



# Algorithmic (In)Tolerance: Experimenting with Beethoven's Music on Social Media Platforms

RESEARCH ARTICLE

ADAM ERIC BERKOWITZ

**]u[**ubiquity press

## ABSTRACT

Popular social media platforms, such as YouTube and Facebook, see insurmountable volumes of media uploaded every day. The extent of which cannot be feasibly monitored through human efforts alone to identify infringing activity. Such companies employ algorithmic methods to enforce copyright by removing or monetizing content on behalf of copyright owners; however, occasionally flagged material is misidentified as infringing. These instances represent a potential loss of income for independent artists by way of misappropriated revenue or removal of material, and time spent challenging automated decisions. This article discusses an experiment seeking to ascertain the false positive rate of YouTube's Content ID and Facebook's Rights Manager. This is put in the context of existing legal precedence in the United States, including the Digital Millennium Copyright Act (DMCA). The article makes recommendations for technological and logistical modifications to these systems, and it encourages public education and research on the topic.

## CORRESPONDING AUTHOR:

**Adam Eric Berkowitz**Tampa-Hillsborough County  
Public Library, US[berkowitza@hcflgov.net](mailto:berkowitza@hcflgov.net)

## KEYWORDS:

social media; public domain,  
fair use; copyright; algorithmic  
moderation; digital rights  
management

## TO CITE THIS ARTICLE:

Berkowitz, A. E. (2023).  
Algorithmic (In)Tolerance:  
Experimenting with  
Beethoven's Music on Social  
Media Platforms. *Transactions  
of the International Society for  
Music Information Retrieval*,  
6(1), 1–12. DOI: [https://doi.  
org/10.5334/tismir.148](https://doi.org/10.5334/tismir.148)

## 1. INTRODUCTION

“Who owns that tune?” drives YouTube’s Content ID and Facebook’s Rights Manager as they scan video files and livestreams for infringing content. The Digital Millennium Copyright Act (DMCA), supported by court rulings, puts the onus of responsibility to identify infringement on copyright owners. YouTube, in response to pressure from media corporations, developed Content ID to monitor its platform for unauthorized material. YouTube was able to regain the confidence of corporate copyright owners when Content ID began automatically monetizing user uploads for allegedly infringing copyright, thereby establishing a new revenue stream for both itself and the music industry (Soha and McDowell, 2016). Facebook, noticing YouTube’s success, followed suit by launching Rights Manager (Keef and Ben-Kereth, 2016).

Automated copyright enforcement systems, however, often fail to recognize the legal distribution of Public Domain works. The Australian Broadcasting Corporation labeled such incidents “copywrongs” for their frequent occurrence (Lorenzon, 2018). Recently, Covid-19 restrictions forced classical musicians to move live performances online, which posed unforeseen challenges to artists. In *The Washington Post*, Brodeur (2020) recounted the experiences of musicians uploading and broadcasting performances on YouTube and Facebook who were hindered by “copyright bots” misidentifying their recitals for copyrighted recordings (Brodeur, 2020).

The intricacies of YouTube’s and Facebook’s algorithmic models are confidential, and court rulings shield their trade secrets from public scrutiny (Jacques et al., 2018; Perel and Elkin-Koren, 2017). It is, therefore, uncertain how each system identifies infringing material. Scholarship points to digital fingerprinting and perceptual hashing as they are both flexible enough to compare protected content to renditions and copies (Gorwa, Binns and Katzenbach, 2020; Urban, Karganis and Schofield, 2017). These algorithms analyze user-uploaded files for similarity against a database of copyrighted content provided by copyright owners. YouTube and Facebook also offer livestreaming options which transmit audiovisual footage instantaneously, bypassing systematic processing. Without the time to access a reference library for review, artificial neural networks equipped with deep learning apply pre-learned training data to classify video feeds on the fly (Zhang et al., 2018a; Zhang et al., 2018b).

YouTube, by controlling 47% of the on-demand music streaming market, is currently the leading service provider with much of its material contributed by uploaders (Prey, 2015; Reis and Burns, 2020). Music accompanies a wide variety of content including tv shows, movies, videogames, and advertisements in addition to artistic performances. Therefore, music copyright owners have broad claim to videos and broadcasts.

In 2016 and 2018, Google reported that Content ID executed 98% of all copyright enforcement decisions, demonstrating minimal human intervention (Google, 2018; Jacques et al., 2018). This is likely because 95% of copyright claims are automatically acted upon according to the preferences of digital rights management profiles. As a result, 50% of the music industry’s YouTube-sourced profits come from user uploads (Borgsmiller, 2019). Recourse exists for users, but it is seldomly leveraged. Research suggests that users may be unaware of dispute systems, overwhelmed by the necessity for specialized knowledge, apprehensive of the consequences that come with failing to dispute a claim, or intimidated by litigious corporations (Davis, 2018; Lawrence-Williams, 2022; Solomon, 2015; Zapata-Kim, 2016).

This article reports on experimentation by which original recordings of Beethoven’s piano sonatas were uploaded to YouTube and Facebook. It contributes to the limited body of literature demonstrating Content ID’s and Rights Manager’s abilities to recognize original recordings of Public Domain compositions as non-infringing. The remainder of this document is structured as follows: Section 2 recounts different research methods utilized for assessing online copyright enforcement. Section 3 describes how recordings were produced and uploaded to YouTube and Facebook. Section 4 documents how each system performed during testing and narrates the dispute process for each platform. Section 5 discusses the results and recommends technological and logistical improvements. Finally, Section 6 explores the ethics of utilizing automated copyright enforcement systems alongside generative artificial intelligence.

## 2. LITERATURE REVIEW

The US Constitution grants Congress the power to pass laws that “promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries,” (Library of Congress and US Copyright Office, n.d.). Congress, through this power, enacted legislation governing copyright. “Copyright protection lasts for the life of the author plus an additional 70 years. For an anonymous work, a pseudonymous work, or a work made for hire, the copyright endures for a term of 95 years from the year of its first publication or a term of 120 years from the year of its creation, whichever expires first,” (US Copyright Office, n.d.a). Once a copyright has expired, the work enters the Public Domain.

This applies to music in two ways. A copyright can protect a scored composition, or a copyright can cover audio and audiovisual recordings of the underlying composition. The print music and the recorded music are two distinct creative works. Artists must retain separate copyrights to secure both their compositions and their

recordings of their compositions. The copyright pertaining to the scored work does not extend to recordings of the music, and likewise, copyright for the recorded music does not extend to the underlying composition (Rae, 2021).

