

Legal Reference Services Quarterly

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/wlrs20>

Perma.cc and Web Archival Dissonance with Copyright Law

Paul Douglas Callister

To cite this article: Paul Douglas Callister (2021) Perma.cc and Web Archival Dissonance with Copyright Law, Legal Reference Services Quarterly, 40:1, 1-57, DOI: [10.1080/0270319X.2021.1886785](https://doi.org/10.1080/0270319X.2021.1886785)

To link to this article: <https://doi.org/10.1080/0270319X.2021.1886785>



Published online: 24 Apr 2021.



Submit your article to this journal



Article views: 159



View related articles



View Crossmark data



Perma.cc and Web Archival Dissonance with Copyright Law

Paul Douglas Callister

Leon E. Bloch Law Library, University of Missouri–Kansas City School of Law,
Kansas City, Missouri, USA

ABSTRACT

Harvard's Perma.cc offers the solution to link rot—the phenomenon that citations in academic journals to Web materials disappear with the passage of time, resulting in "broken links" and disappearance of material from the Web. This article will describe Perma.cc and outline the kinds of copyright issues that may arise, including heavy use of copyright statutes and case law. It will examine the kind of preservation use of copyrighted materials, with reference to fair use, and the library prerogatives as exceptions to the exclusive rights of authors of materials found on the Web. This analysis includes detailed analysis of "transformative use" and the four factors of 17 U.S.C. § 107. It will consider the liability of Perma.cc and participating libraries and institutions under theories of contributory infringement and vicarious liability, including as modified by 17 U.S.C. § 512(c) and (d), governing takedown notices. The article concludes that Perma.cc's archival use is neither firmly grounded in existing fair use nor library exemptions; that Perma.cc, its "registrar" library, institutional affiliates, and its contributors have some (at least theoretical) exposure to risk; and that current copyright doctrines and law do not adequately address Web archival storage for scholarly purposes. In doing so, it will question what the role of the scholarly Perma.cc citation ought to play—confirmation of scholarly propositions or preservation of and access to Web materials. The material and conclusions in this article are important for legal authors, law review editors, and librarians (especially those who use, support, or are considering partnering with Perma.cc) so that they might better assess copyright compliance, especially when selecting materials for archiving, such as articles from news sites, blogs, and professional and scholarly papers, articles, or books.

KEYWORDS

copyright, Perma.cc, fair use, transformative use, contributory infringement, vicarious liability, web archiving, library exceptions, broken links, link rot

Introduction: Perma.cc, link rot, archiving, and copyright

Perma.cc offers the solution to link rot—the phenomenon that citations in academic journals to Web materials disappear with the passage of time, resulting in “broken links.” According to Perma.cc’s own Web site, 70% of all links published between 1999 and 2011 are broken.¹ Perma.cc solves this problem by allowing scholars to load Web materials cited in their papers to a Web archive and by providing a permanent URL known as a “Perma.cc link.”² As an example, Perma.cc preserves the FBI’s *Top Ten Most Wanted List* at a given moment in time. A scholar may want to archive the site as of a given date (May 1, 2015) and can do so with the Perma.cc link to <https://perma.cc/T8U2-994F>.³ The Web page stored is interactive,⁴ although a screenshot version is also available. Besides Web pages, encoded in HTML, there is no reason why other types of file types such as PDFs, MS Word documents, spreadsheets, PowerPoints, and video files cannot be stored if they were linked to the Web.⁵

Perma.cc is not just meant for any kind of archiving. Its fundamental use is indefinite preservation of an otherwise perishable resource for scholarly purposes. Although the risk of a suit, or even a takedown notice, appears to be low, copyright law is fundamentally dissonant with archival practices. Regardless of risk, there are instances when copyright should be considered prior to archiving. Especially concerning might be blogs, articles on news’ sites, scholarly and professional papers, and articles and books in PDF form that do not grant permission for copying.

This article will describe Perma.cc and outline the kinds of copyright issues that may arise, including heavy use of copyright statutes and case law. It will examine the kind of preservation use of copyrighted materials, with reference to fair use, and the library prerogatives as exceptions to the exclusive rights of authors of materials found on the Web. This analysis

¹Perma.cc, *About Perma.cc*, <https://perma.cc/docs#what-is-link-rot> (last visited May 22, 2020).

²*Id.* at <https://perma.cc/docs#how-does-perma-work>.

³Described at Perma.cc, *Perma.cc User Guide/Getting Started*, <https://perma.cc/docs#how-does-perma-work> (last visited May 22, 2020).

⁴The “capture” format uses Web ARChive (or “warc”). See Wikipedia, *Web ARChive*, https://en.wikipedia.org/wiki/Web_ARChive (last visited May 22, 2022). It is widely used in Web harvesting and “Web crawls.” *Id.* The screenshot version is a PNG file. See Perma.cc, *Perma.cc User Guide Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#preservation-formats> (last visited May 22, 2020).

⁵An email from Adam Ziegler of Harvard’s Library Innovation Lab suggests different file types can be archived, although video can be tricky.

When a user directs Perma to try to archive something that’s publicly available on the web, our software goes to the given URL and tries to archive what’s there whether it’s simple HTML, a hosted PDF, DOCx, video, etc. For example, many users want to preserve government reports that are available on the Web as PDFs. Likewise with other types of materials online. We try to support that to the extent the technology allows (video preservation, for one, is often tricky).

Email from Adam Ziegler, Director, Library Innovation Lab, Harvard Law School Library, to author (Apr. 17, 2020, 12:58 pm CST) (on file with author).

includes detailed analysis of “transformative use” and the four factors of 17 U.S.C. § 107. It will consider the liability of Perma.cc and participating libraries and institutions under theories of contributory infringement and vicarious liability, including as modified by 17 U.S.C. § 512(c), governing takedown notices. It observes that frustration with the notice-and-takedown scheme has resulted in a recent, popular legislation, which was passed as part of the Appropriations Bill at the end of 2020 providing relief for COVID-19.⁶ This legislation provides for an alternative to notice-and-takedown notices with a Copyright Small Claims Board. Libraries and archives can preempt jurisdiction by getting on a list with the Register of Copyrights. This could alter the risk analysis for archives relying on takedown notices.

The article concludes that Perma.cc’s archival use is neither firmly grounded in existing fair use nor library exemptions; that Perma.cc, its “registrar” library and institutional affiliates, and its contributors (see [Figure 1](#)) have some (at least theoretical) exposure to risk; and that current copyright doctrines and law do not adequately address Web archival storage for scholarly purposes. In doing so, it will question what the role of the scholarly Perma.cc citation ought to play—confirmation of scholarly propositions or preservation of and access to Web materials. The material and conclusions in this article are important for legal authors, law review editors, and librarians (especially those who use, support, or are considering partnering with Perma.cc) so that they might better assess copyright compliance, especially when selecting materials for archiving, such as articles from news sites, blogs, and professional and scholarly papers, articles, or books.⁷

Perma.cc and terms of service

To understand the obligations of the parties under Perma.cc’s terms of service, it is necessary to understand that there are three parties involved in most transactions—submitting scholars, registrar libraries or organizations,⁸ and Perma.cc. All of them share some risk for infringing contributions to Perma.cc’s archive. Without the scholar’s home libraries affiliating themselves with Perma.cc by becoming registrars, it is only possible to get an individual membership with 10 free links and thereafter a mandatory

⁶See Consolidated Appropriations Act, 2021, H.R. 133, 116th Cong. (2020) (enacted). Title II, Subtitle A, *Chapter 15—Copyright Small Claims*—of the Act, including §§ 1501–1511, covers the creation, claims, and procedures for a Copyright Claims Board. Libraries may “preempt out” of the Small Claims Board proceedings. See *id.* at Title II, Chapter 15, § 1506(aa).

⁷In reviewing an earlier draft of this article, Perma.cc believed the copyright risks to be extremely low and was in fundamental disagreement with the suggestion of risk in an earlier draft of the article (and probably this one too). Copy of review is on file with author. While now deemphasizing risk, I have proceeded with this article to highlight the dissonance of copyright law and online archiving of Web materials.

⁸Registrars may add another party, organizations such as law reviews or research groups. See Perma.cc, *Perma.cc User Guide/For Registrars*, <https://perma.cc/docs/libraries> (lasted visited June 11, 2020).

