributors to build AI disclosure mechanisms [65].

In the middle of this highly dynamic and chaotic ecosystem, we realize there is very little public understanding around AI music and its impact on society today, even on basic questions like its prevalence and popularity. Much of this is due to the fact that there is no understanding of how to even identify AI-generated music. A wide range of questions beckon, from the basic "how widespread is AI music?", to more detailed questions: is AI-generated music gaining traction with users on popular streaming platforms; how does AI-generated music enter the ecosystem; for those who want to label or filter AI-generated music, do we have ways to detect them, and are they robust to adversarial attacks? Answers to these questions quantify the risk to human musicians from the threat of AI-generated music in its many forms.

In this paper, we seek to answer these questions through a large measurement and experimental study, focusing on Spotify, the world's largest streaming music platform with 713 million monthly active users. To our knowledge, this is the first study to address these questions at scale, partly because of the challenge in identifying AI-generated music. For our study, we make use of two Spotify datasets, one containing metadata for 256M tracks (99% of Spotify tracks as of July 2025), and one containing a recommendation graph from which we extract a strongly connected graph of 33M tracks and 3.5B recommendation edges. To identify AI-generated songs, we

[1] https://en.wikipedia.org/wiki/AI_slop

DRAFT

Exhibit # Smith 7 04/09/2026 - B.R
Exhibits#: Nagaraj 318 3/4/2026
exhibitsticker.com

make use of AI labels provided by Deezer, a large streaming service based in France with significant research into detectors for AI-generated music. Our study has three main goals: a) understanding the current prevalence and adoption of AI music in today's streaming platforms, b) understanding the music distribution process, distributors' policies on AI music and their (non-)enforcement, and c) understanding the challenge of detecting AI-generated music for music consumers, professional musicians, and trained AI algorithms.
Our work produces a number of key findings.

First, we analyze the growth of AI music on the Spotify platform. AI-generated tracks were absent from new music arriving on Spotify in 2024, are estimated to surpass human-made music in new releases by early 2026, and are likely to dominate beyond that. AI musicians tend to only submit AI-generated songs and do so at a much higher volume than human musicians. Unlike traditional musicians, they also submit music across many genres, increasing total volume.

Second, we find that most AI music is much less recommended by Spotify in their track recommendation network. AI tracks make up 5.10% of all identified tracks, but only represent 1.19% of tracks in the strongly connected subgraph of our crawled Spotify recommendation graph. Consumers strongly prefer human-made music (human-made tracks average 3127 plays per month, while AI-generated tracks average less than 1/10th of that, at 232 plays per month).

Next, we explore the music distribution pipeline by generating new music using AI models, and submitting to 11 independent music distributors (5 of whom committed to AI disclosure). Regardless of their policies on AI music, most distributors approved our submissions for distribution (2 denied our submissions on grounds of violating platform policy). The large majority of our submissions made it to 6 large streaming platforms (Spotify, Apple Music, Amazon Music, Deezer, YouTube Music, and Tidal). Only 2 distributors labeled our submissions as AI-generated, highlighting the limited usage of AI detection tools and their ineffectiveness.

Fourth, we study the task of detecting AI-generated music. We first evaluate automated detectors, and confirm that Deezer's detection approach is by far the most accurate, but it is vulnerable to adversarial attacks. We also confirm that Deezer's production system is easily fooled by simple transformations like compression or pitch shift as well. We also perform a large scale user study (N=910), and find that regular listeners perform poorly at AI music detection, on par with a coin flip. Professional musicians do slightly better, at roughly 60% accuracy. Given these poor results, we propose an alternative approach that uses economic incentives to shape AI music distribution behavior, making them easier to detect.

Our results highlight the risks that large volumes of low-quality AI music pose to the music industry, and effectively managing AI music slop hinges on the challenging task of reliably identifying AI music. Without further progress on more robust detectors, music distributors and streaming platforms will be unable to enforce their AI policies, whatever those policies may be.

2 Background

In this section, we provide necessary background and context for our study, including information on AI music generators and detectors, music streaming platforms and distributors, and their polices on AI music.

2.1 AI Music Generators and Detectors

AI music generators. Modern AI music generators leverage either autoregressive [13, 18] or latent diffusion models [22, 24, 25, 66]. In this study, we only consider models that produce full-length songs including both lyrical and instrumental music, and do not consider models that synthesize short beats or melodies.

Very few open-source models are capable of generating high-quality, full-length songs with lyrics. We identified only two: DiffRhythm [66] and ACE-Step [25]. Both are diffusion-based, can run on a consumer-grade GPU, and generate complete songs in under a minute. Earlier models (e.g. SongGen [36], Stable Audio Open [22], Musika [41], Riffusion [24]) fail to meet these requirements.

Commercial AI music generators have grown rapidly, with five generators leading the space [51]. Suno [60] and Udio [63] can generate full-length songs, while Boomy [10], AIVA [4], and Soundraw [50] focus on beat sampling, stemming, and instrumentals, with limited functionality to generate full-length songs. For Suno and Udio, an annual subscription of $100 USD allows users to generate roughly 500 songs per month, each taking under a minute to generate.

Our study uses all four full-song generators: DiffRhythm and ACE-Step (open-source), and Suno and Udio (closed-source commercial). Each allows fast and cheap generation of AI music, a key factor that enables the fast, low effort generation of AI music slop [14, 20].

AI music detectors. Existing literature on AI music detection has focused on a binary classification of spectrogram data using convolutional neural networks [2, 3, 35] and transformer-based architectures to improve feature representation [42, 64]. However, their performance depends heavily on the audio datasets used for training. As we show later in §6, current detectors can vary significantly in their detection performance.

There are also commercial detectors only offering paid online services [7, 27, 39, 58]. They are closed-source, with little to no information on the underlying detection method and training data. Only SubmitHub [58] disclosed that it leverages techniques from an open-source detector [42]. At the scale of our large datasets, it was not feasible to evaluate these detectors, but we will consider that study for future work.

DRAFT

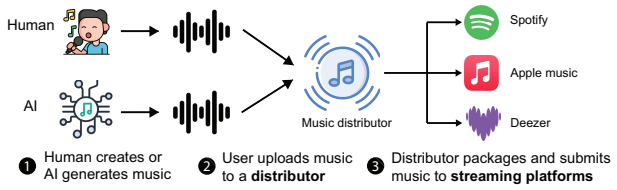


Figure 1: Full pipeline of how music (AI and human-created) are published to music streaming platforms like Spotify.

2.2 Distributors and Streaming Platforms

Streaming Platforms and Spotify. Today, streaming is the dominant method for music consumption, and accounts for 84% of revenue in the music industry [21]. Popular streaming platforms include Spotify, Apple Music, Amazon Music, Tencent Music, Deezer, Tidal, etc. Spotify dominates the market by volume (31%), more than twice that of the second-largest platform, Tencent. It also has the largest subscriber base, with over 200 million users [12, 21].

Musicians distribute their music across all streaming platforms because they earn royalties from each platform [47], which results in music catalogs across major streaming platforms overlapping significantly. Thus, we study Spotify as the largest and most representative streaming platform.

Uploading music via distributors. Today, major streaming platforms like Spotify and Apple Music only accept music uploads through approved distributors (see Figure 1). Distributors act as "middle-men" between musicians and streaming services, managing tasks such as audio formatting, metadata packaging, and collecting royalties [52]. In return, distributors are paid either by upfront fees, or by taking a percentage of the royalties earned. When musicians upload tracks to a distributor, they specify which streaming platforms to target. Later in §5, we explore the AI music distribution process by tracking the flow of AI tracks to streamers via 11 music distributors.

Our study considers music distributed from a wide range of distributors. Major record labels serve some of the most popular musicians, but only a small portion of all music by volume. Studies show that the large majority of music (96%) is distributed outside the major labels and their subsidiaries [28]. Thus, our study focuses on music distributors (rather than large record labels) as the backbone of the music ecosystem.

Distributors' policy on AI music. Distributors vary in their policies on AI-generated music. Only a few explicitly prohibit AI music. However, many have proactively deployed AI-detection systems to enable quality control and enforce copyright compliance. For instance, two distributors, Amuse and Ditto, collaborate with SubmitHub [59], a commercial AI music detection service.

Streaming services' policy on AI music. Notably, Deezer is the only streaming platform that publicly labels the AI music they host [15, 34] using an in-house detector. According to their reports, 39% of their daily incoming music is fully AI-generated, and 85% of AI tracks' "streams" are fraudulent [16]. Other platforms such as Spotify have also taken actions against AI-generated music spam. Spotify announced in 2025 that the platform has already removed over 75 million 'spammy' AI tracks, and that they are working on AI disclosure in collaboration with 15 distributors [55].

3 Goals and Methodology

Our study seeks to understand, at scale, the current and potential future impacts of AI-generated music on the music industry. We are particularly interested in quantifying its prevalence and popularity on the largest streaming service (Spotify), and how the challenges of detecting AI music will impact efforts to regulate and label it. Next we describe the key questions, our methodology, and the datasets we used in our study.

3.1 Key Questions

Our study is focused around several key questions.

Q1 How much AI-generated music is available on Spotify? How fast is it growing relative to new music created by human artists? Do AI musicians behave fundamentally differently from human musicians? What is the economic impact of these AI-generated songs in terms of streaming revenue?

Q2 What is the perceived quality of AI-generated music? Are they popular with listeners, or do they exhibit the same low-quality, high effort characteristics common to AI slop in other mediums? Are AI-generated tracks recommended by Spotify's advanced recommendation systems?

Q3 What is the process for distributing AI-generated music? What are current distributors' policies for AI-generated music, and are they successful in enforcing them?

Q4 Enforcing AI music policies require identifying them at scale. Are there reliable methods today for detecting AI music? Can human listeners identify AI music when they hear it? How well can automated detectors perform?

3.2 Methodology

We answer the above questions through a large measurement and experimental study focusing on the largest streaming platform (Spotify, see §2.2 for background). Getting access to large-scale data on AI music is challenging. Existing datasets of platforms such as MusicBrainz [57], last.fm [8], and FMA [17], were compiled from open libraries prior to the emergence of generative AI. As a result, we use a recently released dataset of Spotify metadata, along with a dataset of Spotify recommendation links we crawled (more details next in §3.3). We cross-reference tracks with labels of AI music on Deezer to label AI tracks on Spotify, and use the results to

DRAFT

analyze the growth, popularity of AI music, characteristics of AI musicians, and their role in Spotify recommendations.