Although original works are protected by copyright, Fair Use Doctrine allows for unauthorized borrowing from copyrighted works and serves as a legal defense against copyright infringement claims (Stim, 2019). The following four criteria must be considered before rendering a decision on fair use: “the purpose and character of the use, including whether the use is of a commercial nature or is for nonprofit, educational purposes; the nature of the copyrighted work; the amount and substantiality of the portion used in relation to the copyrighted work as a whole; the effect of the use upon the potential market for or value of the copyrighted work,” (US Copyright Office, n.d.b).

The DMCA outlines the provisions for enforcing copyright online. Copyright owners must comply with the following to have allegedly infringing content removed from websites: notify the service provider in writing; clearly identify both the allegedly infringing content and the matching protected work; include the copyright owner’s contact information; affirm that the petition for removal is made in good faith and that all details are accurate. Web services must demonstrate the following when petitioning for safe harbor protections which guarantee amnesty for hosting infringing content: absence of knowledge that infringing activity has taken place; efforts to quickly remove infringing content when they gain such knowledge; a lack of benefit from infringing activity in cases where they have the right and ability to control infringing material; policies established to dissuade users from violating copyright law (Solomon, 2015). Despite these guidelines, case law provides necessary clarity on interpretations of the DMCA as applied in the United States.

*Viacom v. YouTube* defined “knowledge of infringing activity.” It was not clear whether the DMCA meant inferred awareness of infringing activities or knowledge of specific incidents of infringement. The courts determined the latter to be the standard because the DMCA requires that specific cases be identified by the copyright owner under the notice-and-takedown provision. Furthermore, the court’s decision affirmed that online service providers are not obligated to actively monitor their platforms for infringing activity. The court also clarified the meaning of “right and ability to control infringing material.” It was assumed that having administrative power over the platform met this criterion, but the court determined otherwise. The court stated that because YouTube’s media storage and access processes are entirely governed by algorithms, they do not possess the right and ability to control infringing material (Lawrence-Williams, 2022).

*Lenz v. Universal Music* focused on Stephanie Lenz who received a DMCA takedown notice administered by YouTube on behalf of Universal Music. They claimed that a video of her child dancing to a song by Prince allegedly infringed copyright. YouTube deleted the video in compliance with the DMCA. Lenz filed a counter-notification, a form of legal recourse users can take against takedown notices, which prompted the restoration of the video, and she sued Universal Music for not considering fair use. Lenz asserted that ignoring fair use violated the DMCA’s good faith belief mandate. No clear victor emerged in this case, but the court ruled that copyright owners are required to review allegedly infringing content for fair use prior to issuing takedown notices (Reymond, 2016).

The corpus of research conducted by Urban and Quilter (2006), Quilter and Heins (2007), Seng (2014; 2021), Bar-Ziv and Elkin-Koren (2017), and Urban, Karganis, and Schofield (2017) showed the results of surveyed databases and recipients of DMCA takedown notices. They indicated that copyright owners misunderstood citizen rights under copyright law and abused the notice-and-takedown process. Takedown notices were often autogenerated and incomplete which overburdened online service providers. These web services were forced to either expend vast resources to review thousands of often invalid takedown notices or comply without reviewing them to demonstrate swift action against allegedly infringing material. The latter of which does not result in legal consequences for service providers but does infringe on individual freedoms. Each study raised concerns over rights to freedom of expression and due process (Bar-Ziv and Elkin-Koren, 2017; Quilter and Heins, 2007; Seng, 2014; Seng, 2021; Urban, Karganis and Schofield, 2017; Urban and Quilter, 2006).

Urban, Karganis, and Schofield (2017) also surveyed online service providers to learn how they address DMCA takedown notices. Some respondents indicated they dedicated resources to manually review petitions, and others stated they used automated means to handle autogenerated notices. A third group noted they implemented algorithmic solutions to preemptively respond to copyright issues while also pursuing licensing deals with corporate copyright holders (Urban, Karganis and Schofield, 2017). YouTube and Facebook fall under both the second and third categories.

In legal analyses of US copyright law, Solomon (2015) alleged that automated copyright enforcement systems like Content ID are incapable of making allowances for fair use which prevents users from expressing themselves online. When such videos are claimed, Content ID unfairly monetizes these videos on behalf of copyright owners, preventing uploaders from generating revenue (Solomon, 2015). Bartholomew (2015) went further to suggest that Content ID subverts fair use, setting Fair Use Doctrine against users. By doing so, Content ID issues

copyright claims indiscriminately and fosters inequity (Bartholomew, 2015). Zapata-Kim (2016) concurred and pointed to how Content ID relieves copyright owners from conducting fair use reviews of allegedly infringing content. Online service providers preemptively blocking and monetizing videos give copyright owners the satisfaction of knowing that their intellectual property is actively protected independent of the DMCA (Zapata-Kim, 2016). YouTube benefits by retaining a portion of the revenue from monetized videos. According to Borgsmiller (2019), “YouTube is not required to create Content ID per the DMCA, but they did not do it for nothing. YouTube saw an opportunity to automate and streamline the notice and takedown process while making money for themselves and the music industry at the same time,” (Borgsmiller, 2019: 673).

Gray and Suzor (2020) determined that music claimants were far more likely to opt for monetization over blocking. This study used natural language processing to analyze a corpus of YouTube videos to evaluate Content ID’s reactions to cases of alleged copyright infringement. The study revealed that recordings were most frequently removed by uploaders themselves, followed by account termination and blocking. DMCA takedowns were the least common reason for removal, showing that recordings were seven times more likely to be disabled by Content ID and five times more likely to be removed for violating terms of service (Gray and Suzor, 2020).

Davis (2018) asserted that it benefits users for copyright owners to rely on Content ID and Rights Manager because copyright claims are easier to contest compared to DMCA takedowns (Davis, 2018). Other studies, however, show it is no simple matter to rebut copyright claims from YouTube and Facebook. Furthermore, most users do not engage the dispute procedures.

Kaye and Gray (2021) found that out of 144 videos surveyed, 75 expressed negative sentiments towards copyright enforcement on YouTube, and users formed theories to explain perceived unbalanced approaches to copyright enforcement. Users complained about YouTube favoring large corporations and using claims as a form of punishment in addition to feeling that there is no real recourse for small content creators. Most users also believed that music was more likely to be claimed than any other type of content, which is supported by Gray and Suzor (Gray and Suzor, 2020; Kaye and Gray, 2021).