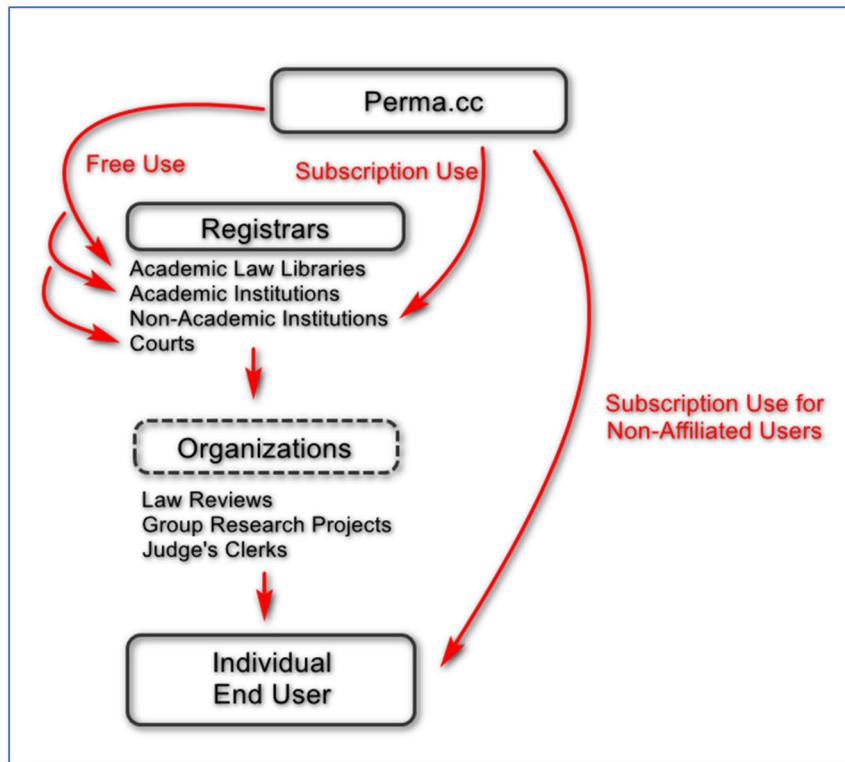


Figure 1. Chart depicting types of membership and subscriptions.

subscription.⁹ Nonacademic institutions can become registrars, but they also have subscription fees.¹⁰

User and library obligations and indemnification

Both scholarly contributors (users) and library (and other institutional) registrars agree to respect the copyright of others and to indemnify Perma.cc for liability.¹¹ Library and other registrars must see that their users “do not abuse their accounts.”¹² Section 2(a) of the terms of services (applying to both users and library registrars) reads: “You may use the Site

⁹See Perma.cc, *About Perma.cc*, <https://perma.cc/about#accounts> (last visited May 22, 2020). Subscription prices are tiered—to links for \$10.00 a month, 100 links for \$25.00 a month, and 500 links for \$100.00 a month. See Perma.cc, *Settings*, <https://perma.cc/settings/subscription> (last visited June 1, 2020).

¹⁰See Perma.cc, *About Perma.cc*, <https://perma.cc/about#accounts> (last visited May 22, 2020).

¹¹Registrars play a role in administering Perma.cc services with the registrar's "network" of users. They are the main point of contact for questions about Perma.cc. They are responsible for ensuring that network users "do not abuse their accounts." Perma.cc, *Perma.cc User Guide/For Registrars*, <https://perma.cc/docs/libraries> (last visited May 6, 2020). Registers are also bound by Perma.cc's Terms of Service. See Perma.cc, *Perma.cc for Libraries*, <https://perma.cc/libraries> (last visited May 6, 2020) ("By registering [as a library registrar], you agree to the terms of service.").

¹²See Perma.cc, *Perma.cc User Guide/For Registrars*, <https://perma.cc/docs/libraries#libraries-authority> (last visited May 22, 2020).

and Service, including content stored at the direction of users, only for non-commercial scholarly, research, reporting, criticism and commenting purposes that do not infringe or violate anyone's copyright or other rights.”¹³ Furthermore, the users make contributions to Perma.cc at their own risk.¹⁴ Users and registrar libraries warrant that “Submitted Content both (1) is freely available on the Internet to the general public … and (2) is cited in a legal work or in a work of scholarship, reporting, criticism or commentary.”¹⁵ The “Submitted Content” must be lawfully acquired, users must have all the necessary rights to store the content, and the content must not “infringe or violate copyrights.”¹⁶ The users and registrars even grant a nonexclusive sublicense to Perma.cc to

use, reproduce, create derivative works based upon, transmit, distribute, perform, display, and make available the User Submitted Content,¹⁷ in any medium now in existence or later developed, in connection with the Service or otherwise in furtherance of our mission, including but not limited to promotional uses.¹⁸

Finally, users and library registrars agree to indemnify Perma.cc:

You agree to indemnify and hold harmless us, our affiliates, governing board members, officers, employees, agents and representatives, and any party with whom we may contract to provide the Service, from and against any claims, liabilities, losses, damages, costs and expenses, including but not limited to reasonable attorneys' fees and court costs, arising out of or in any way connected to your use of the Site or Service, including but not limited to any allegation or claim that, if true, reflects your violation of these Terms of Service or the infringement or violation by you (or occurring through use of your account) of *any intellectual property or other right* of any person or entity.¹⁹

Every item of *submitted content* from the end user is a risk (perhaps slight) for him or her, his or her host registrar institution, and Perma.cc. If there were no risk, why would Perma.cc bother with an indemnification clause?

Governmental library registrars and immunity from suit

Legal counsel from my own host institution (a state university) refuses to agree to let my law library become a registrar because the indemnification clause does not provide that state-sponsored institutions may be free from law suit

¹³Perma.cc, *Perma Terms of Service*, effective Jan. 22, 2019, <https://perma.cc/terms-of-service>.

¹⁴*Id.* at § 4.

¹⁵*Id.* at § 5(a)(i). In spite of the “freely available” requirement, paradoxically, pages with a “noarchive” metatag are still stored by Perma.cc, but by default, in a folder only accessible to the user and his or her organization. See Perma.cc, *Perma.cc User Guide: Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#private-records> (last visited May 19, 2020).

¹⁶Perma.cc, *Perma Terms of Service*, *supra* n. 15, at § 5(a)(ii)–(iv).

¹⁷“User Submitted Content” is defined in *id.* at § 5(a). It is basically what the user directs Perma.cc to store.

¹⁸*Id.* at § 5(c).

¹⁹*Id.* at § 8 (emphasis added).

under the Amendment XI of the Constitution (and should be free from indemnifications for suits exempted by the Constitution). Current law on the immunity of states, and their respective agencies, from copyright infringement is a constantly evolving quagmire that is best left to university counsel.²⁰ However, by accepting the indemnification clause as is, a state school library registrar may effectively waive its Amendment XI protections from liability, at least with respect to indemnifying Perma.cc. Thus, without going any further into copyright risks, state school libraries should hesitate to enter into Perma.cc registrar agreements, at least under their terms for indemnification as of this writing.

Illustration and nature of the problem

Some may wonder if this is an article in search of a problem. It is not. Much Web archival activity is dissonant with current law, although there is almost no litigation over the issue, and the common remedy is a takedown request. The problem is that when library and institutional registrars use Perma.cc, they, and the people who use their Perma.cc links, are violating copyright by publishing content without the copyright owners consent. This problem is compounded by lack of clear copyright defenses for digital archival uses.

To illustrate concern, without my asking, the editors at *Law Library Journal* added 50 Perma.cc links to Web materials cited in my recent article, “Law, Artificial Intelligence, and Natural Language Processing: A Funny Thing Happened on the Way to My Search Results.” Most of the articles I cited were already archived elsewhere.²¹ Well-meaning *Law Library Journal* editors uploaded all of my citations to Web materials to Perma.cc, regardless of licensing, and placed corresponding links in my article. The issue is care and thoughtfulness with respect to copyright in using the Perma.cc archive. Ultimately, I chose to have the links removed.

For instance, in the article,²² I cited to “DANIEL JURAFSKY & JAMES H. MARTIN, *Vector Semantics*, in SPEECH AND LANGUAGE PROCESSING 11–12 (3d ed., draft Sept. 23, 2018), <https://web.stanford.edu/~jurafsky/slp3/6.pdf>.²³” This online book chapter is revised from earlier published print editions of the same book. Currently, the online draft of the third edition is available for viewing on a Stanford Web site.²⁴ It is possible that the chapter will remain indefinitely available on the Web, or the author might place the

²⁰See David Nimmer, 3 Nimmer on Copyright § 12.01 (Matthew Bender 2019). “One of the few constants, however, in the ‘complex, often contradictory, and in some cases baffling’ Supreme Court pronouncements about the Eleventh Amendment is that it is seldom, if ever, taken literally.” See *id.* at [E][2][b].

²¹Paul D. Callister, *Law, Artificial Intelligence, and Natural Language Processing: A Funny Thing Happened on the Way to My Search Results*, 112 Law Libr. J. 161 (2020), <https://doi.org/10.31228/osf.io/dw29y>.

²²See *id.* at 174, n. 58.

²³See Dan Jurafsky & James H. Martin, *Speech and Language Processing* (3d ed. draft), <https://web.stanford.edu/~jurafsky/slp3/> (last visited Jan. 29, 2021).

new edition with a print publisher and, either by choice or instruction of the publisher, remove draft versions from the Web. Perhaps an ebook will even be available, but behind a paywall.²⁴

Naturally, my desire would be to use a Perma.cc link to preserve access to the all of my citations to Web materials, but I would wager that if I archive copyrighted content that is not an ordinary Web site (such as Weather.com), there is a much higher risk that an owner of whatever book, article, or other content I preserved may seek redress for infringing on copyright or at least demand that materials be taken down from the archive. In particular, news articles, blogs, and professional and scholarly publications may trigger objections of their owners.²⁵ In my same article, besides Web sites I cite to many articles, papers, and even videos that are available for free on the Internet. Admittedly, some of the articles are already in scholarly archives, and placement in Perma.cc would be redundant. Some articles may grant permission for copying. However, if they are all unquestioningly preserved by and linked to Perma.cc, there is heightened risk of objection by owners. Yet at the same time, I want my citations to be available in perpetuity and to avoid embarrassing and annoying broken links. It's a catch-22: You can either follow copyright rules or risk dead links in your articles, but you can't do both.

In another instance, in a June of 2019 presentation, I referred to a free ABA Web site for a compilation of ABA Ethics Opinions on attorney use of the “cloud,” but in June of 2020 that Web link is broken and is only available behind a paywall (see Figure 2).²⁶ If I had used Perma.cc, I could have preserved the material, but I would have violated copyright by doing so.

Another troubling scenario is when people use Perma.cc to archive news blogs or services that have their own subscription archives but allow free access for a monthly “peek” at a defined number of articles.²⁷ The copyright holders only intend the material to be available for free for a limited time. Other news sites may make current headlines available but archive

²⁴“A paywall is a method of restricting access to content via a purchase or paid subscription.” Wikipedia, *Paywall*, <https://en.wikipedia.org/wiki/Paywall> (last visited May 22, 2020).