Real-world experiments on publishing AI music. To obtain a deeper understanding of how AI music enters and propagates inside streaming platforms, we run experiments in which we publish AI music through 11 music distributors, and observe the outcomes. We observe distributors' current policies on AI music, their enforcement of those policies, and the resulting role they play in the growth of AI music today. The results also shed light on potential controls on AI music moving forward, and suggest what roles AI detectors and economic factors might play.

Understanding the challenges of detection using user studies and experimental studies. We study the problem of detecting AI-generated music from both a human and algorithmic perspective. Using a variety of human-made and AI-generated music, we perform an IRB-approved user study with 910 participants (250 general users, 600 motivated volunteers, and 60 professional musicians) to examine whether human listeners of different music backgrounds and experiences can distinguish between the two. We also test the reliability of 4 automated AI music detectors under both normal and adversarial conditions.

Given the limited efficacy of both human-based and automated AI detection, we suggest a possible alternative for robust detection leveraging economic incentives associated with publishing and streaming music.

3.3 Spotify Datasets and AI Labels

We now describe details of the two Spotify datasets used in our study, and how we label the AI music tracks in each.

Spotify's Music Metadata adapted from Anna's Archive (SMMA). In December 2025, Anna's Archive released a dataset [5] containing metadata of 256 million song tracks, claiming to cover 99% of Spotify's music catalog as of July 2025. The dataset contains metadata for each track, including its Spotify ID, International Standard Recording Code (ISRC), total play count, release date, artist, and album name. We downloaded this metadata dataset without accessing any audio files. Using the ISRC (a universal identifier for recorded music) to remove duplicated entries, we identified 185 million unique tracks, of which 53 million were released after 2024. We refer to this dataset as SMMA.

Spotify's Core Recommendation Graph (SCRG). The SMMA dataset does not capture relationships between tracks or how users might discover music through their listening patterns. For this, we collected a separate dataset of recommendation links from Spotify's recommendation network using their recommendation API. This API is publicly accessible without account credentials or authentication[2]. Using this API, we obtained the largest strongly connected component of Spotify's recommendation network up to November 2025 (details on our crawl are listed in Appendix A.1). In this graph, tracks are represented as nodes, and their recommendation relationships are defined by directed edges. In total, we have 33 million tracks and 3.5 billion edges, as well as the metadata for each track (similar to SMMA). We refer to this dataset as SCRG. Notably, 99.4% of tracks in SCRG are also in SMMA. 0.28% (97k tracks) were posted after July 2025, thus are missed by SMMA. We do not know the origin of the remaining 120k tracks, though they represent a tiny (0.05%) portion of the Spotify catalog.

Dataset	# of tracks	# of tracks posted on/after 2024
SMMA	185,361,997	52,915,713
SCRG	33,512,395	6,804,105
Combined	185,578,495	53,042,110
w/ label	-	40,748,961

Table 1: The volume of Spotify tracks captured by each of the two datasets. We also report the tracks posted on/after 2024 because the two major AI generators, Suno and Udio, were released in early 2024. For these tracks, we identify their label (AI-generated vs. human-created) by matching the track on Deezer and using Deezer's AI label.

Table 1 summarizes the two datasets. In the rest of the paper, unless otherwise stated, we use the combined dataset to analyze the entirety of Spotify's music catalog.

Labeling AI music using Deezer labels. Neither dataset includes labels indicating whether a track is human-made or AI-generated. Testing each track using AI music detectors would have required us to download music, which would involve violating artists' copyright in the process. Instead, we propose leveraging Deezer's existing labels of AI music. Deezer is the only streaming platform that explicitly labels AI music, and does so using a proprietary detector developed by their research team. As we show later in Section 6.1, Deezer's detector is also the most accurate AI detector in our tests. We identify and label AI music tracks in our two datasets by cross-referencing its corresponding label on Deezer[3] in Deezer.

We limit our matching based labeling to tracks released on/after January 2024, since it predates the arrival of the advanced commercial AI music generation models: Suno and Udio. Filtering by this timeframe, we have 53 million unique tracks introduced to Spotify after January 2024. We were

[2] We discuss our use of this API, ethical considerations, and disclosure in the Ethical Considerations section.

[3] We note that Deezer reports the AI label for each album rather than each track. After studying key artists who have adopted AI, such as will.i.am [44], we found that artists who use AI to generate music tend to apply it consistently across the entire album, rather than mixing AI-generated tracks with human-made songs. Therefore, we label all tracks on an album using the album's label in Deezer.

DRAFT

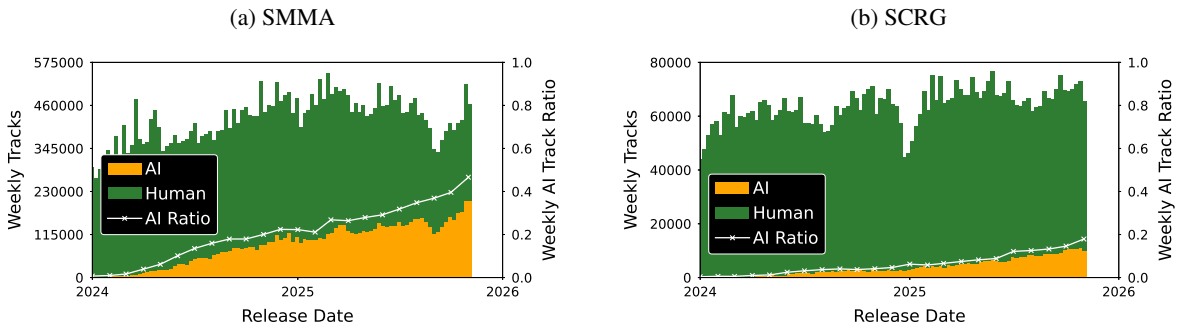


Figure 2: Stacked volumes of weekly releases of AI-generated and human-made tracks for (a) SMMA (overall music catalog) and (b) SCRG (strongly connected tracks), and the ratio of AI music. Both datasets indicate a strong growth of AI music since 2024.

able to definitively match more than 77% of these tracks with Deezer (40.7 million), where each track is labeled as AI-generated or not (details on how we performed matching in Appendix A.2). When we studied the remaining 12.3M tracks that were not found on Deezer, most were tracks from countries where Deezer was not available or blocked, e.g., India, Belarus, Russia, and Sudan. We believe the 40.7M matched songs represent a nearly complete set of music tracks in Western countries released since 2024, and use it as the primary target of our subsequent analysis.

4 Prevalence, Growth & Behavior of AI Music

In this section, we leverage the two Spotify datasets (SMMA and SCRG) to study the prevalence, growth, and perception of AI-generated music and its economic impact. We also examine whether (and in what ways) AI-generated music differs from human-created music, and how the two "interact" with each other on the streaming platform's recommendation system. Our analysis in this section seeks to answer the first two questions (Q1, Q2) listed in §3.1.

Overall, we find that AI-generated music has been growing rapidly on Spotify since 2024, and is likely to surpass human-made music in new releases in early 2026. Yet a majority of AI music is rarely or never played by listeners and generate no revenue, meaning that the current growth is likely driven by AI slop. AI musicians consistently upload a larger volume of AI-generated music across many different genres, hoping to attract exposure and obtain royalty payouts. Finally, AI music is much less recommended by Spotify's recommendation system compared to human-made music, though it is getting slowly embedded into mainstream listening.

4.1 Growth and Popularity

We start with a longitudinal analysis on the growth of AI music on Spotify, using the labeled data of the two Spotify datasets (SMMA and SCRG). As discussed earlier, SMMA represents Spotify's overall music catalog while SCRG captures its strongly connected tracks.

Significant AI growth since 2024. Figure 2 plots, for both datasets, the amount of AI tracks released on Spotify every week since 2024, stacked with human-made tracks in the same time frame. Both indicate a steady growth of AI music.

In terms of the overall music catalog (SMMA), AI music is growing at a tremendous speed. Less than 1% of weekly new releases were AI-generated in January 2024, a figure that has increased beyond 40% by Nov. 3rd 2025, the last full week in which we have data (Figure 2a). In that last week, 202,046 AI tracks were released, compared to 238,646 human tracks (60,022 tracks remain unclassified, since they were not found on Deezer). At this pace, AI tracks has likely already surpassed human-created music in new releases in early 2026. Overall, AI-generated music already accounts for 5.1% of the entire music catalog, despite being around for only the last 2 years.

The growth of AI music is much more muted on SCRG, the strongly connected recommendation graph (see Figure 2b). There, AI-generated music accounts for only 1.2% of the SCRG tracks. Since tracks on SCRG capture what the streaming platform predicts listeners will prefer (which triggered the platform recommendations), this observation suggests that listeners still favor human-made music.

AI music slop. We confirm the above observation by comparing the play count for AI and human-made music. Interestingly, 93% of AI music tracks received less than 1,000 plays since release, the minimum threshold set by Spotify for monetization [53]. Even for AI music released more recently (i.e. between June 2025 and November 2025), 92% have less than 1,000 plays. In contrast, only 78% of human-made tracks display such behavior. We believe this is a strong indicator that most of the AI-generated music behind Spotify's explosive growth (and likely other streaming platforms) amounts to little more than AI slop.

This observation in music streaming mirror trends in other domains such as text and image, where large volumes of low-

DRAFT

quality, low-effort, AI-generated content are being injected into platforms purposefully to capture user engagement and monetization through sheer volume.

4.2 Streaming Royalty and Its Distributions

Next, we study the economic impact of AI music on streaming platforms by analyzing the streaming royalties earned by AI tracks and studying whether (and how) their royalty patterns differ from those of human-created music.

Royalty earned by AI music. Since Spotify does not make royalty earnings of individual tracks public, our datasets do not have such information by default. However, based on Spotify's official documentation [54], royalty payouts are derived from play count via a monthly streamshare model. For each track, Spotify calculates the ratio of the track's play count in a given month to the platform's total play count for all tracks for the same month. Royalties are then distributed proportionally from a total pool based on these ratios.

To perform this royalty calculation, we need data on the monthly play counts for tracks on Spotify. However, our two datasets (SMMA and SCRG) only record the accumulated play count at the time of crawl. Thus, we performed two additional crawls on Spotify for all tracks (AI or not) found on the combined SMMA and SCRG set, and obtained their monthly play count for January 2026 by subtracting their accumulated play counts spaced out by a month. Note that, any track with less than 1,000 total plays is considered negligible on Spotify, and does not earn royalties. Spotify displays playcounts for these tracks as 0. We also mark these tracks as having "negligible" playcount.