Berkowitz (2021; 2022) showed further evidence that Content ID and Rights Manager often over-enforce against music uploads. In these studies, accounts of classical music performances were misidentified for copyright infringement, and it described the experiences of musicians in their attempts to challenge claims. None of which were easily accomplished, and some were unsuccessful. Furthermore, it outlined the lengthy dispute and appeals process on YouTube, showing that

it can take over two months to successfully resolve a copyright claim and an additional three months if the result is a copyright strike (Berkowitz, 2021; Berkowitz, 2022).

Copyright strikes are the result of either a failed attempt at appealing a copyright claim or receipt of a DMCA takedown notice. They can be resolved by waiting a period of 90 days, or if it is the first strike incurred by the account, users can complete YouTube’s Copyright School. Other options include contacting the claimant independently and asking them to release the claim or submitting a counter-notification. After three strikes, the account is deleted, and all associated videos are removed (Google, 2022a).

Users and researchers have expressed that Content ID and Rights Manager are not yet able to meet the general public’s expectations in their ability to fairly enforce copyright, yet YouTube claimed in 2016 that Content ID correctly identifies infringing videos with a success rate of 99.7% (Karp, 2016). This figure may refer to true positives, the rate at which content is correctly identified as infringing, or a combination of true positives and true negatives, the rate at which content is correctly identified as non-infringing. In either case, this figure likely does not include false positives, the rate at which content is misidentified as infringing, and false negatives, the rate at which content is misidentified as non-infringing (Lester and Pachamanova, 2017).

The innerworkings of such systems are protected trade secrets and corporate reports lack transparency. Therefore, the best way to learn about their mechanics is through independent testing. Perel and Elkin-Koren (2017) discuss the practice of black-box tinkering, a way of discerning a system’s nature, function, and operability through experimentation. In this study, researchers uploaded infringing and non-infringing content to Israeli websites and documented whether these platforms utilized automated means to remove material or if they waited for takedown notices (Perel and Elkin-Koren, 2017). Zhang et al. (2018a; 2018b) livestreamed both infringing and non-infringing content on YouTube to document how quickly and accurately Content ID would identify and block the test sample of broadcasts. They found that Content ID was unable to recognize 26% of live feeds as infringing after 30 minutes of streaming while blocking 22% of non-infringing footage (Zhang et al., 2018a; Zhang et al., 2018b). This experiment shows evidence of the system’s false negative and false positive rates respectively.

### 3. METHODOLOGY

This experiment continues the practice of black-box tinkering by uploading recordings of Beethoven’s piano sonatas to YouTube and Facebook to ascertain the false

positive rate of their antipiracy systems. Music files<sup>1</sup> were downloaded in MIDI format from Classical Archives,<sup>2</sup> chosen from a study on piano midi datasets (Kong et al., 2022). A computer system for expressive music performance (CSEMP) converted the MIDI files to MP3 format. A CSEMP is a program able to audibly reproduce computer codified notation, such as MIDI or MusicXML, in such a way that imitates authentic instrument sounds and human performance. Notion,<sup>3</sup> described as possessing an advanced expressive model, facilitated experimentation (Kirke and Miranda, 2009).

All MIDI files were opened in Notion, saved as MP3 files, and subsequently reviewed for errors which may have occurred during conversion. Most of the MIDI/MP3 files contained an entire sonata as a single unit. 26 of 32 sonatas had to be separated into individual movements before converting them to video files. Each music file was then opened in a video editing suite and saved as an MP4 showing a continuous black screen. Each file was surveyed for anomalies that may have transpired during conversion.

Next, YouTube and Facebook accounts were created. Each required a phone number and an e-mail address to set-up user profiles. A nondescript, pseudonymous e-mail address was provided. YouTube also required a copy of a government issued photo ID before allowing more than four uploads in a single 24-hour period. As such, a driver license was provided, or otherwise, uploading videos to YouTube would have taken 26 days. Facebook has no such restriction.

A total of 102 videos were uploaded to YouTube and Facebook. No metadata aside from the file names, which did not disclose any details, were included in each upload. Both profiles are still active as of writing this article;<sup>4</sup> however, the Facebook account requires maintenance to prevent deactivation, and as such, may not be available by the time this article is published.

## 4. RESULTS

### 4.1 YOUTUBE

All files were successfully uploaded to YouTube within the span of approximately 3.5 hours. Immediately, Content ID identified instances of copyright infringement (i.e., claims) despite that each video was non-infringing. Because they were claimed during the upload process, copyright enforcement was likely carried out autonomously by Content ID.

Out of 102 recordings, 29 were claimed by purported copyright owners (i.e., claimants). Seven companies were listed as claimants, and those which most frequently issued claims included AdShare on behalf of AudioSparx with nine claims and The Orchard Music on behalf of Pennrose Media with 12 claims. Notably, Content ID correctly identified the composer, composition, and movement for all 29 videos even without accompanying

descriptors. This suggests that Content ID does not rely on any metadata to help identify uploaded material.

YouTube possesses a useful feature that indicates which parts of a video Content ID deems infringing. On average, Content ID claimed 97.36% of content from these recordings, with the highest claimed percentage being 100% and the lowest being 74.34%. Out of the entire corpus, which amounts to just over 10 hours of music, Content ID claimed 2.67 hours (26.6%) for infringement.

Each claim was challenged via the Content ID dispute procedure which is broken down into a few simple steps. It asks the user to select a reason for the dispute from the following choices: fair use, Public Domain, original work owned by the uploader, or permission to distribute from the copyright owner. Public Domain was selected, and the webform included an explanation to support the dispute.<sup>5</sup> Claimants have up to 30 days to respond.

Out of 29 disputes, various claimants approved eight. Of the disputes that were approved, responses took as little as four days and as long as 27 days for delivery. Responses rejecting the other 21 disputes ranged from 12 days to 26 days for delivery. No explanation was given for why these disputes were rejected. Before escalating disputes to appeals, an account must remain active for 30 days. As such, 18 days transpired before the first rejected dispute could be appealed.