²⁵Certainly, many authors license their materials for reuse using licensing such as the Creative Commons provides. See Creative Commons, *Share Your Work*, <https://creativecommons.org/share-your-work/> (last visited Oct. 5, 2020). These may be archived according to the term set by the Creative Commons license.

²⁶The broken link is to http://www.americanbar.org/groups/departments_offices/legal_technology_resources/resources/charts_fyis/cloud-ethics-chart.html. The material is still available but behind a paywall. See ABA, *Final, Cloud-Based Ethics Opinions 2019*, https://www.americanbar.org/digital-asset-abstract.html/content/dam/aba/images/legal_technology_resources/CloudEthicsOpinions2019/cloudethicsopinions2019.pdf (last visited June 10, 2020).

²⁷For example, online news sites with free views of a specified articles each month. The objective of these “teasers” is to induce viewers to sign up for subscriptions. *The New York Times*, *Washington Post*, and *Wired* all do this. Theoretically, it should be possible for such pages from such sites to be archived in Perma.cc if the number of “free” views by Perma.cc had not been exceeded.

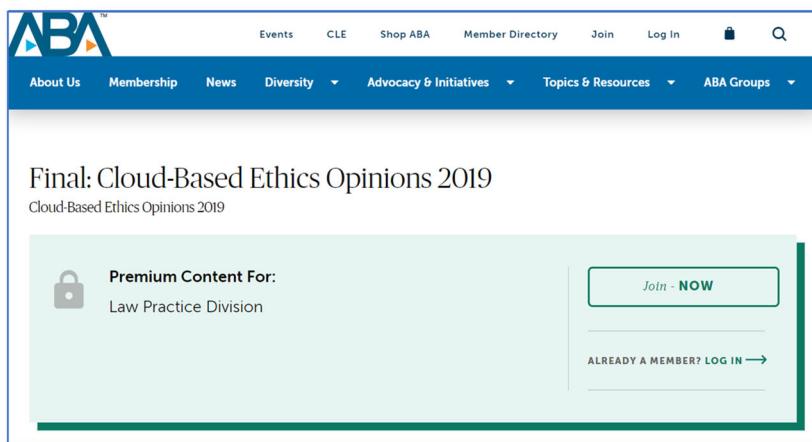


Figure 2. Illustration of ABA paywall.

older articles, which can only be accessed with a subscription.²⁸ Thus, news articles are eventually moved behind paywalls. Users might successfully submit such news articles to Perma.cc, not realizing they have bypassed the access protocols put in place by news organizations.²⁹ Even if the submitted site page has a “noarchive” metatag, Perma.cc will still archive the page, but by default the page “is accessible only to the individual account, organization and registrar responsible for the Perma Record.”³⁰ That default can be changed, however, by users and their organizations.³¹

What form will redress for copyright infringement take? Most likely copyright owners will pursue the notice-and-takedown remedy offered as part of 17 U.S.C. § 512(c).³² But, although seldom used,³³ copyright owners may seek other remedies against Perma.cc and perhaps registrars under § 512(c) based upon theories of contributory infringement and vicarious liability.³⁴ They may also sue end users for infringement under § 501 and the

²⁸The Kansas City Star used to do this and essentially does so now by only offering teaser, headline articles. See <https://www.kansascity.com/>.

²⁹For instance, using my individual account (using up one of my free 10 Perma.cc links), I linked a news story, Pamela Samuelson, *The US Copyright Office Section 512 Study: Why the Entertainment Industry is Claiming Victory*, Kluwer Copyright Blog, May 25, 2020, http://copyrightblog.kluweriplaw.com/2020/05/25/the-us-copyright-office-section-512-study-why-the-entertainment-industry-is-claiming-victory/?doing_wp_cron=1590520010.

³⁰See infra n. 40 for an example of how online news becomes inaccessible with use and over time.

³¹See Perma.cc, *Perma.cc User Guide: Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#private-records> (last visited Oct. 5, 2020).

³²See *id.*

³³See 17 U.S.C.A. § 512(c)(1)(C), (c)(3) (Westlaw through Pub. L. No. 116-152).

³⁴See *infra* section “How the DMCA Works to Protect Internet Service Providers.”

³⁵See *infra* section “Liability for Perma.cc and Registrars under 17 U.S.C. § 512(c), Contributory Infringement and Vicarious Liability.”

following sections.³⁵ Furthermore, even if plaintiffs utilize the notice-and-takedown provisions of § 512(c), the provisions only protect “service providers,” which would be Perma.cc, and might not include library or other registrars, who have no role in Perma.cc’s notice-and-takedown process.³⁶ Not being responsible for answering to § 512(c), it is unlikely that the protections afforded in the section would protect registrar libraries and institutions. More concretely, notice-and-takedown provisions do not protect end users of Perma.cc.³⁷

For Perma.cc users, navigating legal pitfalls for archiving content when some Web materials are unauthorized for copying is difficult at best. The origin of these problems is that many Web materials are unauthorized for copying, and some that are authorized may transition from being freely copyable (and accessible) to uncopyrightable (and inaccessible) as they become formally published.³⁸ Most critically, authorized use probably does not include Web archiving, which creates back-end access to “paywalled” proprietary materials.

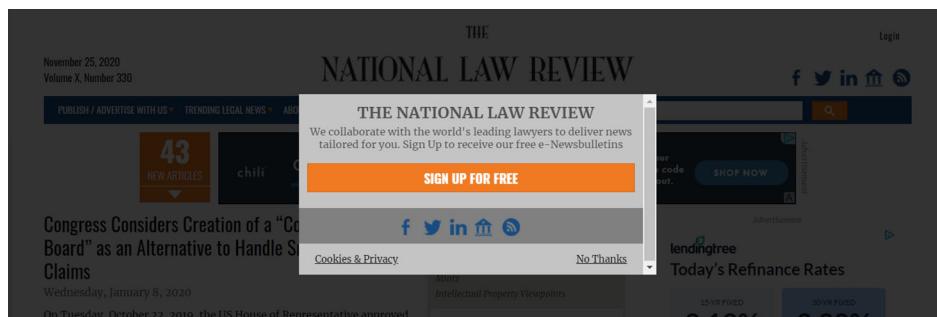
The financial forces driving publication of scholarly and news materials may push what was once freely available behind paywalls.³⁹ Ideally, every

³⁵See 17 U.S.C.A. §§ 501-504 (Westlaw through Pub. L. No. 116-152).

³⁶See Perma.cc, *Perma.cc User Guide/For Registrars*, <https://perma.cc/docs/libraries#libraries-authority> (last visited May 22, 2020). Registrar libraries are the main point of contact between Perma.cc and end users; they remove users and ensure that “those in [their] network do not abuse their account,” but there is nothing about DMCA required notices or takedown procedures, which seem to be the sole responsibility of Perma.cc).

³⁷See 17 U.S.C.A. § 512(k)(1) (Westlaw through Pub. L. No. 116-152) (definition of “service provider” covered by the section). This issue is explored in more depth in the following. See text accompanying *infra* n. 232 through 233. The effect of § 512(d), which parallels §512(c), but with respect to hyperlinking rather than archiving, is also discussed in the context of electronic law reviews and journals with Perma.cc hyperlinks. See *infra* n. 235 through 241 and accompanying text.

³⁸An example of a Web news item that became inaccessible is Susan Neuberger Weller, *Congress Considers Creation of a “Copyright Claims Board” as an Alternative to Handle Small Copyright Claims*, Nat'l L. Rev., January 8, 2020, <https://www.natlawreview.com/article/congress-considers-creation-copyright-claims-board-alternative-to-handle-small>, archived at <https://perma.cc/A8XJ-JGR5>. Spending time over the original Web site brings a demand for subscription fees. The subscription is free, but that is not the point. The issue is permissive archiving and accessible Web sites becoming inaccessible.



Yet the perma.cc link still provides access to the archived article. See <https://perma.cc/A8XJ-JGR5>. This is an example of content moving from accessible to inaccessible but being copied as accessible by Perma.cc.

³⁹See Lindsey Ellefson, *Fortune Begins Moving Its Stories Behind Digital Paywall*, The Wrap, Jan. 15, 2020, <https://www.thewrap.com/fortune-heads-behind-digital-paywall>; Stephen Altrogge, *Pay for Play: Why Putting Your Content Behind a Paywall Is the Right Move for Your Web site*, Paperform, Jan. 8, 2019, <https://paperform.co/blog/put-content-behind-paywall> (Apparently, there are “hard” paywalls like the *Wall Street Journal*, and “soft”

contributing user to Perma.cc would get permission for his or her contributions. However, given the nature of academic legal scholars and law review editors, and their need to work quickly and efficiently with hundreds of sources cited in a single article (even if only a fraction of these are to Web materials), this may not be realistic, regardless of having library registrars to train and issue policy for contribution to Perma.cc archives. Furthermore, scholarly authors and journal editors historically are not accustomed to seeking permission for citations (and using a Perma.cc link might trigger such an obligation).

Fair use and “transformative use”

Generally copyright analysis includes analysis of the subject matter for originality or copyrightability,⁴⁰ determination of copyright,⁴¹ and finding illicit copying.⁴² However, there are exceptions to copyright law that allow copying in certain situations. Perma.cc may be able to defend its copying practices by claiming that its archiving is *fair use* or *transformative use*. Our analysis starts with these exceptions or defenses. This section shall consider the application of transformative use to Perma.cc and its users. The following section will consider *fair use* when *transformative use* does not apply.