Table 2 lists the statistical breakdown of tracks by their play counts, separately for AI and human-made music. Similar to our finding in §4.1, human-made tracks are much more popular than AI tracks. In fact, a majority of AI tracks (94%) received negligible plays during this entire month, i.e., 0 plays that contributed to royalties. In comparison, 69% of human-made tracks received negligible plays. Furthermore, while AI tracks averaged only 232 plays that contributed to royalties in January 2026, human-made tracks averaged 3,127 (>10x).

Using this monthly play count data, we compute the royalties earned by each track. In total, AI music contributes to 0.3% of the monthly royalty payout. Assuming the total royalty equals the average monthly royalty in 2025 reported by Spotify [56], this means roughly $3.5 million USD.

Inequality of royalty distributions. The music industry is known to be highly unequal, where a small number of extremely popular songs capture the majority of revenue/streamshare. We seek to understand whether such inequality also exists among AI music.

Figure 3 plots the cumulative distribution of royalties earned in January 2026 for the most popular AI tracks, which we define as those with at least 1 non-negligible play (and thus earn royalties). This subset only represents the top 6.27%

	% of tracks based on # of plays in Jan. 2026			
	negligible plays	≥1k	≥100k	≥ 10mil
AI-generated	0.94	0.03	0.0003	0
Human-made	0.69	0.05	0.003	0.00004

Table 2: We report the ratio of tracks under different buckets of play count during January 2026, separated by AI and human-made music from our combined Spotify dataset. The royalties earned by each is proportional to their play count.

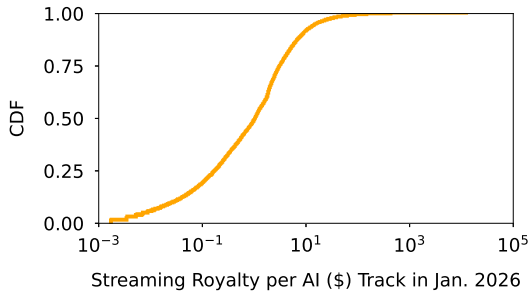


Figure 3: CDF of royalties earned by the most popular 6.27% AI tracks in Jan. 2026 (the remaining 93.63% did not receive enough plays to earn royalties, i.e., <1k total plays). Even in this subset, the majority of AI tracks earned less than $10.

of all AI tracks, and even within this subset, 92% earned less than $10. Only 267 total AI tracks (0.05%) earned more than $1,000.

We further measure the degree of inequality using the Gini Coefficient [19], where a score of 0 means perfect equality and 1 means maximal inequality. Interestingly, the Gini Coefficient for non zero-play AI tracks is 0.89, considerably lower than that of human-made tracks (0.98). This suggests that while it is extremely difficult for human musicians to break into the top earning tiers, AI users are more likely to reach top rankings when competing against their AI peers.

Causes of inequality. For human-made music, it is known that a key reason behind streamshare/royalty inequality among musicians is ***repeated engagement*** [1]. Listeners tend to return to listen to artists they already discovered and like, thus musicians who have released one or more hit songs are more likely to experience continued success in future releases.

Interestingly, we find that this pattern of repeated engagement is less prevalent in AI music compared to human-created music. We measured this pattern by comparing the per-track play counts of a musician's tracks before and after the release of their first top hit. A track is defined as a top hit if its total play count falls within the top 1 percentile of all songs on the platform. For each artist with at least one top hit, we divide their songs into two groups: those released before the first top hit, and those after. For each track in group 1, we measure the average play counts per day, from its release to the date of

DRAFT

# of human musicians	# of AI musicians		
	AI-only	switched	mixed
14,198,574	268,520	19,432	44,601

Table 3: The number of human musician and the three types of AI musicians found in our combined Spotify dataset.

the top hit. For each track in group 2, we measure the average play counts per day, since its release. We then compare the median value of group 2 to that of group 1.

For human musicians, the median number of plays per day increases by 16x for tracks after the first top hit. This confirms that releasing a top hit boosts the success of future releases. For AI music, the boost reduces to 5x. This result, combined with our earlier finding on royalty distribution, suggests that success of top AI music tracks is much more random than human-made music. This raises the question of whether AI musicians behave differently as a result, which we study next.

4.3 Characteristics of AI Musicians

We refer to an artist as human musicians if they have not released any AI music, and as AI musician if they released at least one AI-generated track. Intuitively, there are three relevant types of AI musicians: *AI-only* (who only release AI-generated tracks), *switched* (who used to release human-made music and then switched to AI-only), and *mixed* (who release both human-made and AI-generated tracks),

Table 3 showcases the number of human musicians compared to the three types of AI musicians found in our combined Spotify dataset. One key finding is that AI musicians tend to use AI to generate music whenever possible: the majority of AI musicians (81%) are AI only, while only 5.8% were previously human musicians who flipped to AI-only. Even for the mixed musician, once they released the first AI album, their subsequent releases are dominantly AI (median=88%).

Next, we continue our examination of AI musicians by comparing them to human musicians.

Track count and upload frequency. We first study the behavior of music uploads. Since 2024, each AI musician has uploaded 29 AI tracks in average, double the average volume of a human artist (14 human-made tracks). Additionally, the average gap between releases is a mere 16 days for AI musicians and 50+ days for human musicians. Clearly, AI musicians are more "productive" than human musicians.

Album size. In terms of the number of tracks on each album, human and AI musicians are quite similar to each other. The average track count per album is 2.7 for human musicians and 3.1 for AI musicians. This suggests that while generating a full-length song takes only minutes, AI musicians still follow common album sizes in their submissions.

Genres. Human musicians often focus on a single genre in their releases, as mastering multiple genres at the same time is challenging. This is not the case for AI musicians. When examining the genre diversity of AI musicians, we find that they cover much more distinct genres than human musicians. The most diverse (i.e. top 1%) AI musicians have each covered more than 12 different genres (some even reach 57 genres), twice than that of human musicians (6 genres).

We believe this is because today's music generators can produce full-length tracks across many genres. As a result, AI musicians often generate a large and diverse catalog of music to increase the likelihood that some tracks will attract streams and generate royalty payments.

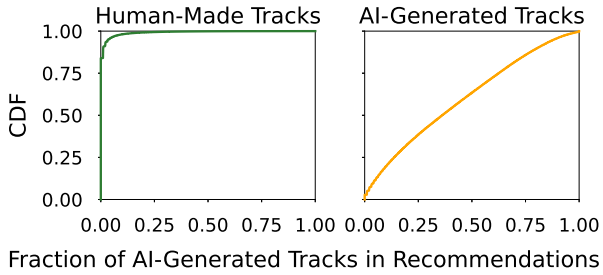


Figure 4: In Spotify's strongly connected recommendation graph, 84% of human-made tracks never recommend any AI tracks; 97% of AI tracks recommended other AI tracks.

4.4 Accessibility of AI Music via Recommendations

Finally, we study how each AI music track "connects" to other tracks (AI or human-made). In Spotify, music tracks are "connected" through a recommendation network [23]. After a user listens to some tracks, Spotify actively recommends additional tracks, which is a core feature of the platform. We are interested in understanding how AI-generated music are integrated into this recommendation network and then discovered by listeners.

We focus our analysis on the SCRG dataset, which represents the largest strongly connected component of Spotify's recommendation network up to November 2025, with 33 million tracks and 3.5 billion edges. Here each edge is directed: an edge $a \to b$ means that after a user listens to track a, Spotify had recommended the user to listen to track b. For simplicity, we call a **recommended** b if $a \to b$ appears in SCRG.

Recommending AI music. AI-generated tracks make up 1.2% of the tracks in SCRG. The fact that they appeared in the SCRG graph means that AI-generated tracks were recommended to listeners. But we are interested in what tracks would trigger recommendation of AI tracks. Figure 4 plots the cumulative distribution of the fraction of AI tracks being recommended by each track in the SCRG graph, where we report separately for human-created tracks (left figure) and

DRAFT

AI-generated tracks (right figure).

Interestingly, the recommendation behaviors of human-created and AI tracks are very different[4]: 84% of human-created tracks have never recommended any AI track, while 93% of AI tracks have recommended other AI tracks. Furthermore, each AI track has recommended 36 other AI tracks on average, mapping to nearly 40% of its total recommendations. Human-made tracks, however, only recommended 0.73 AI tracks, mapping to 0.8% of its total recommendations.

Such a stark contrast between AI- and human-made music is not entirely surprising, given that AI-generated tracks often share similar characteristics, particularly when they originate from a small number of leading generators.

What is surprising is that 16% of human-made tracks have recommended at least one AI track, despite AI tracks count for only 1.2% of the total tracks. This suggests that AI music is not a minor, isolated bubble, but is getting gradually embedded in mainstream listening.

Simulating listening behavior via recommendations. Because Spotify's recommendations encourage users to discover previously unheard tracks, we want to understand whether this system may eventually direct users to AI-generated music. Thus, we run a simulation, using the observed recommendation graph, to estimate what could happen when 500,000 listeners blindly follow non-personalized music recommendations in passive listening. Specifically, we simulate a one-hour listening session for each listener by starting with a track randomly sampled from the top 20% of most played human-created tracks. The "listener" then randomly chooses the next track from the current track's list of recommendations, until it has listened to 20 additional tracks.

Our results show that 4.40% of these simulated listening sessions contain at least one AI track. But more importantly, after listening to one AI track, the probability that the next song will also be an AI track is 39.39%. **This is alarming, because once a user listens to an AI-generated track, the streaming platform is likely to recommend (and play) additional AI tracks shortly after.**

5 Understanding AI Music Distribution

Our static analysis of AI-generated music on Spotify shows that it does indeed exhibit typical characteristics associated with AI slop: low-cost, high-volume, low-engagement content. As this content continues to grow, the consequences will likely mirror that of AI slop in other modalities, overwhelming volume that floods marketplaces, enabling scams while making quality content difficult to locate [11, 29, 37].

[4]These results should not be interpreted lightly, as Spotify's recommendation system is powered by a complex, machine-learning model that accounts for many factors, such as similarity in audio characteristics, artists, and users' listening histories and habits. In light of this, we perform additional analysis to rule out similarity in music genre as the leading cause, which we detail in Appendix A.4.

Empirical evidence and our analysis both show that AI can generate music that is popular and achieves hit status with listeners. Unfortunately, as low quality AI music floods streaming platforms, high quality AI tracks would be buried in the noise along with songs performed by human musicians.

