UNIVERSAL MUSIC GROUP

JEFFREY S. HARLESTON
GENERAL COUNSEL AND EXECUTIVE VICE PRESIDENT
BUSINESS AND LEGAL AFFAIRS

December 3, 2024

Suzanne Wilson, Esq.
General Counsel and Associate Register of Copyrights
U.S. Copyright Office
101 Independence Avenue, SE
Washington, DC 20559-6003

Dear Ms. Wilson:

During our *ex parte* meeting on April 22, 2024, regarding the Copyright Office's NOI on AI and Copyright, you noted that the Office is carefully watching litigation in the area and invited us to share any additional information that might be helpful to your inquiry.

As you are aware, several recording companies, including Universal Music Group, have sued certain music generative AI companies ("MAIs") for copyright infringement. We presume the Office has read our complaints against two MAIs, Suno and Udio, as well as the answers filed by each company. Those answers and various public statements by the defendants offer new thematic arguments and justifications for unlicensed training on copyrighted sound recordings and are rife with specious analogies and misstatements. We write to address them directly with the Office, because (1) these companies were launched after the submission of our comments filed with the Office; (2) their highly atypical litigation responses signal an attempt to influence policymakers, (3) their arguments necessitate a corrective response, and (4) we do not anticipate our litigation timeline (and ability to respond to their arguments) will align with the Office's AI study timeline.

General Counsel and Executive Vice President of Business & Legal Affairs

2220 COLORADO AVENUE SANTA MONICA CA 90404 TEL 310 865 8981 FAX 310 865 5613
E-MAIL jeff.harleston@umusic.com
www.umusic.com
A VIVENDI COMPANY

I. THE NATURE OF MAIs

A. What MAIs Do

Until recently, the legal and policy discussions concerning AI have focused principally on general purpose AI models, such as chatbots that respond to human prompts with plain-language outputs. The models offered by MAI are far more specialized, however, and have the singular purpose of generating audio files that simulate music created by humans. They are not designed, for example, to facilitate academic research or scientific discovery, but to make it simple and nearly instantaneous for anyone to generate files that sound like, but are not, the products of human musical expression. Indeed, MAIs often tout the fact that users need no musical training or skill to generate credible sounding "music." *See, e.g., https://suno.com/about* ("Whether you're a shower singer or a charting artist, we break barriers between you and the song you dream of making. No instrument needed, just imagination. From your mind to music"); *https://www.aiva.ai/* ("Whether a complete beginner or seasoned professional in music making, use the power of generative AI to create your own songs"); *www.boomy.com* ("Create original songs in seconds, even if you've never made music before"); *https://www.loudly.com/blog/ai-music-generator-from-text* ("Imaging being able to transform your words into melodies, your ideas into symphonies, with just a few clicks … Say goodbye to the time-consuming process of traditional composition and hello to instant creativity").

Although multiple variations of MAI models are emerging, they typically work as follows. By inputting plain-English text prompts and/or indicating certain preselected parameters, users select various characteristics of sound files they want the MAI model to generate. They can specify, for example, the desired genre, beat, instrumentation, lyrics, style, or historical era of music, and the model will generate sound files conforming to those specifications. *See, e.g., https://www.udio.com/about-us* ("udio.com allows users to create music from simple text prompts by specifying topics, genres, and other descriptors

which are then transformed into professional quality tracks"). Users can even specify the style of a particular artist they want the resulting sound file to resemble. *See, e.g., https://www.udio.com/guide* ("a prompt can have a reference to an artist style … For example … 'a song about summer rain, jazz, mellow, warm, in the style of Billie Holiday'").

Notably, and consistent with large language models generally, it is the MAI that digitally generates the sound file in response to the uncopyrightable ideas expressed in the prompt. As one popular service explains:

> No matter how detailed text prompts cannot fully define an actual piece of music - the same text describes an infinite number of possible audio tracks. That's why to get close to the music idea you have in your head you might want to create multiple clips with identical prompts (and other settings). Udio by default produces two outputs for each input you provide to the model.

https://www.udio.com/guide. In other words, while text prompts can describe abstractly the desired general characteristics and genre of a desired sound file, it is the MAI algorithm that digitally determines the features and expressive content of the resulting output. In this sense, these audio file outputs are no different than the outputs of other AI services – they are the uncopyrightable products of digital generation rather than the copyrightable product of human creativity. *See* Initial Comments of Universal Music Group ("UMG's Initial Comments") at 75-79 (setting forth UMG's view on why AI outputs are not copyrightable).