Statutory fair use and transformative use

The inquiry into whether the use of copyrighted work is transformative in expression and character⁴³ has modified traditional fair use analysis to such

paywalls that allow “metered” free content or place premium content behind paywalls). *But see* Jol Ito, *The Quest to Topple Science-Stymying Academic Paywalls*, Wired, Jan. 4, 2019, <https://www.wired.com/story/ideas-jol-ito-academic-paywalls/>; “[A]cademic publishers started to consolidate. They solidified their grip on the rights to prestigious journals, allowing them to charge for access and exclude the majority of the world from reading research publications—all while extracting billions in dollars of subscription fees from university libraries and corporations.”; Dean Horowitz, *You've Decided to Implement a Paywall. Now What?*, Folio, Jan. 24, 2019, <https://www.foliomag.com/youve-decided-implement-paywall-now/>; “Instituting a paid subscription model online is frankly unwise unless you're focused on customer experience feedback. It is that simple.”; Lindsay McKenzie, *Open Access at the Movies*, Inside Higher Ed, Sept. 10, 2018, <https://www.insidehighered.com/news/2018/09/10/open-access-movement-hits-silver-screen> (film takes on for-profit academic publishers and favors open-access movement).

⁴⁰See *Feist Publ'n.s, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 348 (1991) (“Originality remains the sine qua non of copyright”). See generally, Alfred C. Yen & Joseph P. Liu, *Copyright Law: Essential Cases and Materials* 14–58 (3d ed., West Academic 2016) (coverage of copyrightability and originality).

⁴¹See 17 U.S.C.A. § 501(b) (Westlaw through Pub. L. No. 116-152) (right of legal owner to bring suit for infringement, subject to registration requirements of § 411). All references to the U.S.C. are to Westlaw Edge’s U.S.C.A. due to the closing of law libraries during the COVID pandemic concurrent with the writing and publishing of this article.

⁴²See generally, Yen & Liu, *supra* n. 42, at 229–274 (substantial similarity and improper appropriation tests).

⁴³Transformative use “‘supplant[s]’ the original), or instead adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message; it asks, in other words, whether and to what extent the new work is ‘transformative.’” *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994).

an extent that we have to consider it separately. How do we get to *transformative use* as playing such an important role? We will answer that in this section but must begin with the traditional four factors of fair use analysis. These four factors of fair use are well known as they have been set forth in 17 U.S.C. § 107:

In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
2. the nature of the copyrighted work;
3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. the effect of the use upon the potential market for or value of the copyrighted work.⁴⁴

The factors are not meant to be exclusive.⁴⁵ How much weight to give to each factor is not spelled out by § 107.⁴⁶ However, certain practices have developed in the case law. For instance, the first factor takes preeminence if there is transformative use (treated immediately below).⁴⁷ Otherwise, the fourth factor has been recognized as preeminent.⁴⁸

History of transformative use

Prior to the doctrine of transformative use, which the Supreme Court debuted in *Campbell v. Acuff-Rose Music, Inc.* (1994),⁴⁹ the majority of the Supreme Court (5–4 split) in *Sony v. Universal City Studios* (1984), when applying the first factor analysis, found “noncommercial, nonprofit activity”

⁴⁴17 U.S.C.A. § 107 (Westlaw through Pub. L. No. 116-152).

⁴⁵Leadsinger, Inc. v. BMG Music Pub., 512 F.3d 522, 529 (9th Cir. 2008) (“When conducting a fair use analysis, we are not restricted to these factors; rather, the analysis is a flexible one that we perform on a case-by-case basis.”) (citation omitted); at 4 Nimmer on Copyright, *supra* n. 22, § 13.05 [A][5][b], (“Consider that the factors listed in the statute are preceded by the words ‘shall include,’ and that the statute as a whole defines use of the term ‘including’ as ‘illustrative and not limitative.’ For that reason, the four factors contained in Section 107 are merely by way of example, rather than being an exhaustive enumeration.”); and William F. Patry, 4 Patry on Copyright § 10:156 (Thomson Reuters, through March 2020 update) (“The listing of fair use purposes and factors is thus “not intended to be exhaustive,” and accordingly courts are free to (but need not) consider factors beyond the four enumerated in the statute”).

⁴⁶Cambridge Univ. Press v. Patton, 769 F.3d 1232, 1260 (11th Cir. 2014) (“Accordingly, we find that the District Court erred in giving each of the four factors equal weight, and in treating the four factors as a simple mathematical formula. As we will explain, because of the circumstances of this case, some of the factors weigh more heavily on the fair use determination than others.”).

⁴⁷See *Campbell*, 510 U.S. at 571 (“The more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use”).

⁴⁸From Nimmer on Copyright, “If one looks to the fair use cases, if not always to their stated rationale, this [fourth factor] emerges the most important, and indeed, central fair use factor.” 4 Nimmer on Copyright, *supra* n. 22, at § 13.05[A][4] (citations omitted). See *infra* n. 121.

⁴⁹*Campbell*, 510 U.S. at 571.

by users videotaping programs broadcast into their homes, described as “time-shifting.”⁵⁰ The dissent found this to be a mischaracterization of time-shifting.⁵¹ In truth, I have trouble imagining how videotaping is a nonprofit activity (lacking even a reference to “education”) without at least a touch of commercial activity, but the Supreme Court seemed to be looking for some way to find a use under the first factor that was *not* “commercial”—to carve out ground in between “commercial” and “nonprofit educational” use. Having opened the door to such a middle ground, it is not surprising that “transformative use” fell into it since the doctrine was capable of recognizing noneducational (and even commercial) uses as fair use.

Campbell v. Acuff-Rose Music, Inc. was a music parody case that pointed out that many of the uses exemplified in the preamble of § 107 had commercial aspects.⁵² “If, indeed, commerciality carried presumptive force against a finding of fairness, the presumption would swallow nearly all of the illustrative uses listed in the preamble paragraph of § 107... .”⁵³ Following such reasoning supported the court’s articulation of transformative use, even though not expressly set forth in the four factors.

The enquiry focuses on whether the new work merely supersedes the objects of the original creation, or whether and to what extent it is “transformative,” altering the original with new expression, meaning, or message. The more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use.⁵⁴

What followed over the years was a litany of cases, many in the 2nd Circuit, finding fair use based on transformative use.⁵⁵

⁵⁰Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 449, (1984) (Blackmun dissenting), superseded by statute, Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2877, as recognized in Realnetworks, Inc. v. DVD Copy Control Ass’n, 641 F. Supp. 2d 913, 941 (N.D. Cal. 2009).

⁵¹Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. at 496 (“As one commentator has observed, time-shifting is noncommercial in the same sense that stealing jewelry and wearing it—instead of reselling it—is noncommercial.”).

⁵²*Campbell*, 510 U.S. at 584.

⁵³*Id.* at 584.

⁵⁴*Id.* at 571.

⁵⁵Mattel Inc. v. Walking Mt. Prods., 353 F.3d 792, 802 (9th Cir. 2003) (popularly known as the “Barbie-in-a-blender case,” the court found defendant artist’s portrayal of the Mattel Barbie doll to be transformative, parodic, a “sexualized perspective,” and a comment on “Barbie’s influence on gender roles and the position of Women in society”); Nunez v. Caribbean Int’l News Corp., 235 F.3d 18, 21–22 (1st Cir. 2000); Kelly v. Arriba Soft Corp., 336 F.3d 811, 818 (9th Cir. 2003); Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 608–615 (2nd Cir. 2006) (court applies transformative use to all four factors in analysis); Brownmark Films, LLC v. Comedy Partners, 800 F. Supp. 2d 991, 1001 (E.D. Wis. 2011), aff’d Brownmark Films, LLC v. Comedy Partners, 682 F.3d 687, 693 (7th Cir. 2012); Cariou v. Prince, 714 F.3d 694 705–711 (2nd Cir., 2013) (court applies transformative use to all of the factors); Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 97–100 (2nd Cir. 2014) (court applies transformative use to first, second, and fourth factors, with respect to search function of database); Authors Guild v. Google, Inc., 804 F.3d 202, 214–224 (2nd Cir. 2015) (court applies transformative use to all four factors, with respect to search snippet view functions); Oracle Am., Inc. v. Google LLC, 886 F.3d 1179, 1186 (Fed. Cir. 2018) (applying but not finding transformative use), cert. granted, 140 S. Ct. 520 (2019).

Limits of transformative use doctrine in *American Geophysical Union v. Texaco*

The transformative use doctrine has its limits. Another Second Circuit case, *American Geophysical Union v. Texaco* (2nd Cir. 1994) defines the limits of transformative use, particularly with reference to the issue of archiving and preservation. In *American Geophysical Union* the infringing, “secondary use,” was to copy a journal article for local reference in the offices of a Texaco employee.⁵⁶ The Second Circuit finds:

To the extent that the secondary use involves merely an untransformed duplication, the value generated by the secondary use is little or nothing more than the value that inheres in the original. Rather than making some contribution of new intellectual value and thereby fostering the advancement of the arts and sciences, an untransformed copy is likely to be used simply for the same intrinsic purpose as the original, thereby providing limited justification for a finding of fair use.⁵⁷

Research alone is not transformative use.⁵⁸ By analogy, the Web pages, articles, documents, spreadsheets, etc., on Perma.cc likewise undergo no transformation. There is no new “intellectual value” through transformation. The intrinsic purpose is the same—to provide the information on the Web page, article, document, etc. Perma.cc just does so over a longer period of time. Perhaps that is of “intellectual value,” but where is the transformation? The Web site, article, document, etc., is hosted at a new location that is dedicated to long-term access, but how does that fit into transformative analysis? Perhaps some court will find long-term preservation to be transformative, but where is there a natural limiting principle to protect against abuse? Everything could be copied for purpose of preservation.

