



May 21, 2021

**VIA EMAIL ONLY**

Regan Smith  
General Counsel and Associate Register of  
Copyrights  
U.S. Copyright Office, Library of Congress  
101 Independence Ave, SE  
Washington, DC 20559-6000

**Re: Response to Post-Hearing Letter on Proposed Class 7 – Docket No. 2020–11,  
Exemptions to Prohibition Against Circumvention of Technological Measures  
Protecting Copyrighted Works**

Dear Ms. Smith:

On behalf of the Joint Creators and Copyright Owners (the “JCCO”), I respectfully submit this response to your letter of April 16, 2021. This letter responds to your specific questions, and to issues raised by petitioners’ reply comments.

First, we explain why we continue to oppose the issuance of any exemption for proposed Class 7. Second, with a focus on motion pictures, we critique the revised proposed regulatory language proffered in petitioners’ reply comments. Third, we respond to your questions regarding security measures. Fourth, we respond to your specific question concerning how to craft potential “regulatory language that would specify that researchers would not be permitted to view the text or images from works included in a corpus.” With respect to the latter two questions, we emphasize that we are responding only to address your requests, and not to endorse the notion that the proposed exemption for text and data mining (“TDM”) should be recommended or granted. The Copyright Office should deny proposed Class 7 for the reasons discussed in our prior opposition comments, during the hearing, and in this letter.

**1. The Exemption Should be Denied.**

Petitioners argue, at bottom, that the *HathiTrust* and *Google Books* cases,<sup>1</sup> along with law review articles by academics pre-disposed not to side with copyright owners, render the law settled that all text and data mining involving reproducing and using copyrighted works is lawful.<sup>2</sup> We

---

<sup>1</sup> *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014); *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015).

<sup>2</sup> The other cases on which petitioners rely are inapposite. For example, *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630 (4th Cir. 2009), involved the use of unpublished, non-commercial student papers in a searchable database to prevent plagiarism and cheating. *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir.

continue to disagree.<sup>3</sup> Moreover, multiple U.S. government agency studies conducted over the course of the years have not so stated.<sup>4</sup> Instead, it is apparent that sometimes text and data mining *might* be fair use, and sometimes it is not. This proceeding is not the proper venue for the Copyright Office to establish a pseudo-precedent that the courts, the legislative branch, and the executive branch have never announced. The issue is too important for the Register and the Librarian to choose to go out on a limb that no agency has heretofore chosen to risk testing given the weight of the issue.

TDM projects, especially those involving unauthorized circumvention of access controls on digital copies of works, often require extensive copying of large numbers of copyrighted works. Such conduct both challenges the Copyright Act's provision of exclusive rights to authors and generates large databases of unauthorized, unprotected, digital reproductions of valuable copyrighted works. While these projects may prove to be valuable for nonprofit research purposes, there is also a potentially valuable market for copyright owners to enter if they elect to license their works or to engage in TDM themselves. The U.S. government continues to consider how to balance the need for the U.S. to (1) protect intellectual property and (2) compete in areas such as artificial intelligence and other forms of machine learning.<sup>5</sup> In this proceeding, the Copyright Office should not conclude that copyright owners have no role in authorizing and designing TDM projects dependent on the content they have invested to create and disseminate.

## 2. Petitioners' Revised Proposed Language Remains Problematic.

In their reply comments, petitioners significantly narrowed their proposed regulatory language. We appreciate these efforts, and the willingness of petitioners' counsel to discuss our concerns before and after filing their reply comments. However, the proposed regulatory language continues to be imprecise, includes material flaws that put the rights of copyright owners at risk, and requires the Register and the Librarian to focus careful attention on the details of the proposal.

---

2007), and *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003), involved copying and display of copyrighted images by internet search engines where they were not instructed by website owners not to crawl their websites. These cases bear no resemblance to the uses of motion pictures – including blockbusters – envisioned by petitioners.