B. How MAIs are Trained

Most MAIs have been decidedly opaque about their training materials and procedures, but the basic training process is similar to how AI models train on text. Because MAIs aim to produce outputs that sound like real recorded music, their models are trained on sound recordings and the musical compositions embodied in those recordings. The MAIs first copy sound recordings from digital sources (discussed further below) and then clean up the collection of copied recordings to eliminate unwanted materials, such as poor-quality recordings or duplicates.

A generative AI can be thought of as a giant equation with a huge set of parameters. Given a text prompt as input, an ideal set of parameters will generate music as output that matches the prompt. During

training, the AI undergoes a mathematical process that uses the copied recordings to refine the AI's parameters so that it gets better at generating musical output that sounds appropriate to a given text or musical input prompt. During this process, the AI is iteratively refined so it is able to generate outputs that more closely matches the quality of the training material; in other words, the quality of the output is directly related to the quality of the training material.

MAI models may be subjected to additional "fine-tuning," typically using a different set of recordings, to improve the quality of the output or to give it some characteristic of interest (e.g. the ability to generate artificial music that sounds like a particular artist). This process also requires the use of copies of pre-existing recordings.

The goal of this entire process is to establish the statistical correlations and patterns between and among musical features or characteristics so that the MAI model learns how predict what an "80s upbeat pop song in the style of Michael Jackson" should sound like. The model might average various correlations it has observed in a given genre or perhaps the style of a given artist. But, critically, it does not "compose" music anew in the human sense by making creative choices or aesthetic judgments. Rather, it assembles a digital facsimile of human music based on patterns and statistical relationships it has observed in prior recorded music. It is a computer program and not an author.

While the training materials are broadly sound recordings, the locations from where they are sourced differ. Some MAIs are at pains to ensure that they do not train on third party copyrighted content out of a stated respect for the law and rights of creators. *See, e.g.,* *https://blog.soundraw.io/post/soundraw-joins-aiformusic-responsible-ai-music-creation* ("At SOUNDRAW, real producers create original beats in-house to train our AI model and therefore consider ourselves an ethical AI platform. We never train our AI with other artists' music or sounds. This ensures everything on the platform is born from our original content, not borrowed."); *https://rightsify.com/hydra/* ("Hydra cannot generate music based on existing songs or melodies. It was trained exclusively on Rightsify-owned catalog. Rightsify is committed to respecting copyright, and the Hydra dataset is limited

to Rightsify's data to ensure the uniqueness and legality of the generated music");

https://www.loudly.com/ai-music-generator

("By following strict ethical AI guidelines, Loudly guarantees that its proprietary music dataset has been carefully developed via consent, transparency and copyright compliance"). The emergence and rapid improvement of such ethically trained MAIs proves that non-infringing training is possible and practicable.

Many MAIs, however, are not so scrupulous and simply scrape recordings from the Internet without regard to license or authorization. *See, e.g., https://suno.com/blog/future-of-music* ("We train our models on medium- and high-quality music we can find on the open internet … [which] indeed contains copyrighted materials, and some of it is owned by major record labels"); *https://www.udio.com/blog/ai-and-the-future-of-music* ("... our model has 'listened' to and learned from a large collection of recorded music").

For this second and more ubiquitous category of MAIs, unauthorized copying of copyrighted sound recordings, and indeed UMG's sound recordings, is one of the key lynchpins of their multimillion-dollar businesses. By way of comparison, some AI models claim to have trained on essentially all available data of a given media, such as text or images, making the contribution of any one copyright owner's content a small fraction of the universe of training data. MAIs are different, because they train on only one kind of data, sound recordings, and because there is a necessary premium on higher quality sound recordings. The better the training materials – the best recordings by the most beloved artists -- the more convincing and appealing the outputs. And because UMG owns or controls such a quantitatively and qualitatively significant portion of the best recorded music in the world, those MAIs could not credibly simulate popular music and attract the staggering investments they have without pilfering UMG's cherished sound recordings. UMG's catalog, in other words, is not merely a "drop in the bucket" of training data but an essential building block in the MAIs infringing services.