This leads to the question, what kind of policies or actions in today's music ecosystem could realistically address the AI music slop problem? To answer this question, we must first understand the role of stakeholders in today's music distribution network, and how they are impacting the flow of AI music today. We do so through *proactive* experiments where we generate a number of new AI tracks, and submit them to a variety of music distributors. By observing the progression of these tracks through the music publishing pipeline, we gain insights into the roles distributors, streaming platforms and their policies can play in limiting AI music slop.

5.1 Generating and Uploading AI Music

Next, we describe our experiences as we generate and attempt to publish our own AI songs onto streaming platforms.

Selecting music distributors. To distribute music to streaming platforms, we need to choose a set of distributors to work with. From the list of 18 distributors identified by indie musicians [26], we exclude those that required ID verification or exclusive partnerships, leaving 11 distributors to work with. This list includes both Ditto and TuneCore, distributors that have worked with well-known independent artists like Chance the Rapper and Russ. For each distributor, we identify their policy regarding AI-generated music (explicitly allow or prohibit). We show this information in Table 4, along with cost for distribution and any required sharing of royalties.

Generating AI tracks. For each music distributor, we created a different artist persona. All that was required was a name, which we generated using ChatGPT. Next, for each persona, we use 4 different generators to create 8 total AI songs. As mentioned in Section 2.1, we use the two most popular commercial generators (Suno and Udio), and the two highest-quality open-source generators (DiffRhythm and ACE-Step). For each generator, we generated an instrumental and lyrical song, using random text input prompts, and lyrics generated using LLaMA 3.1. Further details on the generation process can be found in Appendix A.3. We submitted each generated track as an album single (album with one track), and used FLUX [9] to generate an album art cover. To meet the resolution requirements for distributors, we used SDXL's 4x image upscaler [46]. Based on our experience, our choices of persona, album size and album art are consistent with those of most AI musicians.

We upload a *unique* subset of songs to each distributor. This is because streaming platforms like Spotify and distributors like DistroKid reportedly employ audio fingerprinting to ensure the same songs are not duplicated on their platform [45]. Rejections based on overlapping audio fingerprints would

DRAFT

Distributor	Cost	Royalty Tax %	AI Music Allowed	Detected AI	# AI Songs Approved	Review Time
CD Baby	$9.99 / album	9%	N	0/8	8/8	14 days
Horus Music	$23.03 / year	0%	N	0/8	1/8	3 days
Offstep	$18.00 / year	0%	N	0/8	8/8	2 days
Soundrop	$4.99 / album	15%	N	0/8	8/8	6 days
TuneCore	$22.99 / year	0%	N	0/8	8/8	5 days
Amuse	$23.99 / year	0%	Y	4/8	8/8	2 days
DistroKid	$26.55 / year	0%	Y	0/8	8/8	1 day
Ditto	$19.00 / year	0%	Y	0/8	8/8	1 day
Landr	$23.99 / year	0%	Y	6/8	8/8	5 days
Routenote	free	15%	Y	0/8	0/8	1 month
UnitedMasters	$19.99 / year	0%	Y	0/8	8/8	5 days

Table 4: Our findings after uploading AI songs to 11 music distributors. Notably, even distributors that prohibit AI music (CD Baby, OFFstep, Soundrop, and TuneCore) approved AI songs and sent them to music streaming platforms.

bias our results.

Uploading and observing. For all 11 music distributors, we manually upload every track, with a maximum of 2 per day per distributor (to avoid spam detection). We selected 6 different music streaming platforms as destinations for our music: Spotify, Apple Music, Deezer, Amazon Music, YouTube Music, and Tidal. These represent a majority of the most popular music streaming apps [21]. We performed these experiments between November and December 2025, and manually verify in January 2026 whether each track successfully made it to the target platform.

5.2 AI Policies and Enforcement

First, we report what we learned about distributors' and streaming platforms' policies on AI music, and possible enforcement.

Distributors disagree on AI policies. Uploaded music must be approved by distributors before they can be sent to streaming platforms. As we see in Table 4, current music distributors are split on whether users should be allowed to upload fully AI-generated songs. All 11 distributors we evaluated mentioned AI usage in their user policies, but varied significantly in what they mention. Some disallowed any use of AI, while others disallowed only fully generated AI songs. All distributors require uploaders to fully own the rights to the songs they upload, but there is no consistent definition of how that applies to AI-generated songs. The only consistent AI policy we found is prohibiting the usage of AI to impersonate existing musicians.

Streaming platforms lack stance on AI. Once music is approved by distributors, it still must be approved by the streaming platforms themselves. None of the 6 major platforms we tested (Spotify, Deezer, Apple Music, Amazon Music, YouTube Music, Tidal) mention whether fully AI-generated songs are allowed on their platform. However, Spotify and Deezer have taken steps to address AI music. Deezer now labels AI music and demonetizes fraudulent AI content, while

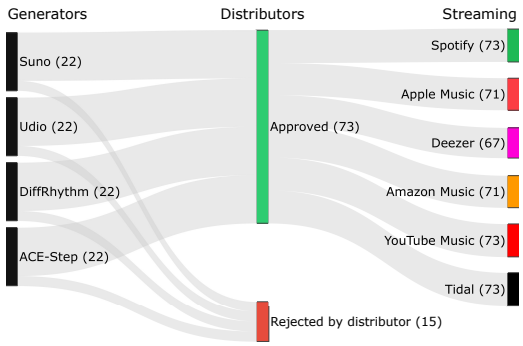


Figure 5: Results of our end-to-end AI music uploading. Rejections mainly occurred at the music distributor. Once songs get approved by distributors, they are rarely rejected by streaming platforms.

Spotify claims to have removed millions of "spammy" tracks.

Enforcement. We found mixed evidence on enforcement of these AI policies. None of the 4 distributors who prohibit AI submissions explain how they enforce it. In fact, in our set of 11 distributors, we only found evidence that Amuse and Ditto use an AI music detector, but both allow AI music.

5.3 Results of Our AI Music Uploads

Distributor approval. We list the results of our uploaded AI music tracks in Table 4. Out of a total of 88 tracks uploaded, a total of 15 songs were rejected by 2 distributors (Horus and Routenote). There was no clear impact across the 4 different AI generators used. Either all are approved by the target distributor, or almost all are rejected. Notably, 4 out of the 5 distributors that explicitly prohibited AI music uploads approved all AI songs we uploaded. When rejecting our uploads, Horus Music and RouteNote did so under a blanket claim that our songs did not meet their platform's requirements. Neither marked the rejected songs as AI-generated.

Lack of reliable AI detection. Based on their documentation, all distributors but one (DistroKid), employ a review process that involves review by a moderation team. DistroKid uses a fully automated review pipeline. Two distributors that prohibit AI music (CD Baby and OffStep), explicitly specify that they use humans to review every submission. Still, only two distributors (Amuse and Landr) identified a subset of our uploads as AI-generated. Recall from Section 2.2 that Amuse and Ditto are known to work with SubmitHub, a commercial AI music detector. Yet Amuse detected only 4 out of 8 tracks as AI, and Ditto detected none. Detection clearly is a significant hurdle to managing AI music, and we study it in detail next in Section 6.

Streaming platforms rarely rejects tracks approved by distributors. In Figure 5, we show how many of our AI tracks appear on streaming platforms after receiving approval

DRAFT

by distributors. Spotify, YouTube Music, and Tidal accepted all songs approved by distributors, while Apple Music and YouTube Music each rejected 2 songs (out of 73). While we received no detailed reasons for these rejections, we infer it was because these songs did not pass Apple/YouTube's platform-specific quality assurance tests.

Although it looks at first glance like Deezer rejected 6 of our tracks, a closer look revealed that Landr was responsible. Landr incorrectly assumed that Deezer does not accept AI-generated music, and thus did not send them the 6 tracks it identified as AI-generated, despite approving them for distribution to other platforms.

Conclusions. Our firsthand experiences creating and uploading AI music reveal a number of insights into the growth of AI music. The current music publishing pipeline, both in distributors and streaming platforms, have very few mechanisms designed to manage AI-generated music. Today, it is extremely easy (and low cost) to generate and publish large volumes of AI-generated tracks. Simple calculation show a maxed out Suno membership would allow 6000 tracks per year for less than $0.02 per track (§2.1), and all of them could be distributed for an annual fee of $26 or less (Table 4). The grand total would be $126 per 6000 tracks.

While AI music is still evolving, the possibility of low-cost and low-effort AI music generation means many will flood the streaming platforms in the hopes of generating a (rare) AI hit that pays off in royalties. This is a similar incentive model that drives and sustains large-scale spam campaigns [32].

6 AI Music Detection

One of the key insights from our experiments uploading AI-generated music (§5) is that current distributors are unable to reliably detect AI music. For any policies regarding AI music to be enforceable (by distributors or streaming services), the critical component is a reliable AI music detector.

In this section, we study this problem in more detail and evaluate detection of AI music by both automated tools and human listeners with different backgrounds and experiences. We begin by evaluating the detection efficacy (under both benign and adversarial conditions) of four open-source detection methods, including one that is the basis for Deezer's commercial detector (see §2.1). We also test the robustness of the commercial Deezer detector using end-to-end tests. We then present results from a large scale user study (N=910) evaluating the ability of general users, volunteers, and musicians to identify AI-generated music.

Our results show that AI music detection is very much an open problem. While one approach (Deezer's Fourier detector) shows high accuracy under benign conditions, it is easily bypassed by applying simple preprocessing operations to the music file, including compression, pitch shift, reverb, and AI-reconstruction. Humans fare much worse, with most users showing similar accuracy as a coin flip, and professional musicians doing marginally better at roughly 60% accuracy.

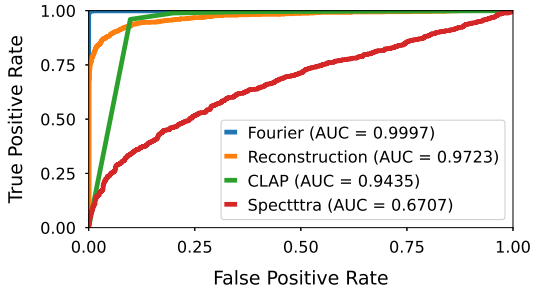


Figure 6: Performance of AI music detectors shown as ROC curves. The Fourier-based detector (used by Deezer's detection system) is nearly perfect at distinguishing AI songs generated by Suno/Udio from human-created songs.

6.1 Automated Detection

We first examine the accuracy of ML-based AI music detectors and consider factors that are most likely to affect robustness. Our experiments test against multiple music generators and apply audio transformations that can act as adversarial attacks. Adversarial attacks are conducted locally against 4 SOTA open source detectors first, then extended to testing on Deezer's commercial detection system.