<sup>3</sup> See, e.g., *Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169 (2d Cir. 2018); see also *Elsevier Inc. v. www.Sci-Hub.org*, No. 15-CV-4282 (RWS), 2015 WL 6657363 (S.D.N.Y. Oct. 30, 2015); *Elsevier Inc. v. Sci-Hub*, No. 15-CV-4282 (RWS), 2017 WL 3868800 (S.D.N.Y. June 21, 2017).

<sup>4</sup> See, e.g., U.S. PATENT AND TRADEMARK OFFICE, PUBLIC VIEWS ON ARTIFICIAL INTELLIGENCE AND INTELLECTUAL PROPERTY POLICY 25 (2020); U.S. COPYRIGHT OFFICE, LEGAL ISSUES IN MASS DIGITIZATION: A PRELIMINARY ANALYSIS AND DISCUSSION DOCUMENT 30 (2011); U.S. COPYRIGHT OFFICE, ORPHAN WORKS AND MASS DIGITIZATION 76 (2015).

<sup>5</sup> See, e.g., NATIONAL SECURITY COMMISSION ON ARTIFICIAL INTELLIGENCE, FINAL REPORT, <https://www.nscai.gov/2021-final-report/> (last visited May 21, 2021). The NSCAI report calls for the Commerce Department, in consultation with the Copyright Office, to develop proposals to address the relationship between intellectual property law and machine learning.

The revised regulatory language proffered by petitioners in their reply comments is as follows:<sup>6</sup>

*Proposed Class 7(a) (revised): Motion pictures, where the motion picture is lawfully made and obtained on a DVD protected by the Content Scramble System, on a Blu-ray disc protected by the Advanced Access Content System, or via a digital transmission protected by a technological protection measure, where: (1) the circumvention is undertaken by a researcher affiliated with a nonprofit library, archive, museum, or institution of higher education to deploy text and data mining techniques for the purpose of scholarly research and teaching; and (2) the researcher uses reasonable security measures to limit access to the corpus of circumvented works only to other researchers affiliated with qualifying institutions for purposes of collaboration or the replication and verification of research findings.*

*Proposed Class 7(b) (revised): Literary works, excluding computer programs, distributed electronically and lawfully obtained, that are protected by technological measures that interfere with text and data mining, where: (1) the circumvention is undertaken by a researcher affiliated with a nonprofit library, archive, museum, or institution of higher education to deploy text and data mining techniques for the purpose of scholarly research and teaching; and (2) the researcher uses reasonable security measures to limit access to the corpus of circumvented works only to other researchers affiliated with qualifying institutions for purposes of collaboration or the replication and verification of research findings.*

The flaws in this language are numerous. The following, non-exhaustive list identifies the most important problems this language poses with regards to motion pictures:

- Petitioners remain unable, and unwilling, to define “text and datamining” in a comprehensible manner. Although petitioners claim they need flexibility, without a definition, any regulation that allows circumvention of access controls on copies of motion pictures for “computational analysis” will be a minefield of uncertainty.
- The proposed language limits the copies of motion pictures that may be circumvented to those “lawfully made and obtained.” As discussed during the hearing, the issue to address is not whether a *motion picture* was lawfully made, but rather whether the *copy* used to reproduce another copy to be ingested into a TDM database was lawfully made. Petitioners disclaim any intent to cover copies of motion pictures obtained from illicit databases such as Sci-Hub, cyber lockers or P2P networks. Thus, any regulatory language must state that the ingested copy must be created from a lawfully made copy.<sup>7</sup> In addition, whether a copy is “lawfully obtained” should turn on whether the institution engaged in circumvention purchased a copy. We are pleased that petitioners, in their reply comments and during the hearing, excluded any copies of motion pictures created from subscription streams/services and time-limited rentals from the scope of the proposal. However, we still have a disagreement with petitioners with respect to including downloaded copies of motion pictures that allow for long-term access within

---

<sup>6</sup> Petitioners appeared to further narrow their proposals during the hearing. Unfortunately, we do not have access to a transcript of that hearing.