C. Sound Files Generated by MAIs Have Flooded the Market

MAIs are breathtakingly prolific in producing audio files. Even back in September of 2023, before the most popular MAIs had emerged, reports from some of the developers estimated the presence of 170 million MAI-generated tracks. *See https://www.musicbusinessworldwide.com/these-ai-music-generators-claim-to-have-created-least-170-million-tracks/.* Since that time, the quality of outputs has improved to the point where some are nearly indistinguishable from real recordings, and they are proliferating at an enormous rate. They have flooded streaming platforms, imposing the dangers of "content oversupply" discussed in UMG's Initial Comments at 12-13, and displacing real music created by real human beings by diverting royalties away from those artists.

For example, one disturbing phenomenon is the explosion of AI-created audio files uploaded to Spotify and held out as recordings by well-known artists. *See https://idioteq.com/ai-generated-metalcore-floods-spotify/.* Apart from the obvious deception to listeners, these tracks (often falsely included in "playlists" of the named band) divert streaming income away from human artists and to the accounts of internet fraudsters who, with the push of a button, generate competitive content. In one recent instance, abuse of MAIs has become the instrument of criminal conduct. At the end of August 2024, the U.S. Attorney's Office in the Southern District of New York indicted a defendant for uploading hundreds of thousands of AI-generated recordings to the major streaming platforms, using bots to stream those recordings, and fraudulently collecting over $10 million in royalties. *See U.S. v. Smith,* 24 Crim. 504, pending in the United States District Court for the Southern District of New York.

Where this synthetic "music" is so easy to pass off as the real music by real people, the threats to the recorded music industry are impossible to overstate. The unlicensed use of the artistry of real musicians and performers in the training of these systems has become the key building block for technology that competes with and crowds out human composers and performing artists and the industry that supports and promotes them.

II. MAI'S FLAWED ARGUMENTS

Faulty MAI Premise No. 1: "Intermediate Copying To Create Non-Infringing Outputs Is Always Fair Use"

Like other AI companies, the MAIs assert that training on copyrighted materials is fair use, because it represents only "intermediate copying" as an interim step to generating non-infringing content. UMG addressed the fallacy in this argument in its original submission and incorporates that discussion here. *See* UMG's Initial Comments at 49-51. However, the MAIs take this theory one step further and insist more broadly that there is an unbroken line of precedent that "intermediate" copying is *always* permissible if it technologically leads to a non-infringing "output":

> Under longstanding precedent, it is fair use to make a copy of a protected work as part of a back-end technological process, invisible to the public, in the service of creating an ultimately non-infringing new product … no case has ever—not one single time—reached a contrary conclusion. Each time the question has been presented—and it has been presented over and over and over again—the ultimate conclusion has been that making an "intermediate" copy of a protected work, in the service of generating noninfringing outputs, is permissible, not actionable.

See Answer Of Defendant Uncharted Labs, Inc. To Complaint (emphasis added), *UMG Recordings, Inc. v. Uncharted Labs, Inc.*, Case No. 1:24-cv-04777-AKH, United States District Court for The Southern District Of New York ("Udio Answer") at 4. **This analysis is as hyperbolic as it is wrong**.

First, there is no broad "intermediate copy" doctrine of the type the MAIs describe. Indeed, of the five cases cited in support the above-quoted reimagination of the doctrine, *see id.*, four do not even discuss "intermediate" copying. *See Google LLC v. Oracle Am., Inc.*, 593 U.S. 1 (2021) (mentioned only in a parenthetical); *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015) (not mentioned); *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630 (4th Cir. 2009) (not mentioned); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003) (not mentioned). Rather, in those cases, fair use applied for different reasons, such as the overall justification for the copying to achieve some beneficial purpose or increase access to the copyrighted work, the noncompetitive nature of the "outputs" enabled by the copying, or the nature and inherent limitations in the plaintiff's copyrights:

- In *Authors Guild* and *Kelly*, copying enabled a useful research capability – a search function that is a paradigmatic fair use – and the search results (snippets of text or

thumbnail images) were so diminished that they could not serve as market substitutes for the copyrighted works they helped users find. *See Authors Guild v. Google, Inc.*, 804 F.3d at 207 (service "augments public knowledge by making available information about Plaintiffs' books without providing the public with a substantial substitute for matter protected"); *Kelly*, 336 F.3d at 821 (thumbnail images helped "guide" users to plaintiffs' work and were so small and unclear that they could not substitute for plaintiff's work).