During the appeals process, users are asked to read an explanation of the appeals process and a description of the copyright strike. They are also asked to confirm that they understand the terms for filing an appeal. YouTube also requests personally identifying information such as the user's full name and address. Finally, YouTube requires an explanation for why rejected disputes are being appealed.<sup>6</sup> After waiting an additional 30 days, the maximum amount of time allotted to claimants for responses, all 21 appeals were approved.

Claimants have the option to block, monetize, or track the viewership statistics of infringing uploads (Google, 2022b). During this project, claimants did not monetize videos as there were no ads present, nor could the uploader monetize any of the videos, a limitation imposed for the duration of the claims. Claimants also did not block videos since they could still be publicly viewed; although, it is possible that claimants tracked audience metrics. Furthermore, no DMCA takedown notices were issued against any recordings.

### 4.2 FACEBOOK

All files were uploaded to Facebook successfully within the span of three hours. Videos were immediately identified for copyright infringement despite all evidence to the contrary. All allegedly infringing videos were claimed during the upload process, and therefore, copyright enforcement seems to have been implemented automatically by Rights Manager.

Out of 102 videos, 29 received copyright claims, but they were not the same 29 recordings claimed by Content ID. Only 12 of the 29 allegedly infringing videos were shared between the two platforms. Unlike Content ID, Rights Manager shows the names of individuals issuing claims. For the 29 claims, 21 different artists were identified and five were claimed anonymously. An unsuccessful attempt was made to determine if any of the artists were associated with the companies named by YouTube. Like Content ID, Rights Manager was able to correctly label all 29 uploads with the name of the composer, composition, and movement. This shows that Rights Manager also does not rely on metadata to aid in identifying content shared on the platform.

The procedure to dispute claims on Facebook is not as robust as the Content ID dispute system. Each copyright claim notice indicated that 20 seconds of audio had been muted in allegedly infringing videos. The interface then gave the option to post videos partially muted or to post videos entirely unmuted. Restoring the audio required confirmation that relevant rights or permissions to share and distribute the content had been secured. This option was successfully chosen for four videos. The five anonymously claimed videos gave no option to repost, and Facebook rendered them entirely muted. For the remaining 20 claims, Facebook stated, “We are generating new audio for you. You will be able to take action in a few minutes.” After waiting several hours, options to restore audio and post videos partially muted remained inoperable; however, despite having active claims against them, these recordings seem to be publicly available in their entirety.

Only six videos showed which portions of the content were allegedly infringing: the five muted recordings and one claimed by a named artist. In these cases, as little as 13.9% to as much as 88.17% of audio allegedly infringed copyright. Oddly, Rights Manager also indicated that one video contained 14 minutes and 22 seconds of infringing content out of the 7-minute and 4-second runtime, meaning that 203.3% of the recording was claimed. The remaining 23 claims did not indicate the degree of infringement.

There was no visible weblink leading to the copyright claim dispute form, and there was no easy way to find it on the Facebook website. Instead, it was located using a Google search. Once found, the simple form asked for personally identifying information such as a full name and street address, and it asked for the reason for the dispute. The same text used in the Content ID dispute was used here, and links to the claimed videos were added since there was no information connecting the dispute to the claims.

Facebook responded to the dispute via e-mail three days later and requested additional information. In

replying, explanations of the following were offered: the limitations of US copyright law as they pertain to Public Domain content; how recordings were produced; how Rights Manager struggles to distinguish copyrighted recordings from independent recordings of the same underlying compositions. Correspondence also included links to the claimed videos and requests for aid from human moderators. After a string of several e-mails, Facebook ceased communications.

Facebook, in the following weeks, suspended the account for not following its “Community Standards.” When this occurs, the user has 30 days to disagree with Facebook’s decision or the account will be terminated. This procedure requires users to submit their phone numbers to receive a six-digit code which is entered into a webform. It then initiates an “I’m not a robot” test, and finally, it asks users to upload images of themselves. After one to two days, Facebook restored the account. At the time of writing this, the account has been suspended and reinstated 12 times over the course of five months. Because of this issue, it may not be feasible to continue monitoring the account as these suspensions are assumed to persist.

Claimants can react to infringing uploads in four ways. They can block, monetize, advertise, or report videos. Blocking and monetizing on Facebook are similar to YouTube, except that blocking hides recordings from public view while still making them available to the uploader. Advertising allows copyright owners to add a banner promoting their profile, and reporting sends notices to Facebook petitioning to have videos removed (Meta, 2022). Claimants took no action against uploads for this experiment as there were no imbedded ads of any kind, and aside from the five that were muted, videos could still be publicly viewed. Also, no DMCA takedown notices were issued.

## 5. DISCUSSION

As mentioned previously, YouTube reported that Content ID correctly identifies infringing content in 99.7% of instances (Karp, 2016). This likely only reflects true positives; although it is uncertain if this number also includes true negatives, it is unlikely that this figure accounts for false positives and false negatives (Lester and Pachamano, 2017). Unfortunately for users and researchers, neither YouTube nor Facebook are meaningfully transparent in their public reports. Furthermore, studies showed that Content ID’s false positive rate was 22%, and its false negative rate was 26% when testing both infringing and non-infringing broadcasts (Zhang et al., 2018a; Zhang et al., 2018b). These reported figures were used to form baseline expectations for Content ID’s and Rights Manager’s performances throughout this experiment.<sup>7</sup>

## 5.1 YOUTUBE

Trying to measure a false positive rate for Content ID from this data can be misleading as there are many variables to consider (e.g., tempi, timbre, audio quality, reverberation, instrumentation, etc.). The test data used for this experiment could be reasonably described as generic renditions of Beethoven's piano sonatas. All recordings were rendered using audio samples from a grand piano, the tempo for each synthesized recording follows conventional pacing for their respective compositions, and the sound quality simulates a moderate reverberation similar to a recital hall. Altering any of these characteristics to extreme ends of the spectrum could have significant impacts on results.

Findings showed that Content ID misidentified 29 of 102 video uploads, incorrectly labeling them as infringing. Furthermore, YouTube also indicated the proportion of allegedly infringing content for each video that received a copyright claim. 2.67 hours out of the approximately 10 hours of music was incorrectly labeled infringing. For this specific sample, the false positive rate was shown to be 28.4% for misidentified uploads and 26.6% for incorrectly claimed audio runtime. Because none of the recordings were actually infringing, it is not possible to calculate the performance rate for true positives nor false negatives.