Dataset and methodology. For our test set, we randomly sample and download a total of 1,000 songs generated by two most widely used music generators, Suno [48], and Udio [49]. They are the only commercial models[5] producing full-length songs. For human-made music, we randomly sample 1,000 songs released before 2021 and download 30s previews.

We evaluate four SOTA ML-based detectors:

- **Spectttra [42]:** transformer trained on spectrograms to detect Suno/Udio songs.
- **CLAP-based [64]:** hierarchical classifier trained on CLAP embeddings to detect Suno/Udio songs.
- **Reconstruction-based [3]:** CNN trained on spectrograms of reconstructed audio to detect VAE usage in AI music generation.
- **Fourier-based [2]:** Linear classifier trained to identify synthetic peaks in spectrograms of Suno/Udio songs. Deezer's built-in AI labeler is based on this type of models.

For each detector, we use the models released by the authors, and if unavailable, we reuse their code release to train a model using separate training data from Suno and Udio, different from our test set. For each detector, we convert the audio to the same format used to train the detector before evaluation.

[5] Detection of open-source generators can be found in Appendix A.5.

DRAFT

Evasion Technique	Success Rate
MP3 Compression	5/5
Pitch Shift	5/5
Reverb	5/5
Reconstruction	5/5

Table 5: All 5 AI-generated songs modified by each of the 4 techniques have successfully bypassed Deezer's AI detector.

Detection performance. We evaluate each detector on the test set (1,000 AI, 1,000 human) and plot performance via ROC curve in Figure 6. Overall, the Fourier-based detector outperforms the other three, with 99% accuracy. It also maintains an extremely low false positive regime, i.e. 99% TPR at 0.1% FPR. Reconstruction-based detector is the next best, achieving 68% TPR at 0.1% FPR.

ML-based detectors are known to be vulnerable to adversarial attacks, which authors of both the Fourier and CLAP-based detectors have confirmed in their work [2,64]. We also verified using experiments that compression can bypass both detectors. Next, we show that similar audio processing techniques can also be used to bypass Deezer's commercial AI music detector, which is currently being used to label its tracks.

Fooling Deezer's built-in AI detector. While there is no direct access to Deezer's AI detector, we test it by utilizing the distribution pipeline established in §5 to upload music to Deezer and manually observe their labels. Specifically, we generate 20 unique AI songs using Suno (the best performing generator), and modify each song by selecting one of 4 audio processing techniques: compression, pitch shift, reverb, and audio reconstruction using an autoencoder such that we upload 5 songs for each technique. Each song is then uploaded through DistroKid (fastest review turnaround of 1 day), specifying Deezer as the streaming target. None of these changes besides pitch shift alter the audio in a distinguishable manner.

Table 5 shows that all 4 techniques bypass Deezer's detector with 100% success, which we suspect is likely because Deezer's detector is trained without adversarial training methods. Using a local version of the Fourier detector that Deezer's is based on, we confirm that training on adversarially perturbed audio improves robustness significantly (95%).

Disclosure to Deezer. We have disclosed this result to Deezer's research team, who responded that they are currently working to improve the robustness of the detector.

6.2 User Study on Human Detection

Next, we study whether human listeners can reliably separate AI-generated and human-made songs. For this we performed a large scale user study, with N=910 participants of different music backgrounds and experiences. In the following, we begin with the design and setup of our user study, and then present the key results.

User study design. We present the core components and decisions of our user study. Additional details, including specific questions used by our study are included in Appendix A.6. Our study has been reviewed and approved by the IRB office at our institute.

Participants: We recruited 910 participants with different music backgrounds, including 250 general users (paid participants recruited via Prolific), 600 highly motivated volunteers (recruited from a social media channel), and 60 musicians (volunteers recruited via the social media channel and connections in the music industry).

Music selection: We selected a diverse set of songs out of six genres: Rap/Hip Hop, Pop, Electro, Jazz, Classical, Rock. These genres were chosen to achieve comprehensiveness while maintaining distances among the chosen genres (e.g. we did not select Dance due to its similarity to Hip Hop and Electro). For each genre, we selected 20 human-made and 20 AI-generated songs from Spotify based on the labels from Deezer. To ensure a fair comparison, we only chose songs released after 2024. When selecting songs, we avoided both highly popular songs and low-play songs, to ensure quality while minimizing the chance of participants already knowing the song. Thus we chose songs whose play count is within [10,000, 100,000].

Impact of visual information: Streaming services present each music track visually to listeners via its title, artist, album and album cover. Thus we varied information provided to participants to evaluate whether visual cues may affect human detection. Specifically, we first ask participants to listen to an audio clip (without any visual cues) and rate their confidence of it being AI or human-made on a 5-point Likert Scale. Immediately after, participants listen to the same clip *and* we reveal track title, artist name, album name, and album cover, and ask them to decide again.

User study setup. We configure our study as an online survey, where participants listen to a 30s song clip and decide it being AI or human-made on a 5-point Likert Scale. Participants are able to pause and repeat each audio clip as they wish. To avoid decision fatigue, we only show each participant 12 songs randomly selected from our pool. We include attention check questions to ensure quality of responses.

We also collect metrics on participants' familiarity with each music genre (for Prolific users and volunteers) and professional experience (for musicians). Near the end of the study, we ask them to rate the factors that helped detection and which genres were easiest to classify. Each Prolific participant was compensated at the rate of $12/hour, with the expected time of 10 mins to take the survey.

Finally, we compute detection accuracy from the survey results on a 5-point Likert Scale. We consider high or low confidence in the true type as correct detection, and the rest three entries as wrong detection.

Key Results. Overall, our study reveals that human users perform poorly at AI music detection. Prolific users are worse

DRAFT

User Group	Participants	Accuracy (%)	Δ Accuracy	p-value
Prolific	250	48.4	–	–
Volunteers	600	52.4	+4.0	$3.4e-7$
Musicians	60	60.7	+8.3	$2.0e-9$

Table 6: Overall detection accuracy by participant group. We calculate Δ Accuracy and p-values using two-sided z-tests on the accuracy difference over the Prolific group. Increases in detection accuracy from volunteer and musician groups are statistically significant.

User Group	Visual Cue	Accuracy (%)	Δ Accuracy	p-value
Prolific	No	44.7	+7.4	$< 1e-8$
	Yes	52.1		
Volunteers	No	45.8	+13.3	$< 1e-50$
	Yes	59.1		
Musicians	No	56.5	+8.4	$1.2e-3$
	Yes	64.9		

Table 7: Detection accuracy before and after visual cue of the music track is shown to participants. Δ Accuracy and p-values are computed using two-sided z-tests.

than a coin flip; volunteers are on par with a coin flip, and musicians do slightly better, at roughly 60% accuracy. Next we discuss multiple factors that affect the detection accuracy.

1) Volunteers outperform Prolific users by spending more time/efforts: Table 6 lists the average detection accuracy in each of the three participant groups as well as the statistical significance of these results. While all three groups cannot reliably detect AI music, the performance of the volunteer and musician groups are statistically better than the Prolific group. One key difference among the three groups is the average time spend on the study: 15 minutes, 42 minutes, and 72 minutes, for Prolific, volunteer and musician groups, respectively. This means that participants in the latter two groups listened to the song tracks repeatedly to improve detection.

2) Visual information of the music helps detection. Table 7 lists the detection accuracy before and after visual cues (album art & name, song title and artist) are displayed. All three groups show 7%-13% increases in detection accuracy after seeing this information. Later Figure 8 also shows that participants self-report album cover and artist name as the key factors guiding their decision.

3) Music genre affects detection: Genre played a key role in detection performance. Figure 7 shows that Electro and Jazz both see a lower accuracy. We ask participants to self-select which genres they believe are easiest to identify. Table 9 reports that while Jazz and Electro are consistently among the bottom across participants (congruent with measured accuracy results), Classical is actually ranked even lower. This suggests that participants show disparity between their beliefs and results, claiming Classical is harder to identify than Jazz and

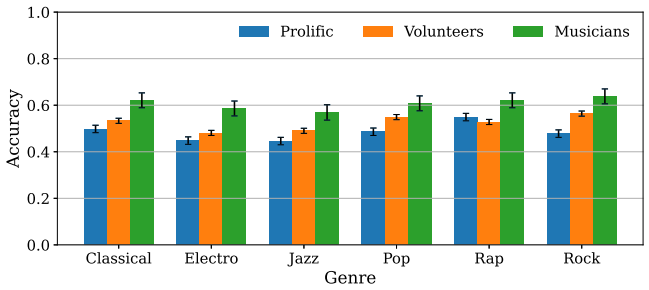


Figure 7: Accuracy by music genre and participant group. Error bars show 95% confidence intervals to indicate variability within each category. Musicians consistently outperformed the rest. All groups performed worst on Electro and Jazz.

User Group	Familiar with Genre	Accuracy (%)	Δ Accuracy	p-value
Prolific	False	46.0	+4.3	$9e-4$
	True	50.3		
Volunteers	False	51.8	+1.6	$4e-2$
	True	53.4		
Musicians	False	60.3	+0.8	0.73
	True	61.1		

Table 8: Accuracy by self-reported familiarity (Prolific, Volunteers) or professional experience (Musicians) with music genre. Δ Accuracy and p-values are computed using two-sided z-tests between familiarity conditions in each group.

Electronic, but actually performing significantly better on it.

4) Familiarity with music genre is consequential, professional experience is not: We ask Prolific users and volunteers which music genres they are most familiar with and ask musicians which genres they have professional experience in. Table 8 shows detection accuracy increases largely with familiarity with genres. Accuracy of Prolific users increases by 4.3% in genres they report being familiar with, that of volunteers increases by 1.6%, while musicians show increases of less than 1% (statistically insignificant due to the large p-value) with genres they have professional experience in.

Genre Easiest to Identify	**Prolific (%)**	**Volunteers (%)**	**Musicians (%)**
Rock	34.4 (48%)	48.0 (56%)	59.6 (64%)
Pop	38.2 (49%)	37.8 (55%)	36.2 (61%)
Rap / Hip Hop	43.3 (55%)	29.1 (53%)	31.9 (62%)
Electronic / EDM	29.3 (45%)	27.9 (48%)	31.9 (59%)
Jazz	24.8 (45%)	21.6 (49%)	29.8 (57%)
Classical	23.6 (50%)	21.0 (53%)	21.3 (62%)
None	0.0	0.6	0.0

Table 9: Distribution of genres reported as easiest to identify by user group. Percentages indicate the proportion of participants selecting each genre; values in parentheses represent the corresponding classification accuracy (%) for that genre and user group. Multiple selections were allowed.