- In *iParadigms,* 562 F.3d at 640, 644 (4th Cir. 2009), plaintiffs' works were copied for the sole purpose of detecting plagiarism, without any corresponding harm to the marketability of the copyrighted works.

- In *Oracle*, the Supreme Court held that because computer programs "always serve functional purposes," fair use plays a critical role "in determining the lawful scope of a computer program copyright." *Oracle*, 593 U.S. at 21-22. The Court contrasted those limitations to the broader protections for artistic works. *Id.* at 20. Here, Google's copying of certain lines of code from Oracles Java programming language functionally allowed programmers to use Oracle's Java in new computing environments. *Id.* at 37-38.

Even the one "intermediate copying" case the MAIs cite, *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992), invoked justifications unavailable to the MAIs. In *Sega*, the defendant reverse-engineered the computer code for plaintiff's video game console so that consumers could play defendant's games on plaintiff's device. The result was to increase the number of games compatible with plaintiff's console. *Id.* at 1522-24. Like the cases cited above, the use found to be fair enhanced legitimate access to the copyrighted work and did not otherwise jeopardize the integrity of the copyright. There are many more distinctions to be drawn, but it is simply a misreading of the caselaw to perceive a broad intermediate copying defense when each case stands on its on facts and weighs different justifications against different threats to the copyright.

More importantly, the services offered by MAIs are incalculably different in scope, operation and impact from any technology previously considered by the courts, and UMG rejects facile analogies to research tools, plagiarism trackers, or software that permits interoperability of different programs. To borrow the "no case in history" framing asserted by the MAIs, no case in history has ever exonerated an infringer who copies every single copyright owned by a rightsholder, builds a multimillion-dollar business based on that copying, and then floods the market with machine-generated substitutes for the very copyrights it pilfered to create the business. And no case in history has ever found copying to be fair

when it quite literally threatened to supplant entire categories of copyrighted works. Analogies have their obvious limits, and MAI models must be assessed in the context of the existential threat they pose to human authors and those who represent them.

Faulty MAI Premise No. 2: "Training MAI Models On Copyrighted Sound Recordings Is No Different Than Musicians Learning from Past Recorded Music"

Perhaps the most familiar defensive gambit of the MAIs is to liken their staggering copying of sound recordings to musicians "listening" to recordings to learn the basic elements and genres of music. MAIs, so the argument goes, simply absorb "influences" from past music to learn how music works and how to create something new and original, just as the artist learns from the music she references. Consider the following justifications:

- "Just as students listen to music and study scores, our model has 'listened' to and learned from a large collection of recorded music ... The goal of model training is to develop an understanding of musical ideas — the basic building blocks of musical expression that are owned by no one." *See https://www.udio.com/blog/ai-and-the-future-of-music*

- "Suno helps people create music through a similar process to one humans have used forever: by learning styles, patterns, and forms (in essence, the 'grammar' of music), and then inventing new music around them … model training looks a lot more like a kid learning to write new rock songs by listening religiously to rock music … But, just like the kid writing their own rock songs after listening to the genre — or a teacher or a journalist reviewing existing materials to draw new insights — learning is not infringing. It never has been, and it is not now." *See https://suno.com/blog/future-of-music.*

This analogy is rife with error in logic, scale, and context.

First, while composers and musicians certainly learn about music by listening to it, they do not literally copy and encode into their brains millions of copyrighted sound recordings. Even if they remember every note of a favorite recording – and no human being can – they cannot faithfully duplicate in their minds the actual recorded sounds of a sound recording and the nuances and idiosyncrasies of the performances those recordings embody. Their minds do not faithfully capture the timbres, inflections, tones, and unique qualities of John Coltrane's saxophone or Frank Sinatra's voice the way a downloaded audio file does. Moreover, the scale of copying is not even remotely comparable. A musician learns from

perhaps hundreds or thousands of recordings but cannot faithfully reproduce every recording on the internet.