Calculating the true negative performance rate from this data can be equally misleading. On the surface, Content ID correctly identified the remaining 73 uploads and 7.33 hours of music as non-infringing which would equate to a true negative performance rate of 71.6% and 73.4% respectively. However, in addition to the aforementioned variables, it is not clear if these videos were rightfully identified as non-infringing or if recordings of the remaining compositions were absent from YouTube's reference database.

In comparison to baseline expectations, this study's reported false positive rate of 28.4% or 26.6% is slightly greater than the 22% reported by Zhang et al. (2018a; 2018b). The greatest difference between the two studies is the former focused on computer-synthesized recordings of Beethoven's piano sonatas and the latter tested video broadcasts of live sporting events and concerts (Zhang et al., 2018a; Zhang et al., 2018b). It is possible that Content ID is better at screening livestreams, but more likely, Content ID experiences greater difficulty when rendering decisions pertaining to Public Domain music. If the true negative rating is taken at face value, the results from this study, a performance rate of 71.6% or 73.4%, were considerably lower than YouTube's self-reported 99.7% performance rating.

## 5.2 FACEBOOK

Facebook also misidentified 29 of 102 uploads as infringing, rendering the same results as Content ID, 28.4% for false positives. For the same reasons, performance ratings for true positives and false negatives cannot be calculated

based on the data available. To reiterate, attempting to render a true negative performance rating based on this experiment could be misleading, but if it is assumed that the information available is accurate, Rights Manager's true negative performance rating for this test sample would likewise be 71.6%.

Facebook did not consistently indicate which portions of claimed videos infringed copyright the way that YouTube did, and so, comparisons cannot be drawn based on this metric. Additionally, for the few videos that did indicate which portions of each recording allegedly infringed copyright, the information provided appeared unreliable. Content ID and Rights Manager performed nearly the same regarding false positives and true negatives during active scanning of the test sample; however, they did not issue claims against the same 29 recordings. This further suggests that the true negative rate for each system is likely inaccurate due to the absence of recordings from their respective reference libraries. Both systems show the greatest similarity in the way they issued claims. Content ID and Rights Manager claimed videos during the upload process, accurately identified composers and compositions for each claimed upload without relying on metadata, and provided similar information about claimants.

Berkowitz (2021) asserted, "Since Facebook's Rights Manager is designed to work similarly to YouTube's Content ID, users should expect uploads to be monitored and treated the same way on Facebook as they are on YouTube," (Berkowitz, 2021: 191). This statement was shown to be reasonably accurate, but where the two significantly diverge is the way they handle disputes. Based on this analysis, YouTube is the more user-friendly platform, providing comprehensible and relatively effective recourse for musicians.

## 5.3 IMPROVEMENTS

Scholars have debated automating fair use reviews. One study suggests, "Courts should allow copyright owners, with computerized methods comparable to YouTube's Content ID, to form a good faith belief solely through the use of computerized methods," (Davis, 2018: 262). Other previously mentioned studies disagree. Meanwhile, the courts have wavered on the issue. "The Ninth Circuit released an amended opinion in *Lenz* no longer discussing the effects of its holding on computer algorithm systems. The fact that the court redacted language hinting that computer algorithms are capable of sufficiently considering fair use suggests that computer algorithms may in fact not be safe under its holding," (Zapata-Kim, 2016: 1865).

The rate of false positives demonstrated in this study suggests that algorithmic systems are currently ill equipped to render decisions on fair use. US courts are not unanimous in their rulings on fair use because of its stochastic nature, and therefore, a system which cannot

separate recorded music from underlying compositions should not be relied upon to make such decisions on their own. Perhaps with human intervention, transparency from corporations, and significant improvements to software powering copyright enforcement systems, good faith affirmations and fair use reviews can be automated.

Experimentation in this article reveals that Content ID and Rights Manager do not incorporate linked data technologies. No metadata was provided to YouTube nor Facebook, yet all recordings mischaracterized as infringing were accurately labeled with the correct titles and composer. Content ID and Rights Manager seem to rely solely on active scanning for content recognition. Embedding linked data algorithms would create a more robust system.

Algorithms could analyze user-generated content to find matching copyrighted recordings from the system's reference database in the way they already do, and the metadata from the matching copyrighted content would serve as a built-in controlled vocabulary which eliminates the need to rely upon user-generated descriptors. Linked data algorithms could utilize metadata from copyrighted recordings to search reliable databases for corroborating information, confirming the presence of Public Domain content. For example, [WorldCat.org](http://WorldCat.org)<sup>8</sup> adheres to linked data standards and is trusted by a global community of librarians and archivists. Upon recognizing user uploads as Public Domain music, additional algorithms could conduct waveform analyses to determine whether the music is an unauthorized reproduction or a rendition of the underlying composition. This should be feasible considering Alphabet and Meta are industry leaders in this area.

YouTube offers a feature that indicates which parts of an allegedly infringing upload has been claimed. Users noted the benefits of this feature saying, "Before, YouTubers had no idea what was being claimed, but now they know exactly what is being claimed... this is a step in the right direction," (Kaye and Gray, 2021: 6). While there was some evidence of this feature in Facebook's copyright claims, it was inconsistently applied and did not provide useful indicators. Furthermore, links to dispute webforms should be embedded in their copyright notices. Facebook is more than capable of reproducing YouTube's success.

Davis (2018) has suggested that amendments to the DMCA include a statutory requirement for copyright owners to prove that they are acting in good faith and have conducted a fair use review before issuing takedown notices (Davis, 2018). Zapata-Kim (2016) recommends that online service providers employing automated copyright enforcement be held to the same standard (Zapata-Kim, 2016). While modifications to the DMCA may be warranted, pivoting to logistical improvements that YouTube and Facebook could enact may satisfy these calls for change.

Berkowitz (2021; 2022) reported occurrences of YouTube and Facebook users encountering situations where multiple claims from different entities were simultaneously levied against a single recording (Berkowitz, 2021; Berkowitz, 2022). This can occur when content creators rely on fair use and material from the Public Domain. Systems should prompt human moderators to conduct fair use and Public Domain assessments under such circumstances which would satisfy the good faith belief and fair use review requirements outlined in the DMCA and *Lenz v. Universal*.