DRAFT

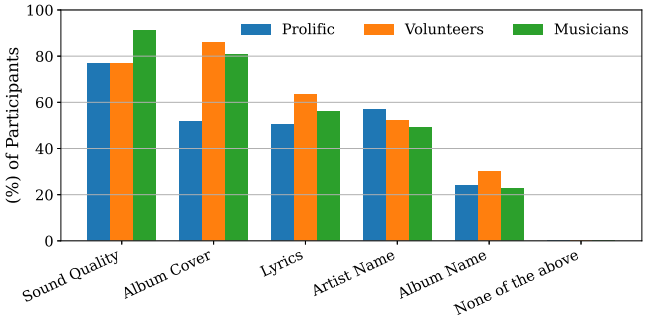


Figure 8: Properties Used to Identify Music, by User Group.

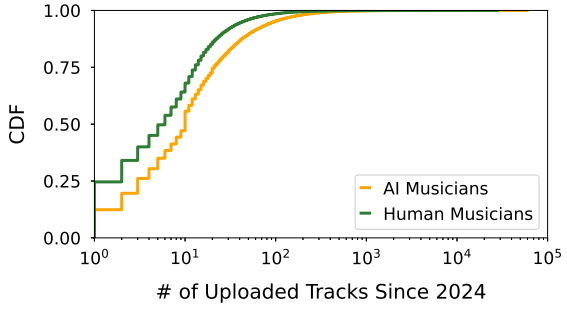


Figure 9: CDF of the number of uploaded tracks for both human and AI musicians (uploaded $\geq$ 1 AI track). While 90% of human musicians uploaded $\leq$26 tracks, 90% of AI musicians uploaded $\leq$52 tracks.

We also measure association with AI-generated music, prompting participants to chose 'Positive', 'Negative' or 'Both'. We find unsurprisingly that Prolific users have a more mixed view; 18% select 'Positive' while 37% select 'Both', whereas no volunteers or musicians selected 'Positive' and only roughly 6% select 'Both' instead of 'Negative'.

7 Economic-based Tools to Curb Music Slop

Our tests (§6) show that both automated tools (vulnerable to evasion attacks) and human listeners (very low accuracy) are ultimately limited when it comes to detecting AI-generated music. As AI generators mature, we expect third parties producing slop to gain access to faster and cheaper compute resources, making detection even more challenging.

Here, we outline an alternative, complementary approach that leverages the fundamental motivation behind AI slop: money. Ultimately, AI music slop is about high volume, low-cost content that floods the consumers, hoping that one or two songs out of thousands posted will catch an audience and generate revenue that pays for the entire campaign. The economics behind this is identical to that of large scale spam campaigns [32] and pharmaceutical scams [40]. What makes this scenario different is that we have stakeholders who are motivated to curb AI music slop, and who hold key controls over the economics of AI music: *music distributors*.

Initial observations. The reality is that the cost structure of music distribution remains optimized for human musicians (low volume, high quality), not potential AI slop producers (high volume, low quality). Most distributors charge a monthly/annual fee to upload music (Table 4). Properly amortized over a year, third parties can reduce per track costs significantly below $0.02 (§5.3). This low cost structure that encourages indie musicians to produce more tracks is being exploited by AI slop producers. Today, AI musicians consistently upload more tracks than human musicians (see Figure 9).

A key factor working in our favor is the relatively small number of music distributors with relationships to streaming services. Any transparent changes in costs will be clearly visible to streaming services. As the AI music slop problem grows in severity, distributors can be incentivized to change their cost structure and be perceived as preferred sources of higher quality music.

Shaping AI music using cost incentives. There are multiple ways to curtail AI music slop. While we lack visibility into the revenue and cost models inside music distributors, we can offer a number of potential changes, some of which might be more attractive to distributors than others.

- Charge distribution costs per track, increase costs for those uploading large number of tracks while keeping costs stable for others.
- Increase distribution costs as a whole, decreasing probability of AI music slop turning a profit.
- Increase per track costs as function of upload frequency, a disproportionally higher cost increase for those who upload large volumes of music.
- Cap number of tracks per musician per time, forcing slop producers to create numerous fake accounts to scale.

Some of these approaches have clear tradeoffs that not all distributors can accept. Some might dislike the higher overhead of per-track costs, a strict quota per musician or the potential loss of customers to others with lower costs.

Alternatives to assist AI music identification. We note that curbing AI slop does not help the accuracy of detecting AI music, a feature that users seem to demand. One simple idea is to verify musician and AI usage before distributing their music. While this is imperfect and can be gamed by bad actors, it nonetheless raises the friction of uploading slop, and would strongly discourage organized spammers from participation. Today, only 3 out of the 18 total distributors from [26] required identity verification. Many AI musicians on Reddit comment that they avoid those distributors out of fear of copyright violation claims or legal accountability.

Finally, charging distribution costs per album would shift AI slop producers to large albums with hundreds of tracks (we

DRAFT

already observe this in some AI musicians). Tracks per album can then be used as a reliable complementary mechanism for detecting AI music slop.

8 Conclusion

Leveraging a large-scale measurement analysis on Spotify, end-to-end experiments on music distribution in the wild, and a large-scale user study with 910 participants, our work sheds light on the growth of AI music in today's music ecosystem, and the potential problems of AI music slop. Our experiments show why AI music is easily injected into streaming platforms and that their robust detection (by machines and human listeners) is very challenging. On top of these observations, we also propose a potential direction for managing AI music using cost incentives. We hope our work spurs steps by distributors and streamers to limit slop while supporting non-slop artists, human and AI alike.

Ethical Considerations

Our study involves potential impact on several primary stakeholder groups. (1) *Streaming platforms* Spotify and Deezer. We used a public dataset of Spotify metadata in our analyses. We also used public APIs to gather data measurements from Spotify's recommendation graph and Deezer's AI labels. During measurements, we kept our query rate low to near-human levels to avoid incurring load on servers. We used 6 or fewer machines to crawl Spotify, and made sure they had publicly identifiable IP addresses so we could be easily identified and reached if we caused problems. We observed no server errors or throttling or limiting during data measurements. (2) *11 Music distributors* that we uploaded music to using standard practice for indie music artists. We also analyzed their policies on AI music and included summaries in the paper. (3) *Participants in our user study*, which includes paid workers from Prolific crowdsource platform, as well as volunteers and self-identified professional musicians from social media. Our user study was approved by our university IRB office prior to conducting this research. (4) *Users of streaming platforms* who have been vocal about proliferation of AI music and asked for transparency and tools to label AI music and limit AI slop. (5) *Researchers and Practitioners* including the ML and security community. Groups (4) and (5) were not impacted during the research process. (6) *Human music artists* want to better understand the impact of AI music and potential slop on their industry and livelihoods. Our uploads of a small number of AI music tracks received only a handful of plays, and should not impact human musicians.

Disclosures. We submitted disclosure to Spotify, informing them that their public API could put them at risk for data extraction. Spotify replied thanking us, and stating there is no significant security impact from this disclosure and they classified our report as "informative." We also communicated our use of the Deezer API and our results directly to the Deezer research team, and their response was quite positive and cordial.We were careful to not gather any data that could be deemed sensitive or private. All of our measurements used publicly available APIs, and all data were public fields that any user could see by browsing Spotify and Deezer apps. Our interactions with data distributors all followed standard processes based on their policies and did not disclose any private or sensitive data.

Copyright. Respecting copyright of musicians whose samples we included in our user study, we did not at any time download or store music. Instead, our user study showed participants 30-second samples that were hosted by Deezer and made available via their public APIs.

Justification for research. This work sheds much needed transparency on the growth and impact of AI music into popular streaming platforms like Spotify. Conducting this work is critical to identify potential risks to both human and AI musicians from AI music slop, and contributes to insights on potential mitigation mechanisms to limit the harm of AI music slop.

References

[1] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Controlling popularity bias in learning-to-rank recommendation. In *Proc of RecSys*, 2017.

[2] Darius Afchar, Gabriel Meseguer-Brocal, Kamil Akesbi, and Romain Hennequin. A Fourier explanation of AI-music artifacts. In *Proc. of ISMIR*, 2025.

[3] Darius Afchar, Gabriel Meseguer-Brocal, and Romain Hennequin. AI-generated music detection and its challenges. In *Proc of ICASSP*, 2025.

[4] AIVA. Your personal AI music generation assistant. https://www.aiva.ai (accessed Feb. 4, 2026).

[5] Anna's Archive. Backing up Spotify, Dec. 2025. https://annas-archive.li/blog/backing-up-spotify.html.

[6] Tom Beasley. Artificial/intelligent: AI slop dominates online bookstores, so what now? Voice Magazine UK, Jun. 2025. https://www.voicemag.uk/feature/15327/ai-books-online-future-amazon.

[7] BeatsToRapOn. AI music detector & checker. https://beatstorapon.com/ai-music-detector (accessed Feb. 4, 2026).

[8] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The million song dataset. In *Proc. of ISMIR*, 2011.

DRAFT

[9] Black Forest Labs. FLUX, 2024. https://github.com/black-forest-labs/flux.

[10] Boomy. Unleash your creativity make music with Boomy AI. https://boomy.com/ (accessed Feb. 4, 2026).

[11] Lauren Brown. 'AI' slop versions' of books on retailers like amazon 'risk harming consumer confidence'. The Bookseller, Aug. 2025. https://www.thebookseller.com/news/ai-slop-versions-of-books-on-retailers-like-amazon-risk-harming-consumer-confidence.

[12] Barbara Collins. How Spotify stayed no. 1 in streaming audio even with Apple, YouTube and Amazon aiming for it. CNBC, Nov. 2022. https://www.cnbc.com/2022/11/10/how-spotify-stayed-no-1-in-streaming-music-vs-apple-youtube-amazon.html.

[13] Jade Copet et al. Simple and controllable music generation. In *Proc. of NeurIPS*, 2023.

[14] Rowan Davies. AI music is flooding spotify, and subscribers are furious – here's why music fans 'no longer trust discover weekly'. Tech Radar, Jan. 2026. https://www.techradar.com/audio/spotify/ai-music-is-flooding-spotify-and-subscribers-are-furious-heres-why-music-fans-no-longer-trust-discover-weekly.

[15] Deezer. Deezer launches world's first AI tagging system for music streaming, Jun. 2025. https://newsroom-deezer.com/2025/06/deezer-launches-worlds-first-ai-tagging-system-for-music-streaming/.