Unless they are infringers, composers and performers also access their musical training materials lawfully by purchasing instructional books, downloading recordings, or subscribing to streaming services. A budding singer listening to YouTube videos is still helping the rightsholders generate advertising income, for example. MAIs, in contrast, simply scrape the internet for anything that is "publicly available," without paying the price of admission. It would be no defense to infringement for composers or recording artists who unlawfully copied recordings to insist that they stole the music only for "training" or "learning" purposes or to promise that their musical output will not infringe the recordings or music they stole in the first place. In other words, while "learning" is not infringing, making unauthorized copies is, even if one's ultimate goal is to learn something about the music embodied in those recordings. And if humans must abide by the rules of copyright law, even if their ultimate purpose is to learn something, then certainly MAIs deserve no greater solicitude.

Faulty MAI Premise No. 3: "MAIs Simply Facilitate Human Beings In Creating Music"

As we stated in our initial submission, UMG fully supports artists who use AI tools to advance their own original human authorship. *See* UMG's Initial Comments at 5-7. We in fact provided numerous examples of the salutary use of AI to extend works across different languages and connect artists and/or fans remote in place and time. *See id.* In those instances, where all rightsholders approved, AI enhances creativity without replacing it.

However, UMG also roundly rejects the myth that providing a pushbutton, instant music MAI service built on the backs of UMG's copyrighted content is justifiable and fair, because it merely "assists" or "enables" human beings to create their own original music. One MAI boasts on its website that it will "enable the next generation of music creators," *https://www.udio.com/about-us*, and insists in pleadings that it merely uses "existing sound recordings as data to mine and analyze for the purpose of identifying patterns in the sounds of various musical styles, *all to enable people to make their own new creations*..."

See Udio Answer at 1-2. Other MAIs press the same faulty argument: "Suno is a new kind of musical instrument … Suno is designed for original music, and we prize originality, both in how we build our product and in how people use it. *People who use Suno are using the product to create their own, original music.*" *https://suno.com/blog/future-of-music* (emphasis added).

This justification ignores the critical distinction between human authorship and digital processing, resorts to additional empty analogies, and falsely aggrandizes what the user of these type of MAIs do. As a general matter, users prompting MAIs to generate audio files are not composing or writing anything, much less "their own, original music." They simply supply an uncopyrightable idea in a text prompt – "electronic funk with a repeating bassline and saxophone riff" – and the software itself generates an audio track based on its own predictive algorithms. The prompting user is no more an author than someone who tells a musician friend to "write me a pretty love song in a major key" and then falsely claims co-ownership. As noted above, despite their marketing spin to the contrary, the MAIs concede as much when they advise users that "text prompts cannot fully define an actual piece of music" because each prompt can produce "an infinite number of possible audio tracks." *https://www.udio.com/guide#prompting*. The reason text prompts cannot "fully define a piece of music" is because prompts represent only an uncopyrightable "idea" for music and not expressive authorship. The fact that a single prompt might "describe an infinite number of possible audio tracks" betrays how little the user/prompter creative controls or contributes to the resulting output.

UMG stressed this very point in its first submission:

> When generative AI responds to a human text prompt, it draws upon its analysis of those statistical patterns to create something that matches or "autocompletes" the human prompt, but it is the AI system that creates the output, not a human being. Whatever the nature of the output – text, music, image – its expressive content was created and arranged by a computer making statistical predictions rather than creative choices.

UMG's Initial Comments at 12.

This reality effectively eliminates the key justifications made by the MAIs. First, it undermines the tenuous analogies used to justify infringement. An MAI is nothing like a "new kind of musical

instrument," whose sounds, notes, phrasing and overall musicality are controlled by a real human being. It is rather an autonomous digital generator of sounds based on statistical correlations, far beyond the authorial control of human being.

Second, despite their arguments to the contrary, MAIs *do not* facilitate human authorship but simply let users farm out that task to a machine, and one that is built entirely and completely on the unlawful copying of copyrighted human expression. The outputs, devoid of human authorship, are not only uncopyrightable but also ultimately unoriginal in an important sense. A human author can transcend her influences by creating new genres, new styles, new sounds, and indeed new building blocks of music. MAIs cannot. They are inherently cabined by their process of reassembling the "patterns" and sounds of prior recorded music. They are designed to simulate authorship, not promote it.