⁵⁶See Am. Geophysical Union v. Texaco Inc., 60 F.3d 913, 915 (2d Cir. 1994).

⁵⁷*Id.*at 923. American Geophysical Union has been cited three times within the First Circuit: Soc'y of the Holy Transfiguration Monastery, Inc. v. Gregory, 689 F.3d 29, 60–61 (1st Cir. 2012) (for proposition that “[t]o the extent that the secondary [work] involves merely an untransformed duplication, the value generated by the secondary [work] is little or nothing more than the value that inheres in the original. . . . [T]hereby providing limited justification for a finding of fair use”); Nunez v. Caribbean Int'l News Corp., 235 F.3d 18, 22 (1st Cir. 2000) (for proposition that American Geophysical “express[es] wariness of emphasis on commercial use because most secondary users seek commercial gain.”); and Sony BMG Music Entm't v. Tenenbaum, 672 F. Supp. 2d 217, 235 (D. Mass. 2009) (for the proposition, “[I]t is sensible that a particular unauthorized use should be considered ‘more fair’ when there is no ready market or means to pay for the use, while such an unauthorized use should be considered ‘less fair’ when there is a ready market or means to pay for the use.”).

⁵⁸The Second Circuit clarified this in *Fox News Network v. TVEyes*:

In Texaco, a company was allowing each of its 400 to 500 scientists to photocopy journal articles pertinent to their individual research projects, thus enabling three subscriptions to service the needs of hundreds of scientists We stated that if copying were deemed transformative “simply because [it was done] in the course of doing research,” then “*the concept of a ‘transformative’ use would be extended beyond recognition.*” (*Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 178 n.4 (2d Cir. 2018) (citations omitted) (emphasis added).

Limits of transformative use in scholarly archiving: HathiTrust and Google Books

Another case denoting the boundaries of transformative use analysis is *Authors Guild, Inc. v. HathiTrust* (2nd Cir. 2014).⁵⁹ Defendant HathiTrust listed among their arguments for fair use that by

preserving the copyrighted books in digital form, the HDL [HathiTrust Digital Library] permits members to create a replacement copy of the work, if the member already owned an original copy, the member's original copy is lost, destroyed, or stolen, and a replacement copy is unobtainable at a "fair" price elsewhere.⁶⁰

The court, although finding fair use on other grounds for defendant HathiTrust,⁶¹ did not find it appropriate to rule on this issue because:

The record before the district court does not reflect whether the plaintiffs own copyrights in any works that would be effectively irreplaceable at a fair price by the Libraries and, thus, would be potentially subject to being copied by the Libraries in case of the loss or destruction of an original. The Authors are not entitled to make this argument on behalf of others, because § 501 of "the Copyright Act does not permit copyright holders to choose third parties to bring suits on their behalf."⁶²

This is an interesting development in transformative use analysis. Before the court even begins to answer the question of transformative use (or fair use) based on preservation of old and decaying materials, it has to know whether the plaintiffs are the specific copyright owners of such items. Lacking such proof, the Second Circuit dismissed the case for want of subject-matter jurisdiction. There is no "live controversy for adjudication."⁶³

The Court's failure to quiet the questions about whether it's okay to publicly, digitally archive copyrighted materials left a gray area in the law, making it hard to know whether Perma.cc could successfully invoke a transformative use defense. But wouldn't this issue always arise with any long-term data preservation? There are two possible cases: (1) the materials are linked to and archived in Perma.cc and are still available on the Web (either at their original location or elsewhere), or (2) the materials are linked to and archived in Permal.cc but are not still available on the Web (either at their original location or elsewhere). More of these cases will be discussed in the following, but in case (1) defendants are placed in the

⁵⁹Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014). HathiTrust has been cited once within the First Circuit (the jurisdiction most likely to see litigation over Perma.cc), in *Comerica Bank & Tr., N.A. v. Habib*, No. 17-12418-LTS, 2020 U.S. Dist. LEXIS 1343, at *21 (D. Mass. Jan. 6, 2020) (but in a case that distinguishes defendant's actions as nontransformative).

⁶⁰HathiTrust, 755 F.3d at 92.

⁶¹See *id.* at 97–103 (the court found that full-text searching and access to the disabled were transformative uses favoring the defendant).

⁶²*Id.* at 103–104.

⁶³*Id.* at 104.

awkward position of arguing that because the materials might disappear one day, archiving them is a fair use. It is an instance of unauthorized copying as a form of “access insurance”—something the law of fair use has not recognized. Maybe the law should find access insurance to be fair use, even transformative use, but no court or statute has done so yet.⁶⁴ Nimmer has pointed out that the very absence of a work may build up its demand for additional publication.⁶⁵ Sometimes removing a work from circulation is strategic.

In case (2) the plaintiffs need to establish that they owned the materials previously linked to the Web, and the defendants need to establish the materials are “effectively irreplaceable at a fair price,”⁶⁶ presumably by not being available elsewhere than the original link. In case (2), there are lots of possibilities for availability elsewhere—located at other locations on the Web, placement behind a cyber “paywall,” distributed in print for purchase, or accessible by contacting the original authors or creators of the materials. Thus, while a suit by Authors Guild or a similar collective rights organization might be blocked in the Second Circuit under *HathiTrust* for lacking a live controversy, there may plausibly be plenty of plaintiffs satisfying case (1) or (2) for a case against Perma.cc and its contributors. In the end, HathiTrust (which did provide fair use services to users through full-text searching and disability access)⁶⁷ does not provide much shelter from litigation for Perma.cc, the service of which are confined to archiving.

In 2015, another Authors Guild case, *Authors Guild v. Google*, also failed to provide much cover from litigation for Perma.cc and its users.⁶⁸ The Second Circuit examined the “search and snippet view” functions of Google Books, which did not allow users to view entire works but only limited, “snippet” portions of text.⁶⁹

As with HathiTrust (and iParadigms), the purpose of Google’s copying of the original copyrighted books is to make available significant information about those books, permitting a searcher to identify those that contain a word or term of interest, as well as those that do not include reference to it. In addition, through the ngrams tool, Google allows readers to learn the frequency of usage of selected words in the aggregate corpus of published books in different historical periods. We have

⁶⁴The only instance of preservation being treated as fair use that this author can find is in a Senate Report to the 1976 Copyright Act, commenting that preservation of decaying nitrate film is fair use. See *infra* n. 101 through 103 and accompanying text for a more thorough treatment.

⁶⁵“In short, the fact that a work is out of print cannot mean that its copyright becomes vitiated. Works out of print are published in new editions when the demand becomes sufficient. Such demand may never arise if competitors may freely copy the out-of-print work.” 4 Nimmer on Copyright, *supra* n. 22, § 13.05[A][2][b][i].

⁶⁶*HathiTrust*, 755 F.3d at 103.

⁶⁷See *id.* at 97–103.

⁶⁸*Authors Guild v. Google, Inc.*, 804 F.3d 202, 206 (2d Cir. 2015).

⁶⁹*Authors Guild*, 804 F.3d at 217 (“Snippet view adds important value to the basic transformative search function, which tells only whether and how often the searched term appears in the book.”).

no doubt that the purpose of this copying is the sort of transformative purpose described in Campbell as strongly favoring satisfaction of the first factor.⁷⁰

Besides the search function, the court also found snippet view “adds importantly to the highly transformative purpose of identifying books of interest to the searcher.”⁷¹ Perma.cc does not provide searchable databases or “snippet” views to facilitate identification of relevant materials. It is an archive of Web material that may extend the availability of some of the materials.⁷² It is a preservation tool only.

Limits of Second Circuit and transformative use

The preeminence of the Second Circuit in “transformative use” cases should give us pause.⁷³ We need to see how uniformly its doctrines are accepted. In *Kienitz v. Sconnie Nation LLC*, Judge Easterbrook of the Seventh Circuit eventually criticized the Second Circuit for a line of cases, ending in *Cariou v. Prince*,⁷⁴ that supported the transformative use doctrine:

We’re skeptical of Cariou’s approach, because asking exclusively whether something is “transformative” not only replaces the list in § 107 but also could override 17 U.S.C. § 106(2), which protects derivative works. To say that a new use transforms the work is precisely to say that it is derivative and thus, one might suppose, protected under § 106(2). Cariou and its predecessors in the Second Circuit do not explain how every “transformative use” can be “fair use” without extinguishing the author’s rights under § 106(2).⁷⁵

In 2016, after the *HathiTrust* and *Google Books* cases, even the Second Circuit, perhaps sensitive to criticism from the Seventh Circuit and *Nimmer on Copyright*, would acknowledge that Cariou was “the high-water

⁷⁰*Id.* at 217.

⁷¹*Id.* at 218.

⁷²Perma.cc does permit the addition of metadata and for members of the same organization to view each other’s submissions. For instance, all of the members of a law review might have access to all of the Perma.cc links in articles being published by the review. See Perma.cc, *Perma.cc User Guide Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#organizing-links> (last visited May 19, 2019).

⁷³See *supra* n. 57.

⁷⁴*Cariou v. Prince*, 714 F.3d 694 (2d Cir. 2013), cert. denied 134 S. Ct. 618 (2013). Cariou was criticized by Nimmer because Prince (the defendant artist) was not “trying to create anything with a new meaning or a new message” 4 Nimmer on Copyright, *supra* n. 22, at § 13.05[B][6]. In addition, the plaintiff suffered significant damages when his art show was canceled, usurped by Cariou’s.