<sup>7</sup> Moreover, any institution attempting to qualify as a beneficiary of the exemption must not intermix reproductions of lawfully made copies of motion pictures with reproductions of unlawfully made copies of motion pictures (which may not require circumvention to obtain because they are disseminated illegally without TPMs).

the category of “lawfully obtained” copies. Such downloads are provided based expressly on terms of service that disclose that the copies are licensed, not purchased, and that if a service ceases to exist the copies may no longer be accessible. Thus, the language should limit the exemption to copies obtained on physical discs (*i.e.*, copies lawfully distributed).

- The proposed language covering literary works states that an access control must “interfere” with the text and datamining. The language proposed for motion pictures does not. This must be corrected. Any circumvention undertaken must be necessary to accomplish the purposes of the research. Where alternatives exist, no circumvention should be allowed.<sup>8</sup> Also, licensing should be considered/explored as an alternative. If institutions forego the opportunity to request permission from rights holders (especially in instances where large catalogues of rights are held by single entities), there is no demonstration of the absence of alternatives to circumvention.
- The proposed language states that “the circumvention is [to be] undertaken by a researcher” but fails to define “researcher” or to state where or how the circumvention will be conducted. If an exemption is granted with respect to this class, it should make clear that the circumvention and reproduction should be conducted in specified, on-campus settings that enable protective security measures to be employed. Any exemption should not allow thousands of “researchers” to circumvent access controls on their home computers to create databases of works and to then transmit them or otherwise leave them exposed to misuse. In addition, the comments, reply comments, and hearing testimony all focus almost exclusively on academic researchers in non-profit university settings. Any exemption should be so limited (see more below).
- The proposed language allows for circumvention by a “researcher affiliated with a nonprofit library, archive, museum, or institution of higher education.” Our view is that “affiliated” is far too vague a term. Almost anyone could be “affiliated” with a library or school. Any exemption should be limited to faculty, necessary staff, and registered graduate students at accredited, nonprofit research universities. There is no evidence to support including libraries, archives and museums in the exemption.<sup>9</sup> And, as conceded by petitioners, no commercial enterprises, or affiliated individuals, should be using or have any access to the resulting databases.
- The proposed language should include a requirement that the circumvention be for the “sole purpose of deploying text and data mining techniques for the sole purpose of scholarly research and teaching.” It should also be made clear that petitioners’ promises that “substitutional uses of expressive content,” “training models to generate output that

---

<sup>8</sup> In our view, petitioners have not demonstrated that screen capture technology is an unviable alternative to circumvention.

<sup>9</sup> Our understanding is that petitioners’ intend to include such institutions as they are defined in Section 108. That is too broad a scope in this context.

mimics or even infringes upon works in the corpus, or unspecified commercial enterprises” are all prohibited.

- The proposed language requires that “the researcher uses reasonable security measures to limit access to the corpus of circumvented works ...”. As discussed below, the security methods must be much more specific. Also, it should not be left up to the researcher to deploy them. The academic institution should be responsible for implementing them. Finally, consistent with the E.U. Directive, copyright owners should be in control concerning best practices for securing databases of reproduced motion pictures.
- “Limit[ing] access” is not enough for sufficient security. All distribution of full or partial copies should be prohibited (see more below). Works in a database should not be in downloadable form and copies should be deleted at the conclusion of a project. Moreover, an express requirement that no infringement may be facilitated is necessary.
- The proposed language states that access must be limited “only to other researchers affiliated with qualifying institutions for purposes of collaboration or the replication and verification of research findings.” As discussed above, “other researchers affiliated with qualifying institutions” should be limited to faculty, necessary staff and registered graduate students at accredited, nonprofit research universities. In addition, “collaboration” is a vague term that could render the exemption unclear. It is our understanding that petitioners are agreeable to limiting the scope of the exemption such that each institution must create its own database of motion pictures, specific to each research project, using copies that institution itself lawfully acquired. They disclaimed a desire for institutions to share databases or to enable linking of multiple databases. If an exemption is granted that allows for collaboration, the term must be defined narrowly. Moreover, there is no explanation as to how the collaborators, or persons engaged in peer review, would be provided access to search results. This importantly relates back to the security measures issue, discussed below.