[16] Deezer. Deezer confirms demonetization of up to 85% of AI-music streams due to fraud and moves to sell AI-detection technology, Jan. 2026. https://newsroom-deezer.com/2026/01/ai-generated-music-deezer-selling-detection-tool/.

[17] Michaël Defferrard, Kirell Benzi, Pierre Vandergheynst, and Xavier Bresson. FMA: A dataset for music analysis. In *Proc. of ISMIR*, 2017.

[18] Prafulla Dhariwal et al. Jukebox: A generative model for music. *arXiv preprint arXiv:2005.00341*, 2020.

[19] Robert Dorfman. A formula for the Gini coefficient. *The Review of Economics and Statistics*, 1979.

[20] Stuart Dredge. Umg boss slams 'exponential growth of AI slop' on streaming services. Music Ally, Jan. 2026. https://musically.com/2026/01/09/umg-boss-slams-exponential-growth-of-ai-slop-on-streaming-services/.

[21] Fabio Duarte. Music streaming services stats (2026). Exploding Topics, Jan. 2026. https://explodingtopics.com/blog/music-streaming-stats.

[22] Zach Evans et al. Stable audio open. In *Proc of ICASSP*, 2025.

[23] Ghazal Fazelnia, Sanket Gupta, Claire Keum, Mark Koh, Timothy Heath, Guillermo Carrasco Hernández, Stephen Xie, Nandini Singh, Ian Anderson, Maya Hristakeva, Petter Pehrson Skidén, and Mounia Lalmas. Generalized user representations for large-scale recommendations and downstream tasks. In *Proc of RecSys*, 2025.

[24] Seth Forsgren and Hayk Martiros. Riffusion - Stable diffusion for real-time music generation. HuggingFace, 2022. https://huggingface.co/riffusion/riffusion-model-v1,.

[25] Junmin Gong, Sean Zhao, Sen Wang, Shengyuan Xu, and Joe Guo. ACE-Step: A step towards music generation foundation model. *arXiv preprint arXiv:2506.00045*, 2025.

[26] Ari Herstand. Best music distribution companies in 2026 - full comparison chart. Ari's Take, 2026. https://aristake.com/digital-distribution-comparison/.

[27] Hive. AI generated music detection. https://docs.thehive.ai/docs/ai-generated-music-detection (accessed Feb. 4, 2026).

[28] Tim Ingham. The major record companies release around 3,900 tracks every day. Their problem? That's a drop in the ocean. Music Business Worldwide, Mar. 2023. https://www.musicbusinessworldwide.com/record-companies-3900-tracks-every-day-problem/.

[29] Harry H. Jiang et al. AI art and its impact on artists. In *Proc. of AIES*, August 2023.

[30] Jonathan Landrum Jr. AI country hit 'Walk My Walk' built on Blanco Brown's sound sparks questions of attribution, ethics. AP News, Nov. 2025. https://apnews.com/article/walk-my-walk-blanco-brown-2c9bbde6e88434365640c50e2998cfe2.

[31] Maura Judkis. The deluge of bonkers AI art is literally surreal. Washington Post, Jun. 2024. https://www.washingtonpost.com/style/of-interest/2024/06/30/ai-art-facebook-slop-artificial-intelligence/.

[32] Chris Kanich et al. Spamalytics: an empirical analysis of spam marketing conversion. In *Proc. of CCS*, 2008.

DRAFT

[33] Leslie Katz. Man charged with $10 million streaming music scam using AI-generated songs. Forbes, Sep. 2024. https://www.forbes.com/sites/lesliekatz/2024/09/08/man-charged-with-10-million-streaming-scam-using-ai-generated-songs/.

[34] John Koetsier. 20,000 AI-generated songs uploaded to Deezer every single day. Forbes, Apr. 2025. https://www.forbes.com/sites/johnkoetsier/2025/04/16/20000-ai-generated-songs-uploaded-to-deezer-every-single-day/.

[35] Yupei Li, Manuel Milling, Lucia Specia, and Björn W Schuller. From audio deepfake detection to AI-generated music detection–a pathway and overview. *arXiv preprint arXiv:2412.00571*, 2024.

[36] Zihan Liu et al. Songgen: A single stage auto-regressive transformer for text-to-song generation. In *Proc. of ICML*, 2025.

[37] Arwa Mahdawi. AI-generated 'slop' is slowly killing the internet, so why is nobody trying to stop it? The Guardian, Jan. 2025. https://www.theguardian.com/global/commentisfree/2025/jan/08/ai-generated-slop-slowly-killing-internet-nobody-trying-to-stop-it.

[38] Emanuel Maiberg. Spotify publishes AI-generated songs from dead artists without permission. 404 Media, Jul. 2025. https://www.404media.co/spotify-publishes-ai-generated-songs-from-dead-artists-without-permission/.

[39] MatchTune. Deepmatch - detect AI-generated music. https://www.matchtune.com/deepmatch (accessed Feb. 4, 2026).

[40] Damon McCoy, Andreas Pitsillidis, Grant Jordan, Nicholas Weaver, Christian Kreibich, Brian Krebs, Geoffrey M. Voelker, Stefan Savage, and Kirill Levchenko. Pharmaleaks: Understanding the business of online pharmaceutical affiliate programs. In *Proc. of USENIX Security*, August 2012.

[41] Marco Pasini and Jan Schlüter. Musika! fast infinite waveform music generation. In *Proc. of ISMIR*, 2022.

[42] Md Awsafur Rahman et al. SONICS: Synthetic or not - identifying counterfeit songs. In *Proc. of ICLR*, 2025.

[43] Max Read. Drowning in slop. New York Magazine, Sep. 2024. https://nymag.com/intelligencer/article/ai-generated-content-internet-online-slop-spam.html.

[44] Nathan Reiff. Famous pop stars and their views on music AI, from Grimes to Drake. Decrypt, Aug. 2023. https://decrypt.co/resources/famous-pop-stars-and-their-views-on-music-ai-from-grimes-to-drake.

[45] Wojciech Reise, Ximena Fernández, Maria Dominguez, Heather A. Harrington, and Mariano Beguerisse-Díaz. Topological fingerprints for audio identification. *SIAM Journal on Mathematics of Data Science*, 2024.

[46] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proc. of CVPR*, 2022.

[47] Salif. The benefits of releasing your music across multiple streaming platforms, Nov. 2024. https://www.vampr.me/faq/the-benefits-of-releasing-your-music-across-multiple-streaming-platforms/.

[48] sleeping-ai. suno-307k. HuggingFace, 2025. https://huggingface.co/datasets/sleeping-ai/suno-307K/tree/main.

[49] sleeping-ai. udio-126k. HuggingFace, 2025. https://huggingface.co/datasets/sleeping-ai/Udio-126K.

[50] Soundraw. 100% Copyright-safe AI music generator. https://soundraw.io/ (accessed Feb. 4, 2026).

[51] Spherical Insights. Top 20 companies in generative AI in music market 2025: strategic overview and future trends (2024-2035), Sep. 2025. https://www.sphericalinsights.com/blogs/top-20-companies-in-generative-ai-in-music-market-2025-strategic-overview-and-future-trends-2024-2035.

[52] Spotify. Getting music on Spotify. https://support.spotify.com/us/artists/article/getting-music-on-spotify/ (accessed Feb. 4, 2026).

[53] Spotify. Track monetization eligibility. https://support.spotify.com/us/artists/article/track-monetization-eligibility/ (accessed Feb. 4, 2026).

[54] Spotify. Understanding Spotify royalties. https://support.spotify.com/us/artists/article/understanding-spotify-royalties/ (accessed Feb. 4, 2026).

[55] Spotify. Spotify strengthens AI protections for artists, songwriters, and producers, Sep. 2025. https://newsroom.spotify.com/2025-09-25/spotify-strengthens-ai-protections/.

[56] Spotify. From $11b in 2025 payouts to what we're building for artists in 2026, Jan. 2026. https://news

DRAFT

room.spotify.com/2026-01-28/2025-music-industry-payouts-whats-next-for-artists/.

[57] Alisa Rata Stutzbach. MusicBrainz. *Notes*, 2011.

[58] SubmitHub. AI song checker. https://www.submithub.com/ai-song-checker (accessed Feb. 4, 2026).

[59] SubmitHub. Who we work with. https://www.submithub.com/ (accessed Feb. 4, 2026).

[60] Suno. https://suno.com/ (accessed Feb. 4, 2026).

[61] Jiaru Tang and Patrik Wikström. 'Side job, self-employed, high-paid': behind the AI slop flooding TikTok and Facebook. The Conversation, Sep. 2024. https://theconversation.com/side-job-self-employed-high-paid-behind-the-ai-slop-flooding-tiktok-and-facebook-237638.

[62] Hugo Touvron et al. LLaMA: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

[63] Udio. AI music generator. https://www.udio.com/ (accessed Feb. 4, 2026).

[64] Laura Vila, Bob Sturm, Luca Casini, and David Dalmazzo. The AI music arms race: On the detection of AI-generated music. In *Proc. of ISMIR*, 2025.

[65] Elissa Welle. Spotify is finally taking steps to address its AI slop and clone problem. The Verge, Sep. 2025. https://www.theverge.com/news/785136/spotify-ai-slop-impersonation-disclosure.

[66] Ning Ziqian et al. DiffRhythm: Blazingly fast and embarrassingly simple end-to-end full-length song generation with latent diffusion. *arXiv preprint arXiv:2503.01183*, 2025.

A Appendix

A.1 Detailed Spotify Crawl

Crawling the largest connected component. To find the most important connected tracks, we leverage a BFS/snowballing methodology using Spotify's recommendations API to find new tracks. First, we started with a set of "seed" tracks, which we should be both comprehensive and diverse. Then, we repeatedly query Spotify recommendations API to establish connections (edges) from each seed track to new seed tracks. This process is repeated until there are no more new tracks recommended. Note that at each sampling layer, we additionally seeded the crawler with other songs released by the musician of the current track to expand our reach.

We seeded our crawl with 200,000 tracks spanning 75 different countries obtained through Spotify's official playlists, which represented the most relevant songs up to that date. Our crawl took 2 weeks to complete in November 2025, and returned ≈65 million unique track IDs. Then, using graph analysis, we computed the largest strongly connected component using the digraph constructed with each track as a node, and recommendations as directed edges. We find that there is a single large component (33 million tracks), and many tiny components, each with less than 20 tracks each. Based on this data, our SCSR dataset consists of two parts: 1) a database of 33 million tracks, where 6.8 million were released on/after 2024, and 2) a recommendation network of 33 million tracks (nodes) and 3.5 billion recommendations (directed edges).