Contrary to the district court’s conclusion, the application of this factor does not focus principally on the question of damage to Cariou’s derivative market. We have made clear that “our concern is not whether the secondary use suppresses or even destroys the market for the original work or its potential derivatives, but whether the secondary use usurps the market of the original work.”

Cariou, 714 F.3d at 708 (citing *Blanch v. Koons*, 467 F. 3d 244, 258 (2nd Cir. 2006)). Cariou was cited only once within the First Circuit (the place most likely to see litigation regarding Perma.cc), *Comerica Bank & Tr., N.A. v. Habib*, No. 17-12418-LTS, 2020 U.S. Dist. LEXIS 1343, at *24 (D. Mass. Jan. 6, 2020) (for the limited proposition that creative works favor the plaintiff).

⁷⁵*Kienitz v. Sconnie Nation LLC*, 766 F.3d 756, 758 (7th Cir. 2014), cert. denied S. Ct. 1555 (2015).

mark of our court's recognition of transformative works . . .”⁷⁶ Furthermore, in 2018, in *Fox News Network v. TVEyes*, the Second Circuit reversed the decision of the District Court to find TVEyes’ archiving of videos *not* to be fair use.⁷⁷ Specifically, the Second Circuit pointed out that *Authors Guild v. Google*, a transformative use case, had “tested[ed] the boundaries of fair use.”⁷⁸ Based on these more recent cases, it is now less likely that the Second Circuit, whose opinions have been so important to the development of the transformative use doctrine, would extend the doctrine further to cover the archiving for preservation that Perma.cc facilitates.

The point about the Second Circuit’s extended reach over its history with “transformative use” is made to inform the reader that reliance upon the Second Circuit as the source for all things “transformative” may be imprudent, although certainly the doctrine is treated in other Circuits.⁷⁹

Limits of transformative use in *Cambridge University Press v. Patton*

For universities and libraries, including law schools and libraries relying on Perma.cc for archiving cited materials in scholarly work, an important

⁷⁶TCA TV Corp. v. McCollum, 839 F.3d 168, 181 (2d Cir. 2016), cert. denied, 137 S. Ct. 2175 (2017). For criticism of “transformative use,” see 4 Nimmer on Copyright, *supra* n. 22, at § 13.05 [A][1][b]. “[T]he transformative use standard has become all things to all people.” Michael J. Madison, *A Pattern-Oriented Approach to Fair Use*, 45 Wm. & Mary L. Rev. 1525, 1670 (2004).

⁷⁷See *Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 174 (2d Cir. 2018), rev’g *Fox News Network v. TVEyes, Inc.*, 124 F. Supp. 3d 325 (S.D.N.Y. 2015). District Court had held that “archiving function is fair use” for video and radio clips from broadcasts. *Id.* at 334. The video archiving originally protected by the Southern District of New York was not strictly for preservation but to compliment “searching and indexing functions.” *Id.* Indeed, it was TVEyes users who selected clips for later use and which consequently were stored on TVEye’s servers for more than the default 32 days, set for unselected clips. *Id.* at 333. Likewise, it is Perma.cc end users who determine what is archived, although there is no searching and indexing function, unless we count the metadata, annotating services offered by Perma.cc. See Perma.cc, *Perma.cc user guide/Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#annotating-links> (last visited May 21, 2020).

⁷⁸*TVEyes*, 883 F.3d at 174 (citing *Authors Guild v. Google, Inc.*, 804 F.3d 202, 206 (2d Cir. 2015)).

⁷⁹At the time of this writing, a search of “fair use” AND “transformative” in “Copyright Law Cases” on Lexis+ produces 453 federal cases, with the most such cases in the Ninth Circuit (155), compared to 136 in the Second Circuit. Of course, not all of these cases may have found “transformative use.” The Circuit Court figures include cases in their respective District Courts. Major cases in the Ninth Circuit include *Dr. Seuss Enters., L.P. v. Penguin Books USA, Inc.*, 109 F.3d 1394 (9th Cir. 1997); *L.A. News Serv. v. CBS Broad., Inc.*, 305 F.3d 924 (9th Cir. 2002), *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007); *Monge v. Maya Magazines, Inc.*, 688 F.3d 1164 (9th Cir. 2012); and *Seltzer v. Green Day, Inc.*, 725 F.3d 1170 (9th Cir. 2013). Important Second Circuit cases have already been set forth in *supra* n. 57. The First Circuit, where Perma.cc and Harvard find their home, includes only 18 cases (of which only two are appellate cases). The most important case from that circuit is perhaps *Nunez v. Caribbean Int’l News Corp.*, 235 F.3d 18 (1st Cir. 2000), which dealt with whether photographs of a naked beauty pageant winner appearing without authorization in a local television news program were fair use—they were. But Soc’y of the Holy Transfiguration Monastery, Inc. v. Gregory, 689 F.3d 29, 60-61 (1st Cir. 2012) is also important for assertion that “[t]o the extent that the secondary [work] involves merely an untransformed duplication, the value generated by the secondary [work] is little or nothing more than the value that inheres in the original. . . . [T]hereby providing limited justification for a finding of fair use.” Similarly, when cases are selected in Westlaw Edge based on Topic No. 99 and Key Number 53.2, “Fair use and other permitted uses in general” and then limited to those cases that use “transformative” in the text, the Second Circuit Court of Appeals has 36 cases and the Ninth Circuit Court of Appeals has 25. In this instance, cases from the district court for each circuit are not included. The Second and Ninth Circuits lead in the number of transformative use cases. In Westlaw Edge findings in the First Circuit include the same cases found in Lexis+.

limitation on transformative use might be found in the Eleventh Circuit case of *Cambridge University Press v. Patton*.⁸⁰ The case comes close to considering *archival preservation* but falls short and is really about providing free, *Web archival access* for students to educational materials. The case centered on the issue of university library electronic course reserves consisting of the digitization of book chapters (sometimes multiple chapters in the same book). The Eleventh Circuit applied the test from *Acuff*: “whether the new work merely supersede[s] the objects of the original creation, or instead adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message . . .”⁸¹ The court ultimately found that the defendant university library’s digitization of course reserve did not qualify as a transformative use:

Here, Defendants’ use of excerpts of Plaintiffs’ works is not transformative. The excerpts of Plaintiffs’ works posted on GSU’s [Georgia State University’s] electronic reserve system are verbatim copies of portions of the original books which have merely been converted into a digital format. Although a professor may arrange these excerpts into a particular order or combination for use in a college course, this does not imbue the excerpts themselves with any more than a de minimis amount of new meaning. See Princeton University Press . . . (“[I]f you make verbatim copies of 95 pages of a 316-page book, you have not transformed the 95 pages very much—even if you juxtapose them to excerpts from other works.”).⁸²

Library archives like Georgia State University’s share common attributes with Perma.cc. The reasoning of the case is apropos to our analysis of Perma.cc’s service and transformative use. Like Perma.cc, the defendant was a nonprofit—or technically the key officers of a nonprofit university (Georgia State University). The use was educational, but the placing of print book chapters into a digital course reserve system, which could deliver items by links, failed transformative use. Perma.cc makes verbatim copies of the Web materials it collects. Under the reasoning of the Eleventh Circuit, Perma.cc’s archival preservation of digital materials already on the Web would not be found to be transformative because such action “does not imbue the excerpts [or Web material] themselves with any more than a de minimis amount of new meaning.”⁸³ The only distinction is that it is Perma.cc end users who are requesting the Perma.cc copy Web materials, and because of that we must look to theories of vicarious liability,

⁸⁰*Cambridge Univ. Press v. Patton*, 769 F.3d 1232 (11th Cir. 2014), remanded to *Cambridge Univ. Press v. Becker*, 371 F. Supp. 3d 1218 (N.D. Ga. 2016), *aff’d in part and vacated in part*, *Cambridge Univ. Press v. Albert*, 906 F.3d 1290 (11th Cir. 2018), remanded to *Cambridge Univ. Press v. Becker*, No. 1:08-CV-1425-ODE, 2020 U.S. Dist. LEXIS 35134 (N.D. Ga. Mar. 2, 2020).

⁸¹*Cambridge*, 769 F.3d at 1262 (quoting *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579, (1994)).

⁸²*Cambridge*, 769 F.3d at 1262 (quoting *Princeton Univ. Press v. Mich. Document Servs.*, 99 F.3d 1381, 1389 (6th Cir. 1996)).

⁸³*Cambridge*, 769 F.3d at 1262.

contributory infringement, and 17 U.S.C. § 512(c).⁸⁴ However, if *Cambridge* turns out to be applicable precedent, a standard fair use analysis, without consideration of transformative use, is the proper mode analysis, even for the activities of Perma.cc's end users, who merely make verbatim copies of works when placing them in Perma.cc's archive.

Transformative use and the problem of preservation

Some may argue that preservation of materials online is, in and of itself, a transformative use, underscored by the phenomenon of *link rot*. However, our inquiry into transformative use has never come across an instance in case law of preservation being a transformative or fair use.⁸⁵ Furthermore, if it were, then almost any copying could be justified as transformative on the off chance that the original material might become unavailable. There is no natural *limiting principle* to the uses to which transformative use could be applied if every preservation were to qualify as transformative.⁸⁶ Where do we draw the line without predicting which Web materials are likely to become inaccessible? Is it enough that a significant percentage of Web materials, accessible today, will suffer link rot tomorrow? There has to be some way to limit the application of preservation as a transformative use, and if not, then preservation alone should not serve as a limiting principle.⁸⁷

Some may point to the long existence of the Internet Archive's WayBack Machine,⁸⁸ which seems to have escaped liability over copyright, as indication that preservation is transformative use. There is no

⁸⁴See *infra* section "Liability for Perma.cc and Registrars under 17 U.S.C. § 512(c), Contributory Infringement and Vicarious Liability."