Lester and Pachamano (2017) encourage greater transparency from technology and media companies regarding the performance of automated antipiracy systems. YouTube mentions only how accurate Content ID is at identifying true positives (and possibly true negatives); however, reports do not contain information about false positive and false negative performance (Lester and Pachamano, 2017). Additionally, it would be worthwhile to report metrics such as the percentage of copyright claims successfully or unsuccessfully disputed and appealed, and it would also be useful to know the most and least common reasons for disputes by category (i.e., fair use, Public Domain, original content, and permission from copyright owners). Furthermore, platforms should offer their users explanations for why disputes and appeals are rejected.

It is also important to educate the public and persist in researching such matters. Two major video sharing platforms were not included in this project. Experimenting with TikTok was unsuccessful as the feature allowing for 10-minute videos is only accessible via the mobile app. Using TikTok's website limits recordings to three minutes which is impractical for classical music performance. Instagram, which is also monitored by Rights Manager, rendered errors when uploading videos. The reasons for this are unclear. It was also noticed that very few studies mention Facebook which likewise necessitates further study.

## 6. CONCLUSION

It has been almost 70 years since the debut of Hiller and Isaacson's *Illiac Suite* and Klein and Bolith's "Push Button Bertha." These were among the first successful attempts to produce music entirely synthesized via computer automation. Many techniques for autonomously generating music have since been explored such as symbolic knowledge-based systems, Markov chains, and artificial neural networks. Deep learning models are currently the most favored approach (Micchi et al., 2021).

Software applications designed to enable users to create their own music through controlled parameters and descriptive inputs have permeated the market in recent years. Melomics Media began its operations



in 2011 as a technology-based startup and brings AI generated print and recorded music to the public for general use via its Iamus “computer composer” (Sanchez Quintana et al., 2013). In 2016, co-founders Pierre Barreau, Denis Shtefan, and Vincent Barreau launched their auto-generative music program called Artificial Intelligence Virtual Artist (AIVA). This “AI composer” allows consumers to modify a variety of parameters such as mood, note density, genre, and style to render unique audible music based on their evolving needs, situations, activities, and disposition. As Pierre Barreau put it, “A personalized live soundtrack based on their story and their personality,” (Aiva Technologies SARL, 2018). Endel was launched in 2018 and is a generative music app compatible with several devices such as those powered by Apple’s iOS and Amazon’s Alexa. Endel’s model generates soundscapes in real time based on the user’s biometrics and geographic location, including the time of day and weather forecasts (Endel, 2022).

Technology that produces entire musical works at the press of a button, and applications that constantly generate and regenerate music instantaneously show how computer synthesis surpasses human productivity. Collins (2018) discusses how “billion song datasets” are easily created with generative AI. For comparison, the size of song catalogues offered by streaming services like Spotify and iTunes have been estimated to number between 30 and 65 million pieces of music. Their experiment shows how they were able to generate one billion unique melodies in approximately an hour and suggests that their AI program could create more music than the cumulative human output in a matter of days. The time it takes to listen to autogenerated music far exceeds the brief moments needed for algorithmic composing, and therefore, such technology may bring about market saturation through mass production (Collins, 2018).

Spotify propagated music by “fake artists” on its own platform to claim ad revenue for itself and reduce overall royalty payouts to unaffiliated musicians (Goldschmitt, 2020). Although this was done with in-house human composers, “media commentators saw in this incident an augur of the way platforms might employ cheap AI-generated music to replace more costly human-created tracks,” (Drott, 2021: 205). Music generative programs like JukeDeck aim to provide video content creators with royalty free works, protecting them from copyright claims online; however, many YouTube users are skeptical of such promises, fearing that sudden changes to licensing agreements would force them to pay fees for music already added to their videos. Likewise, they worry that noncompliance would trigger copyright claims which would redirect ad revenue (Collins, 2018; Kaye and Gray, 2021).

Corporations like Meta and Alphabet operate social media platforms, employ automated copyright enforcement systems, and invest in generative AI programs. This

combination of technology makes it possible for them to flood their websites with autogenerated music under the guise of fake artists. Furthermore, their ability to autonomously enforce copyright via their Rights Manager and Content ID systems may tempt them to store samples of autogenerated music in their antipiracy reference libraries. Doing so would allow Meta and Alphabet to leverage Rights Manager’s and Content ID’s monetization features to automatically appropriate ad revenue from content creators by claiming uploaded content similar to their own autogenerated samples.

Studies have suggested that few artists would attempt to dispute these claims, and the experiment presented in this article has shown how difficult and tedious it can be to resolve claims on Facebook and YouTube, sometimes without success. Without the incentive to justify rejections to disputes and appeals, any challenges to claims can be disregarded in order to continue profiting from the creative efforts of users. Such circumstances may give rise to a technocratic system where the freedoms of individuals are subservient to their consumer/prosumer roles and the surplus value of the content they create. As such, these companies would reap the benefits of free labor while controlling the scarcity and value of music on their platforms (Prey, 2015).

Morreale (2021) advocates for the consideration of how AI music technologies may negatively impact society as technologists continue to fund and develop algorithmic systems (Morreale, 2021). Ignoring this responsibility may result in unintended consequences like having rights to fair use restricted and the Public Domain privatized. Echoing these conclusions, the author encourages continued discussion and engagement in the ethical development and employment of algorithmic music recognition and copyright enforcement systems.

## NOTES

- 1 Recordings have been provided in MIDI format for future testing with different CSEMPs and online media platforms; [https://adamericberk.weebly.com/uploads/9/9/7/5/99756384/test\\_data.zip](https://adamericberk.weebly.com/uploads/9/9/7/5/99756384/test_data.zip).
- 2 [www.classicalarchives.com](http://www.classicalarchives.com), 2022.
- 3 <https://apps.apple.com/us/app/notion/id475820434>, 2022.
- 4 <https://www.youtube.com/channel/UCDyO1qXYUb7D4Lkene0G5yg/videos>, 2022; <https://www.facebook.com/profile.php?id=100080834802491>, 2022.
- 5 [https://adamericberk.weebly.com/uploads/9/9/7/5/99756384/copyright\\_claim\\_dispute\\_text.pdf](https://adamericberk.weebly.com/uploads/9/9/7/5/99756384/copyright_claim_dispute_text.pdf).
- 6 [https://adamericberk.weebly.com/uploads/9/9/7/5/99756384/copyright\\_claim\\_appeal\\_text.pdf](https://adamericberk.weebly.com/uploads/9/9/7/5/99756384/copyright_claim_appeal_text.pdf).
- 7 [https://adamericberk.weebly.com/uploads/9/9/7/5/99756384/recorded\\_results.xlsx](https://adamericberk.weebly.com/uploads/9/9/7/5/99756384/recorded_results.xlsx).
- 8 <https://www.worldcat.org/>, 2022.