Spotify recommendations API. The key to our web crawl was being able to obtain recommendations for any given track. Spotify deprecated their official recommendations API endpoint early 2025. However, when a user visits a track on their webpage, they are still able to view recommended tracks. This information is obtained via an internal GraphQL query, which alongside recommendations returns other metadata, including the track ID, album ID, artists, play count, . . . , etc. We pinpointed the exact query that returned this information, and used it as an API for our web crawler.

While this is all publicly available information, the usage of this API at scale was enabled by a small bug in Spotify's GraphQL authentication. The GraphQL query requires an authentication token, which is normally non-trivial to obtain. However, Spotify's embedding link (often used for iFrame HTML embeddings) automatically generates a valid authentication token that lasts for hours at a time. This allowed us to crawl Spotify's recommendations API at scale by polling a random embedding link (which is nearly instantaneous) whenever our tokens expired. We have since informed Spotify of this vulnerability.

A.2 Detailed Deezer Matching Algorithm

As we described in §3.3, we labeled AI tracks in both our Spotify datasets by matching them to the Deezer platform, which already labels AI music. To accomplish this, we need to 1) match Spotify tracks to a corresponding one on Deezer, and 2) determine if the track on Deezer is labeled as AI. We were able to automate this process using two separate APIs.

Searching and matching tracks on Deezer. For each Spotify track, we first query Deezer's official search API using combinations of the track's title, album title, and artist name to obtain a list of candidate matches. Then, we identify the correct candidate based on whether or not the two track's metadata match. Note that metadata between Spotify and Deezer is often not exactly the same (i.e., released at different times, title formatting, ..., etc.). As a result, we manually defined and tested a set of heuristics for a "match." These rules are meant to be slightly fuzzy so as to catch any inconsistencies

DRAFT

in metadata:

- album title, track title, release year, and at least one artist matches
- album title, track title, and all artists match
- track title, release year, and all artists match
- album title, track title, release year, and track duration matches
- album title, release year, track duration, and copyright label matches

Before we employed this matching algorithm on our entire dataset, we tested it on a random sample of 100 Spotify tracks sampled from SCRG. We were able to obtain matches for 80% of tracks, which we manually verified to be correct. We also found no false positives using our matching rules.

Labeling AI on Deezer. Once we established a match for a Spotify track on Deezer, our final step is to obtain the AI label that Deezer assigned to that matched track. Since Deezer labels AI at the album level, we just needed to identify the AI label for the album of the matched Deezer track. While this label can be publicly viewed on Deezer's web page, it is not a part of Deezer's official developers API. Instead, similar to Spotify, Deezer leverages an internal GraphQL query to identify which albums need to display the AI label. As such, we use a similar methodology as we did with Spotify's recommendations API (querying GraphQL) to obtain labels for any given Deezer album, and tie that back to the original Spotify track.

A.3 Generating AI Tracks

We generated AI tracks using as little human-input as possible. All content, including lyrics, was generated using a generative AI model. Below, we detail how we created each track, split by whether it was done using a closed/open-source model. For all tracks, cover art was generated via FLUX [9] using prompts from ChatGPT, followed by SDXL's 4x upscaler [46] to increase the resolution of the image.

Suno/Udio tracks. Both Suno and Udio came with built-in capabilities to random generate song titles and lyrics. We used the randomize feature in both tools to generate full-length instrumentals and lyrical songs, and reused the song titles that were provided.

ACE-Step/DiffRhythm tracks. For the open-source models, we generated lyrics using LLaMA-3.1 [62] paired with song titles generated by ChatGPT. The prompt for the song was randomly chosen from a bank of 10 different song prompts also generated using ChatGPT.

Predictor	Intercept	X_a	X_{gs}
Coefficient	-5.173	4.555	0.623

Table 10: Logistic regression results for predicting whether a recommended track is AI-generated (Y_a). The dominant predictor (higher coefficient) is whether the source track is AI-generated (X_a), even accounting for genre similarity (X_{gs}). All predictors are statistically significant ($p < 0.001$).

A.4 Genre as a Potential Factor for AI Music Connectivity

In §4.4, we found that AI music is significantly more likely to recommend other AI music compared to human-made music. However, recommendation algorithms are complex, and there are many potential factors for this observation. One of the more intuitive reasons for why AI might recommend more AI is that AI tracks may share similar genres, which recommendation systems are likely to bias connections towards. Here, we conduct an experiment to rule out genres similarities as the primary reason for our AI music connectivity result.

To test this, we trained a logistic regressor on a sample of all edges (90/10 train-test split), equally divided between AI and human-created tracks from our recommendations network via the SCRG dataset, where inputs to the model X_a is 1 if the source node is AI, 0 otherwise, X_{gs} is the genre similarity of the source and recommended node, and Y_a represents if the target node is AI.

$$Y_a = c_0 + c_1 \cdot X_a + c_2 \cdot X_{gs}$$

If genre similarity were to be the more important factor as to why AI tracks recommend other AI tracks, then c_1 should be much larger than c_2 after training. In table 10, we summarize the regression results. Our regressor achieves 81.39% accuracy, and c_1, which is whether the source node is AI, has by far the highest positive coefficient ($p < 0.001$). At the very least, this confirms that the source track being AI has a much stronger effect on the recommended track being AI compared to genre similarity.

A.5 Open Source Music Generators

While we were uploading AI songs ourselves to music streaming platforms, we noticed that all of our DiffRhythm/ACE-step songs were not being labeled as AI by Deezer. This means that our labeling methodology (which uses Deezer's detector) may be underestimating the total amount of AI music. However, given that open source generators still lag behind SOTA in Suno/Udio, it's possible they do not play a large role in the current AI music ecosystem. We do experiments to investigate this here.

First, we trained a highly accurate detector of DiffRhythm and ACE-Step songs using the CNN architecture proposed

DRAFT

in [3]. Note that this approach does not generalize to Suno or Udio songs since it relies on having access to a model's latent encoders/decoders, which we only have for DiffRhythm and ACE-Step. Our final model achieved >99% accuracy on both DiffRhythm and ACE-Step songs, with a 0.1% false positive rate. Finally, we applied the model to a random sample of 10,000 human-made Deezer songs released between October and November 2025.

Our results found that only 0.4% were identified to be generated via DiffRhythm and ACE-Step. In comparison, our labeled data found that current AI releases (Figure 2) already account for well over 40% of all weekly releases. This result confirms our suspicion that open source models still lag behind Suno and Udio, and that any unlabeled tracks generated by DiffRhythm or ACE-Step are unlikely to change the overall conclusions of our analysis.

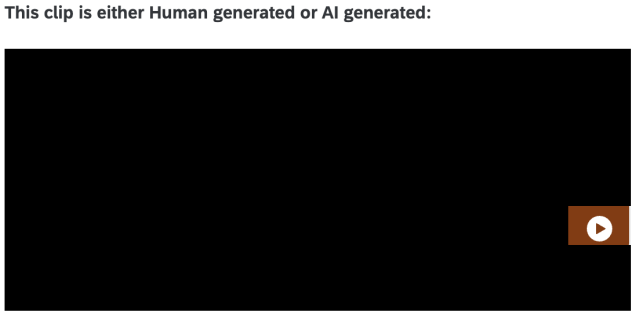
A.6 User Study Detailed Description

Here we provide additional details on our user study questions and protocol to supplement Section 6.2. Table 11 lists the questions we asked participants. Figure 10 shows screenshots of the user study interface, with and without the visual cues (track title, artist name, album name, and album artwork) shown to participants.

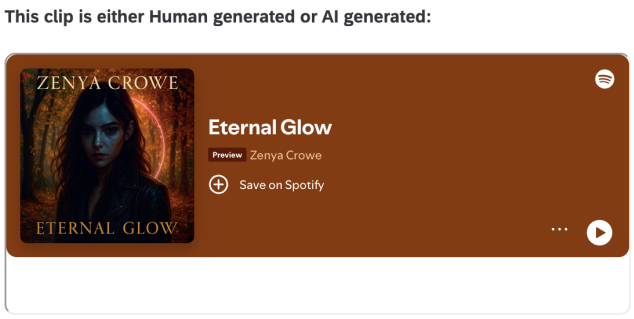
DRAFT

Question	Response Type
"What are your thoughts about the impact of AI-generated music on the music industry?"	Single select
"Do you self-identify as a professional musician?"	Single select
"Which genres of music are you most familiar with as part of your work?"	Multiple select
"Do you think the above clip is produced by a human or generated by an AI?"	Likert scale (1–5)
"Did you find music from any of the following categories easier to distinguish as AI-generated?"	Multiple select
"Which properties do you use to differentiate AI-generated music from human-made music?"	Multiple select

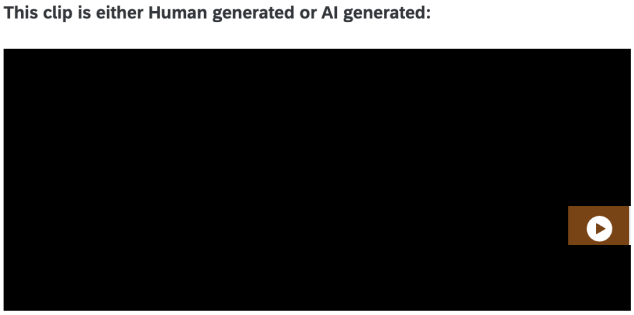
Table 11: User study questions and response types. Response options are listed in analysis of study results in Section 6.2.



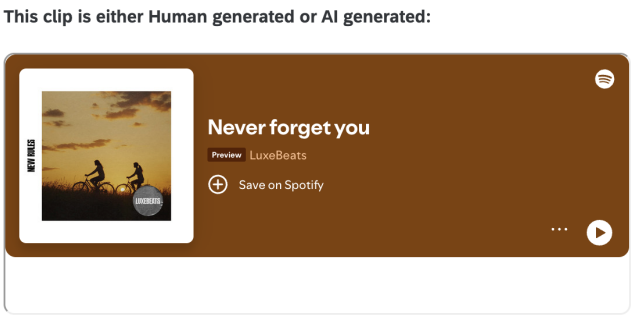
(a) AI-generated audio clip shown *before* visual cues.



(b) AI-generated audio clip shown *after* visual cues.



(c) Human-generated audio clip shown *before* visual cues.



(d) Human-generated audio clip shown *after* visual cues.

Figure 10: User study interface screenshots showing clips *before* and *after* 'visual cues': left column (a, b) shows AI-generated audio clips, right column (c, d) shows human-generated audio clips.