⁸⁵The Second Circuit in *Texaco* refrained from ruling against all archival copying. We do not mean to suggest that *no instance of archival copying would be fair use*, but the first factor tilts against *Texaco* in this case because the making of copies to be placed on the shelf in Chickering's office is part of a systematic process of encouraging employee researchers to copy articles so as to multiply available copies while avoiding payment. *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 920 (2d Cir. 1994). The key to *Texaco* is the avoidance of payment. That is not really that different from Perma.cc's archival copying, regardless of whether it is for nonprofit institutions.

⁸⁶An example of application of the limiting principle and fair use is found in *Princeton Univ. Press v. Mich. Document Servs.*, 99 F.3d 1381, 1386 n.2 (6th Cir. 1996):

And if the fairness of making copies depends on what the ultimate consumer does with the copies, it is hard to see how the manufacture of pirated editions of any copyrighted work of scholarship could ever be an unfair use. . . . [T]he dissenters' suggestion—which proposes no limiting principle—runs counter to the legislative history of the Copyright Act and has properly been rejected by the courts.

Id.

⁸⁷We could argue that the limiting principle is citation in a scholarly article or court document, but even so, not all the materials underlying the Web citation will disappear due to link rot. The same problem of a limiting principle is just as relevant with the preservation of Web material that is in fact cited in scholarship.

⁸⁸Internet Archive, *WayBack Machine*, <https://archive.org/web/> (last visited May 21, 2020) (emphasis added).

case law directly on point,⁸⁹ but the WayBack Machine does something Perma.cc does not—it documents the transition of Web sites over time. The Internet Archive also honors “tagging files for robot exclusion,”⁹⁰ unlike Perma.cc, which archives material with “noarchive” metatags, but with limited access to end users and their organizations.⁹¹ Like Perma.cc, the Internet Archive also adheres to takedown requests by authors and publishers.⁹² If anything, the Internet Archive has appeased those authors and publishers who do not want their materials included in the Archive, perhaps thus escaping liability by taking sites down. Unlike Perma.cc, the Internet Archive allows for Web addresses to be searched. In the end, the Internet Archive, with the right case, may be subject to liability. This article finds no cover for archives under the Copyright Code.

⁸⁹The sole published decision involving the Internet Archive is *Internet Archive v. Shell*, 505 F. Supp. 2d 755 (D. Co. 2007). Oddly, Internet Archive filed a declaratory relief action to substantiate that federal copyright law preempts the contractual provisions of Shell’s Web site, which the Internet Archive had copied. There is nothing in the decision to provide guidance on actual copyright issues of Internet Archive’s copying Shell’s Web site.

Researching further, the Internet Archive’s Motion to Dismiss Counterclaims suggest that a copyright infringement claim was once been part of the pleadings.

After threatening plaintiff and counterclaim-defendant Internet Archive, a non-profit digital library, with a baseless lawsuit challenging the Archive’s mission of creating a historical record of the Internet that is accessible to the public (*a practice that has recently been held lawful with respect to other archivists by at least two courts*), defendant and counterclaimant Suzanne Shell (“Shell”) has now attempted to turn up the heat on this non-profit organization with a series of far-fetched and legally inapposite counterclaims. With the exception of her copyright infringement claim (which is substantively without merit but probably properly alleged), Shell’s counterclaims fail to state a claim on which relief can be granted.

Internet Archive’s Motion to Dismiss Counterclaims at 1, *Internet Archive v. Shell*, 505 F. Supp. 2d. 755 (D. Co. 2007) (No. 06-cv-01726-LTB-CBS), 2006 WL 3851749, at 1 (emphasis added). I have been unable to locate the cases that have upheld archiving the Internet, as mentioned in the Motion. However, the original suit was filed in the Northern District of California. The Internet Archive’s original Complaint for Declaratory Relief of Copyright Non-Infringement did determine the action originally arose under copyright law. Complaint for Declaratory Relief of Copyright Non-Infringement at 3, ¶12, *Internet Archive v. Shell* (2006) (No. 3:06-cv-00397) (N.D. Cal. Jan. 20 2006) (available on Bloomberg Docket Search). The Plaintiff, Shell, was unsatisfied with simple removal of her Web site from the Internet Archive and demanded \$100,000 in damages for infringement. *Id.* at 3–4, ¶ 13–15. Internet Archive’s claim was that “Internet Archive has not violated Shell’s copyrights. Internet Archive made fair and permissible use of the www.profanec-justice.org Web site for the purposes of offering permanent access for researchers, historians, scholars, and the public to historical collections that exist in digital format.” *Id.* at 4, ¶ 20. According to Bloomberg Law Docket Search, this case was eventually transferred to the District of Colorado, effective Sept. 5, 2006 (per *Receipt of transferred case from USDC-Colorado*), and resulted in the published opinion discussed previously. I have been unable to retrieve any further documentation from Bloomberg (or a document delivery service) since the docket includes material in 2006, prior to comprehensive Pacer coverage.

⁹⁰See *Internet Archive’s Terms of Use, Privacy Policy, and Copyright Policy*, Internet Archive, *WayBack Machine*, <https://archive.org/about/terms.php> (last visited May 21, 2020).

While we collect publicly available Internet documents, sometimes authors and publishers express a desire for their documents not to be included in the Collections (by tagging a file for robot exclusion or by contacting us or the original crawler group). If the author or publisher of some part of the Archive does not want his or her work in our Collections, then we may remove that portion of the Collections without notice.

Id.

⁹¹Perma.cc, See *Perma.cc User Guide: Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#private-records> (last visited May 19, 2020).

⁹²See *supra* n. 93.

In fact, four book publishers recently filed suit against the Internet Archive for providing access to over 1.3 million books—that were still within copyright—during COVID-19 as part of its Open Library’s “National Emergency Library.”⁹³ While the digitization and lending (unrestricted during the early weeks of COVID-19) of books is factually different than Perma.cc’s archiving of Web materials, similarities arise because Perma.cc archives whole documents and creates unrestricted access through a footnote link. What is thankfully lacking, in Perma.cc’s instance, is a group of powerful plaintiffs who are motivated to launch litigation, especially when instead, authors of works can always request that materials be taken down (although the Internet Archive’s National Emergency Library also responded to takedown requests).⁹⁴ So far, Perma.cc is “flying under the radar” without potential plaintiffs that care enough to sue. I am not suggesting that they do so.

Fair use analysis without transformative use

Since we cannot pin preservation neatly under the doctrine of transformative use, which is not required for finding fair use,⁹⁵ our analysis of defenses to copyright infringement from Perma.cc archiving needs to return to standard fair use analysis. In doing so, we will place a heavy emphasis on educational nonprofit use and the effect on the market. Before beginning with the four factors, let us consider the preamble and the legislative history of the fair use code section.

Preamble and legislative history of 17 U.S.C. 107

The preamble for § 107 lists “criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research”⁹⁶ as

⁹³Complaint at 1–2, Hachette Book Group, Inc., HarperCollins, LLC, Wiley & Sons, Inc., and Penguin Random House LLC v. Internet Archive, No. 1:20-cv-04160 (S.D.N.Y. June 6, 2020). See also, Campbell Kwan, *Is Internet Archive’s Unrestricted Lending of 1.4 Million Books Legal?*, Tech Republic (June 15, 2020, 12:52 PM), <https://www.techrepublic.com/article/is-internet-archives-unrestricted-lending-of-1-4-million-books-legal/>. Besides the four named publishers serving in the law suit, the Authors Guild, who had been active in other litigation mentioned in this article, see *supra* n. 62 and 71, wrote an open letter demanding the Internet Archive’s National Emergency Library be shut down. The Author’s Guild, *Authors Guild Sends Open Letter to Internet Archive and Brewster Kahle Demanding Open Library’s “National Emergency Library Shut Down*, <https://www.authorsguild.org/industry-advocacy/ag-sends-open-letter-demanding-national-emergency-library-shutdown/> (last visited Feb. 1, 2021).

⁹⁴See Defendant Internet Archive’s Answer and Affirmative Defenses to the Complaint at 2–3, Hachette Book Group, Inc. v. Internet Archive, No. 1:20-cv-04160-JGK (S.D.N.Y. Jul. 28, 2020).

⁹⁵Transformative use is not necessary for finding of fair use, but it does heavily influence the analysis. “Although such transformative use is not absolutely necessary for a finding of fair use, . . . , the goal of copyright, to promote science and the arts, is generally furthered by the creation of transformative works.” *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994) (citations omitted).

⁹⁶17 U.S.C.A. § 107 (Westlaw through Pub. L. No. 116-152).

examples of what might be fair use, but there are many instances when each of these activities has been found not to constitute fair use.⁹⁷ On the other hand, activities considered to be fair use come from the legislative history rather than the text of the statute:

Quotation of excerpts in a review or criticism for purposes of illustration or comment; quotation of short passages in a scholarly or technical work, for illustration or clarification of the author's observations; use in a *parody* of some of the content of the work parodied; summary of an address or article, with brief quotations, in a news report; reproduction by a library of a portion of a work to replace part of a damaged copy; reproduction by a teacher or student of a small part of a work to illustrate a lesson; reproduction of a work in legislative or judicial proceedings or reports; incidental and fortuitous reproduction, in a newsreel or broadcast, of a work located in the scene of an event being reported.⁹⁸

This is the legislative source of *parody* as a fair use because *parody* is not listed in the text of § 107.