## COMPETING INTERESTS

The author has no competing interests to declare.

## AUTHOR AFFILIATION

**Adam Eric Berkowitz**  [orcid.org/0000-0001-8340-4745](https://orcid.org/0000-0001-8340-4745)  
Tampa-Hillsborough County Public Library, US

## REFERENCES

- Aiva Technologies SARL.** (2018). About AIVA. Retrieved November 11, 2022, from <https://www.aiva.ai/about#about>.
- Bar-Ziv, S. and Elkin-Koren, N.** (2017). Behind the scenes of copyright enforcement: Empirical evidence on notice and takedown. *Connecticut Law Review*, 50(2), 3–45.
- Bartholomew, T. B.** (2015). The death of Fair Use in cyberspace: YouTube and the problem with Content ID. *Duke Law & Technology Review*, 13(1), 66–88.
- Berkowitz, A. E.** (2021). Are YouTube and Facebook canceling classical musicians? The harmful effects of automated copyright enforcement on social media platforms. *Notes*, 78(2), 177–202. DOI: <https://doi.org/10.1353/not.2021.0083>
- Berkowitz, A. E.** (2022). Classical musicians v. copyright bots: How libraries can aid in the fight. *Information Technology and Libraries*, 41(2), 1–9. DOI: <https://doi.org/10.6017/ital.v41i2.14027>
- Borgsmiller, K.** (2019). YouTube vs. the music industry: Are online service providers doing enough to prevent piracy? *Southern Illinois University Law Journal*, 43, 647–673.
- Brodeur, M. A.** (2020, May 21). Copyright bots and classical musicians are fighting online. The bots are winning. *The Washington Post*. Retrieved August 24, 2022, from [https://www.washingtonpost.com/entertainment/music/copyright-bots-and-classical-musicians-are-fighting-online-the-bots-are-winning/2020/05/20/a11e349c-98ae-11ea-89fd-28fb313d1886\\_story.html](https://www.washingtonpost.com/entertainment/music/copyright-bots-and-classical-musicians-are-fighting-online-the-bots-are-winning/2020/05/20/a11e349c-98ae-11ea-89fd-28fb313d1886_story.html).
- Collins, N.** (2018). There is no reason why it should ever stop: Large-scale algorithmic composition. *Journal of Creative Music Systems*, 3, 1–25. <https://search.informit.org/doi/10.3316/informit.667494634199085>. DOI: <https://doi.org/10.5920/jcms.525>
- Davis, S. M.** (2018). Computerized takedowns: A balanced approach to protect fair uses and the rights of copyright owners. *Roger Williams University Law Review*, 23(1), 229–264.
- Drott, E.** (2021). Copyright, compensation, and commons in the music AI industry. *Creative Industries Journal*, 14(2), 190–207. DOI: <https://doi.org/10.1080/17510694.2020.1839702>
- Endel.** (2022). Endel: Personalized soundscapes to help you focus, relax, and sleep. Retrieved November 11, 2022, from <https://endel.io/>.
- Goldschmitt, K. E.** (2020). The long history of the 2017 Spotify “fake music” scandal. *American Music*, 38(2), 131–152. DOI: <https://doi.org/10.5406/americanmusic.38.2.0131>
- Google.** (2018, November). How Google fights piracy. Retrieved August 24, 2022, from [https://blog.google/documents/27/How\\_Google\\_Fights\\_Piracy\\_2018.pdf/](https://blog.google/documents/27/How_Google_Fights_Piracy_2018.pdf/).
- Google.** (2022a). Copyright strike basics: YouTube. Retrieved August 24, 2022, from <https://support.google.com/youtube/answer/2814000#zippy=%2Cwhat-happens-when-you-get-a-copyright-strike%2Cresolve-a-copyright-strike>.
- Google.** (2022b). Learn about Content ID claims: YouTube. Retrieved August 24, 2022, from [https://support.google.com/youtube/answer/6013276?hl=en&ref\\_topic=9282678](https://support.google.com/youtube/answer/6013276?hl=en&ref_topic=9282678).
- Gorwa, R., Binns, R., and Katzenbach, C.** (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1). DOI: <https://doi.org/10.1177/2053951719897945>
- Gray, J. E. and Suzor, N. P.** (2020). Playing with machines: Using machine learning to understand automated copyright enforcement at scale. *Big Data & Society*, 7(1). DOI: <https://doi.org/10.1177/2053951720919963>
- Jacques, S., Garstka, K., Hviid, M., and Street, J.** (2018). An empirical study of the use of automated anti-piracy systems and their consequences for cultural diversity. *SCRIPTed*, 15(2), 277–312. DOI: <https://doi.org/10.2966/scrip.150218.277>
- Karp, H.** (2016, June 28). Industry out of harmony with YouTube on tracking of copyrighted music. *The Wall Street Journal*. Retrieved October 26, 2022, from <https://www.wsj.com/articles/industry-out-of-harmony-with-youtube-on-tracking-of-copyrighted-music-1467106213>.
- Kaye, D. B. V. and Gray, J. E.** (2021). Copyright gossip: Exploring copyright opinions, theories, and strategies on YouTube. *Social Media + Society*, 7(3), 1–12. DOI: <https://doi.org/10.1177/205630512111036940>
- Keef, A. T. and Ben-Kereth, L.** (2016, April 12). Introducing Rights Manager. Retrieved August 24, 2022, from <https://www.facebook.com/formedia/blog/introducing-rights-manager>.
- Kirke, A. and Miranda, E. R.** (2009). A survey of computer systems for expressive music performance. *ACM Computing Surveys*, 42(1), 1–41. DOI: <https://doi.org/10.1145/1592451.1592454>
- Kong, Q., Li, B., Chen, J., and Wang, Y.** (2022). GiantMIDI-Piano: A large-scale MIDI dataset for classical piano music. *Transactions of the International Society for Music Information Retrieval*, 5(1), 87–98. DOI: <https://doi.org/10.5334/tismir.80>
- Lawrence-Williams, S.** (2022). Regulating automatic content recognition software in Canada. *Intellectual Property Journal*, 34(3), 317–340.
- Lester, T. and Pachamano, D.** (2017). The dilemma of false positives: Making Content ID algorithms more conducive to fostering innovative Fair Use in music creation. *UCLA Entertainment Law Review*, 24(1), 51–73. DOI: <https://doi.org/10.5070/LR8241035525>