### **3. Exemption Beneficiaries Should Be Required to Follow Best Practices Developed by Copyright Owners and/or Specific Security Measures Should be Required.**

Due to the scope of copying implicated by the proposed exemption, its potential beneficiaries should be required to follow reasonable security guidelines articulated by copyright owners. Article 3 of the European Directive on the Digital Single Market acknowledges the need for such protection for copyright owners.<sup>10</sup> While the Directive allows for national implementation of laws allowing certain text and data mining activities, it also provides: “Rightholders shall be

---

<sup>10</sup> The Directive does not expressly include motion pictures as a category of works that should be subject to text and data mining. See Directive 2019/790, of the European Parliament and of the Council of April 17 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC, Preamble, Para. 8, 2019 O.J. (L 130/92) (hereinafter “Directive”). The Directive overall is also not a good fit for the U.S. market in multiple respects, although the requirement that copyright owners are entitled to issue guidelines to protect their content is a helpful, commonsense benchmark.

allowed to apply measures to ensure the security and integrity of the networks and databases where the works or other subject matter are hosted.”<sup>11</sup> Given the novelty of petitioners’ proposal, such guidelines have not yet been developed by the motion picture industry. Examples of existing security standards discussed during the hearing on this proposed class include those promulgated by the International Organization for Standardization, the National Institute of Standards and Technology, and the University of California at Berkeley. Moreover, MPA has publicly available guidelines for the physical and virtual protection of content.<sup>12</sup> All of these standards are good starting points. However, if an exemption is recommended and granted, the motion picture industry must be given time to articulate a set of guidelines specific to the scope of the TDM projects sanctioned by the exemption.

#### **4. Based on Petitioners’ Reply and Testimony, Content in any TDM Database Should be Inaccessible.**

Given that we oppose the proposed exemption, and that our list of concerns above regarding the petitioners’ proposed drafting should result, at least, in the substantial revision of such proposal, it is difficult to draft language focusing only on this issue without knowing the full context of how an exemption would be crafted. Nevertheless, if an exemption is recommended and/or granted, it should state: “No content from the works to which unauthorized access was gained under this exemption may be viewed or heard by any person who is otherwise allowed to conduct research under this exemption.” To be clear, only faculty, necessary staff and registered graduate students at accredited, nonprofit research universities should have access even to the search results, so this limiting language focuses only on them and does not imply anyone else should have access to the content.

#### **5. Conclusion**

In sum, we remain opposed to proposed class 7. We also have significant concerns regarding the petitioners’ proposed language. Any exemption should require compliance with reasonable security guidelines published by impacted copyright owners, including by the motion picture industry. Finally, content in TDM databases should not be accessible for viewing or listening.

Respectfully,  
/s/J. Matthew Williams  
Partner of  
MITCHELL SILBERBERG & KNUPP LLP  
Cc: Jordana Rubel, Assistant General Counsel

---

<sup>11</sup> Directive, Article 3(3). *See also id.*, Article 3(4) (“Member States shall encourage rightholders, research organisations and cultural heritage institutions to define commonly agreed best practices concerning the application of the obligation and of the measures referred to in paragraphs 2 and 3 respectively.”). Security guidelines for TDM practices have not been issued for motion pictures in Europe and no best practices have been negotiated there for motion pictures.

<sup>12</sup> *See, e.g.*, Motion Picture Association, *Content Security Best Practices Common Guidelines* (Nov. 11, 2020), <https://www.motionpictures.org/wp-content/uploads/2020/11/MPA-Best-Practices-Common-Guidelines-V4.08-FINAL.pdf> (last visited May 20, 2020).