These inherent limitations deprive MAIs of the fair use justifications they proffer. If the broader purpose of copyright law as mediated by fair use is to increase the universe of human expression and let authors build on the work of prior authors, MAIs accomplish none of those goals. They are simply pushbutton services that let any user type a few words to generate an output that "sounds like," but is not human expression. In the manner in which they are currently used, MAIs of this nature do not enable copyrighted expression but simply replace it.

Faulty MAI Premise No. 4: "The Recording Industry Is Trying to Monopolize Style"

Another MAI misdirection is the claim that the recording industry's objection to training on sound recordings is an effort to monopolize musical styles. *See, e.g.,* Udio Answer ("The premise of [the record labels'] case is that musical styles—the characteristic sounds of opera, or jazz, or rap music—are somehow proprietary"); Answer of Defendant Suno, Inc. To Complaint, *UMG Recordings, Inc. v. Suno, Inc.*, Civil Action No. 1:24-cv-11611-FDS, United States District Court For District Of Massachusetts ("No one owns musical styles"). In this telling, objecting to wholesale copying in the training process is an effort to exert exclusive control over genres and styles.

This concern is a wholly manufactured *non sequitur*. Indeed, in our initial comments, we stated quite explicitly that copyright law does not protect "style." *See* UMG's Initial Comments at 95. The legal actions commenced against these MAIs currently focus on their infringement regarding inputs – copying UMG's sound recordings (including the compositions they embody) during the training process. The challenged use is the copying of copyrighted sound recordings to build multimillion dollar businesses. And, as noted above, that copying is not some incidental feature of those businesses. Rather, the MAIs have predicated their entire business model – one that garners breathtaking investment – on making exact reproductions of legitimate sound recordings in the course of training their models. The relevance of the outputs is not what "styles" they feature, but their proven capacity to serve as competitive substitutes for real music. MAIs cannot point to a single case that ever deemed copying of this scale to enable directly substitutable products to be fair use. The charge of monopolizing styles is yet another red herring.

The MAIs indulge in a related diversion when they stress the unique nature of sound recording copyrights as a justification for their copying at the training stage. They argue that "outputs … which do not reprise 'the actual sounds fixed' in any 'recording' owned by any record label, are not and cannot be even *prima facie* copyright infringements." But as stated above, whether or not outputs incorporate actual sounds of sound recordings – and there is strong evidence that they do, *see* Comments Of The American Association Of Independent Music And Recording Industry Association Of America, Inc. at 4-5 – there is no dispute that MAIs copy those actual sounds to train their models in the first place. There is nothing "*per se*" lawful about copying entire catalogs of music to create a commercial enterprise that threatens to displace those catalogs.

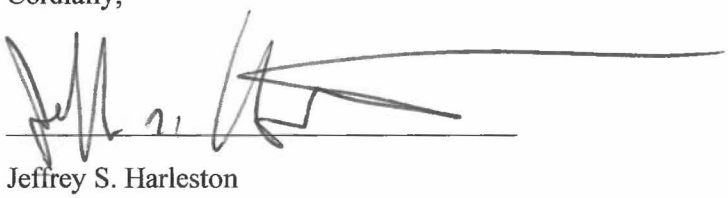
And, as UMG has previously pointed out, the "new sounds" are not copyrightable content, but computer-generated sounds assembled by predictive algorithms. Rather than creating incentives for human beings to produce new copyrighted works, these outputs undermine those incentives. Why create a new composition or sound recording when you know it will be used to generate synthetic substitutes that compete with and potentially displace your original authorship?

III. CONCLUSION

AI holds both great promise and threatens serious harm to the creative communities. That promise can be harnessed, and those threats managed by an appropriate system of licensing and cross-industry collaboration of the type that has solved past technological challenges. But within the spectrum of benefits and harms occasioned by AI, MAIs pose a uniquely immediate and devastating threat to musicians, composers, and the music industry. There is no precedent for a service and technology that copies *all of a copyright owner's copyrighted* content for the exclusive purpose of generating competitive digital content that drowns out the very human authorship that made the service possible in the first place. Neither legal precedent, public policy, equity, nor common sense can justify such an unfair and unwise outcome.

We thank the Copyright Office for considering our viewpoint.

Cordially,



Jeffrey S. Harleston