- Library of Congress and US Copyright Office.** (n.d.). Constitution annotated. Retrieved October 28, 2022, from <https://constitution.congress.gov/browse/article-1/section-8/clause-8/>.
- Lorenzon, M.** (2018, December 20). Why Is Facebook muting classical music videos? ABC Classic FM. Retrieved August 24, 2022, from <https://www.abc.net.au/classic/read-and-watch/music-reads/facebook-copyright/10633928>.
- Meta.** (2022). What tools does Facebook provide to help me protect my intellectual property in my videos? Facebook. Retrieved August 24, 2022, from <https://www.facebook.com/help/348831205149904>.
- Micchi, G., Bigo, L., Giraud M., Groult R., and Levé F.** (2021). I Keep Counting: An experiment in human/AI co-creative songwriting. *Transactions of the International Society for Music Information Retrieval*, 4(1), pp. 263–275. DOI: <https://doi.org/10.5334/tismir.93>
- Morreale, F.** (2021). Where does the buck stop? Ethical and political issues with AI in music creation. *Transactions of the International Society for Music Information Retrieval*, 4(1), 105–113. DOI: <https://doi.org/10.5334/tismir.86>
- Perel, M. and Elkin-Koren, N.** (2017). Black box tinkering: Beyond disclosure in algorithmic enforcement. *Florida Law Review*, 69(1), 181–221.
- Prey, R.** (2015). “Now Playing. You”: *Big Data and the Production of Music Streaming Space*. Dissertation. Simon Fraser University, Burnaby, BC.
- Quilter, L. and Heins, M.** (2007). Intellectual property and free speech in the online world: How educational institutions and other online service providers are coping with cease and desist letters and takedown notices. <http://fairusenetwork.org/resources/OSPreport-2007.pdf>.
- Rae, C.** (2021). *Music Copyright: An Essential Guide for the Digital Age*. Rowman & Littlefield Publishing Group.
- Reis, A. J. and Burns, M. L.** (2020). Who owns that tune? Issues faced by music creators in today’s content-based industry. *Landslide*, 12(3). Retrieved August 24, 2022, from [https://www.americanbar.org/groups/intellectual\\_property\\_law/publications/landslide/2019-20/january-february/who-owns-tune-issues-faced-music-creators-todays-contentbased-industry/](https://www.americanbar.org/groups/intellectual_property_law/publications/landslide/2019-20/january-february/who-owns-tune-issues-faced-music-creators-todays-contentbased-industry/).
- Reymond, M. J.** (2016). Lenz v Universal Music Group: Much ado about nothing. *International Journal of Law and Information Technology*, 24(2), 119–127. DOI: <https://doi.org/10.1093/ijlit/eav021>
- Sanchez Quintana, C., Moreno Arcas, F., Albarracin Molina, D., Fernandez Rodriguez, J. D., and Vico, F. J.** (2013). Melomics: A case-study of AI in Spain. *AI Magazine*, 34(3), 99–103. DOI: <https://doi.org/10.1609/aimag.v34i3.2464>
- Seng, D.** (2021). Copyrighting copywrongs: An empirical analysis of errors with automated DMCA takedown notices. *Santa Clara High Technology Law Journal*, 37(2), 119–192.
- Seng, D. K. B.** (2014). The state of the discordant union: An empirical analysis of DMCA takedown notices. *Virginia Journal of Law & Technology*, 18, 369–428. DOI: <https://doi.org/10.2139/ssrn.2411915>
- Soha, M. and McDowell, Z. J.** (2016). Monetizing a meme: YouTube, Content ID, and the Harlem Shake. *Social Media + Society*, 2(1), 1–12. DOI: <https://doi.org/10.1177/2056305115623801>
- Solomon, L.** (2015). Fair users or content abusers? The automatic flagging of non-infringing videos by Content ID on YouTube. *Hofstra Law Review*, 44(1), 237–268.
- Stim, R.** (2019). What is Fair Use: Copyright overview. Retrieved October 30, 2022, from <https://fairuse.stanford.edu/overview/fair-use/what-is-fair-use/>.
- U.S. Copyright Office.** (n.d.a). How long does copyright protection last? Retrieved August 24, 2022, from <https://www.copyright.gov/help/faq/faq-duration.html>.
- U.S. Copyright Office.** (n.d.b). U.S. Copyright Office Fair Use Index. Retrieved October 29, 2022, from <https://www.copyright.gov/fair-use/>.
- Urban, J. M., Karaganis, J., and Schofield, B. L.** (2017). *Notice and Takedown in Everyday Practice*. BerkeleyLaw, University of California, and American Assembly, University of Columbia. DOI: <https://doi.org/10.31235/osf.io/59m86>
- Urban, J. M. and Quilter, L.** (2006). Efficient process or chilling effects: Takedown notices under Section 512 of the Digital Millennium Copyright Act. *Santa Clara High Technology Law Journal*, 22(4), 621–693. DOI: <https://doi.org/10.31235/osf.io/pyzua>
- Zapata-Kim, L.** (2016). Should YouTube’s Content ID be liable for misrepresentation under the Digital Millennium Copyright Act? *Boston College Law Review*, 57(5), 1847–1874.
- Zhang, D. Y., Badilla, J., Tong, H., and Wang, D.** (2018). An end-to-end scalable copyright detection system for online video sharing platforms. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 626–629. DOI: <https://doi.org/10.1109/ASONAM.2018.8508288>
- Zhang, D. Y., Li, Q., Tong, H., Badilla, J., Zhang, Y., and Wang, D.** (2018). Crowdsourcing-based copyright infringement detection in live video streams. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 367–374. DOI: <https://doi.org/10.1109/ASONAM.2018.8508523>

---

**TO CITE THIS ARTICLE:**

Berkowitz, A. E. (2023). Algorithmic (In)Tolerance: Experimenting with Beethoven's Music on Social Media Platforms. *Transactions of the International Society for Music Information Retrieval*, 6(1), 1–12. DOI: <https://doi.org/10.5334/tismir.148>

**Submitted:** 26 August 2022    **Accepted:** 27 November 2022    **Published:** 03 January 2023

**COPYRIGHT:**

© 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

*Transactions of the International Society for Music Information Retrieval* is a peer-reviewed open access journal published by Ubiquity Press.