Important to finding an example of fair use for archiving, there is an instance when the same House Report for the 1976 Copyright Act encouraged the preservation of prints for motion pictures made before 1942, which is the date of conversion of "film stock" with a nitrate base.⁹⁹ This is an instance when archival preservation was deemed a fair use by a House of Representatives Report. It is the only instance I could find in which preservation is demarcated as fair use,¹⁰⁰ but it comes not from the courts but from a House Report. The certainty of the loss of all films with a nitrate base and the collaboration of the Library of Congress and the American Film Institute to save these films probably influenced the Report's finding of fair use. For those who would argue that preservation is

⁹⁷E.g., *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 584 (1994) ("If, indeed, commerciality carried presumptive force against a finding of fairness, the presumption would swallow nearly all of the illustrative uses listed in the preamble paragraph of § 107"); *Harper & Row, Publrs. v. Nation Enters.*, 471 U.S. 539, 562 (1985) (news scooping not fair use where "supplant[ed] copyright holder's commercially valuable right of first publication"); *Princeton Univ. Press v. Mich. Document Servs.*, 99 F.3d 1381, 1386 (6th Cir. 1996) ("It is true that the use to which the materials are put by the students who purchase the coursepacks is noncommercial in nature. But the use of the materials by the students is not the use that the publishers are challenging."); *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 919-21 (2d Cir. 1994) (copying for research and archival use not fair use); *Rogers v. Koons*, 960 F.2d 301, 310 (2d Cir. 1992) (criticism or satire of society based on prior copyright work is not fair use); *Basic Books, Inc. v. Kinko's Graphics Corp.*, 758 F. Supp. 1522, 1531 (S.D.N.Y. 1991) ("The use of the Kinko's packets, in the hands of the students, was no doubt educational. However, the use in the hands of Kinko's employees is commercial."); *Encyclopedia Britannica Educ. Corp. v. Crooks*, 542 F. Supp. 1156, 1169-70 (W.D.N.Y. 1982) (cumulative copying of educational broadcasting by school district not fair use).

⁹⁸H.R. Rep. No. 94-1476, at 76 (1976) (emphasis added).

⁹⁹*Id.* at 73 (1976). A problem of particular urgency is that of preserving for posterity prints of motion pictures made before 1942. Aside from the deplorable fact that in a great many cases the only existing copy of a film has been deliberately destroyed, those that remain are in immediate danger of disintegration; they were printed on film stock with a nitrate base that will inevitably decompose in time. The efforts of the Library of Congress, the American Film Institute, and other organizations to rescue and preserve this irreplaceable contribution to our cultural life are to be applauded, and the making of duplicate copies for purposes of archival preservation certainly falls within the scope of "fair use." *Id.*

¹⁰⁰Perhaps this is so because most preservation protections for libraries and archives are found under § 108. See *infra* section "Summation of Fair Use Factors."

generally a fair use, there is scant legislative evidence of it. Without further legislation (and hopefully some cooperation of stakeholders of the Internet, libraries, archives, and educational institutions), evidence supporting preservation as a fair use is unlikely.

Having reviewed the preamble to 17 U.S.C. § 107, it is time to apply the four factors in the section to Perma.cc.

Education nonprofit, noncommercial use

Returning to the four factors in the statute, the first factor of fair use analysis requires us to consider whether “(1) the purpose and character of the use, including whether such use is of a commercial nature or is for non-profit educational purposes . . .”¹⁰¹

The creator of Perma.cc is Harvard’s Library Innovation Lab.¹⁰² No advertising appears on the site. Academic institutions may qualify for free, unlimited service. Academic libraries may sign up to be Perma.cc “registrars,” thereby providing unlimited service to their users.¹⁰³ However, individual users may also sign up, but if not affiliated with a library that is a Perma.cc registrar, they must pay a subscription price after 10 uses.¹⁰⁴ Nonacademic institutions also pay subscription fees.¹⁰⁵ Information about the price of the subscription is not available on the Web site prior to actual subscription. Perma.cc’s strategy seems to be to drive users to pressure their home libraries to become Perma.cc registrars and partners, and only then do end users get unlimited, free access.

What Perma.cc has created is a network of mostly educational and non-profit registrars that are in many instances academic law libraries. The nature of the organization of the network suggests educational, nonprofit use; however, it is not only the nature of the user or network enabling the use, but the nature of the *use* that is critical.¹⁰⁶ Some of the preservation may be of commercial Web sites or materials, such as from news sites or blogs.¹⁰⁷ The Supreme Court has explained that “[t]he crux of the profit/nonprofit distinction is not whether the sole motive of the use is monetary gain but whether the user stands to profit from exploitation of the

¹⁰¹17 U.S.C.A. § 107 (Westlaw through Pub. L. No. 116-152).

¹⁰²See Perma.cc, *About Perma.cc*, <https://perma.cc/about> (last visited May 22, 2020).

¹⁰³See Perma.cc, *Perma.cc for Libraries*, <https://perma.cc/libraries> (last visited May 22, 2020).

¹⁰⁴See Perma.cc, *Sign up with Perma.cc*, <https://perma.cc/sign-up> (last visited. May 22, 2020).

¹⁰⁵See Perma.cc, *Perma.cc User Guide/Accounts*, <https://perma.cc/docs/accounts#paid-service> (last visited May 22, 2020).

¹⁰⁶See Cambridge Univ. Press v. Patton, 769 F.3d 1232, 1264 (11th Cir. 2014) (“GSU is a nonprofit educational institution. While this is relevant, our inquiry does not end there: we must consider not only the nature of the user, but the use itself.”) (citing Am. Geophysical Union v. Texaco Inc., 60 F.3d 913, 921-22 (2d Cir. 1994)).

¹⁰⁷See *supra* n. 31.

copyrighted material without paying the customary price.”¹⁰⁸ In this instance, we must consider whether any of four possible “users” of Perma.cc-linked, copyrighted works on the Web—namely, Perma.cc, registrars, scholarly end users, and readers of the scholarly output with Perma.cc links—“profit from exploitation.”

The question is do they exploit archived material without paying the “customary price?” If use of Perma.cc allows the circumvention of paywalls or obviates the need to purchase publications once freely accessible on the Web, it is hard to see how at least one of the four copyright users, and perhaps all, are not profiting without “paying the customary price”:

1. Perma.cc is copying and archiving Web material.
2. Registrars are profiting by providing valuable services (in the case of law libraries) to law journals and faculty researchers, whose own end product is enriched with access to archived Perma.cc-linked works.
3. Scholarly authors, in the case of law professors with a law library registrar, are directing Perma.cc to copy and archive Web material without payment, thus increasing the value of their own works over time by avoiding dead links and by distributing the Perma.cc linked works.¹⁰⁹
4. Ultimately, the “downstream” readers of scholarly work may gain access to, and perhaps even download (thereby copying), the materials cited in the scholarly works, which in some cases, are no longer otherwise accessible by original Web links.

Even in the instance of genuine scholarship, exploitation can occur. Perma.cc links make it possible for the downstream readers to access works (if no longer available online) that in a print environment would likely have at least required a trip to the library (which indeed had paid for copies of the material). Even with respect to Web materials, which are no longer available online, downstream readers would have had to seek out for-profit databases, which they or their host institutions subscribed to, perhaps even paying for just the one Perma.cc-linked work. The paywall has been circumvented without its due payment. Everyone from the downstream reader to those up the stream or Perma.cc archival process benefits from the free access provided by Perma.cc. I write this not because I am in favor of our intellectual material being secured by paywalls (I am not) but because we must consider the “customary price” in current fair use and copyright analysis.

¹⁰⁸Harper & Row, 471 U.S. 539, 562 (1985). Also cited in *Cambridge*, 769 F.3d at 1265.

¹⁰⁹Distribution is an exclusive right of authors, just as reproduction is. See 17 U.S.C.A. § 106 (1), (3) (Westlaw through Pub. L. No. 116-152).

True, in some instances copyright owners may have granted permission for copying their Web material.¹¹⁰ It is also true that some material may be in the public domain, such as federal government reports (the example on Perma.cc's Web site is of the FBI's Top Ten List).¹¹¹ However, the utility of Perma.cc links will likely drive much content into the archive that is outside authorized use and the public domain, just as is the case with the Internet Archive's WayBack Machine.¹¹² Indeed, I was able to do it easily enough in my test trial of Perma.cc.¹¹³

From a policy perspective, it is not a simple case of Perma.cc depriving authors of Web materials' rewards for the "sweat of their brow." Instead, Perma.cc offers a public good¹¹⁴ that might otherwise be unavailable: a solution to link rot. Copyright law should not perfunctorily disallow a solution for scholarly communities to benefit from services like Perma.cc, whether or not they pay the "customary price." More subtle analysis is needed. At this point, to best serve our analysis, we need to skip to the all-important fourth factor, the effect upon the potential market.¹¹⁵

Effect upon the potential market

The fair use factor reads: "(4) the effect of the use upon the potential market for or value of the copyrighted work."¹¹⁶ The Supreme Court has stated: "The fourth factor is the 'most important, and indeed, central fair use factor.'"¹¹⁷ As Nimmer has stated: "If one looks to the fair use cases, if not always to their stated rationale, this [fourth factor] emerges the most important, and indeed, central fair use factor."¹¹⁸

¹¹⁰This was the case in *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 443–444 (1984) (less than 10% of copying television broadcasts was authorized) (Blackmun dissenting), superseded by statute Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2877, as recognized in *Realnetworks, Inc. v. DVD Copy Control Ass'n*, 641 F. Supp. 2d 913, 941 (N.D. Cal. 2009) (superseded by the anticircumvention and antitrafficking sections of the DMCA, which are different than the code section we are dealing with here).

¹¹¹See *supra* n. 4 and accompanying text.

¹¹²See Internet Archive, *WayBack Machine*, <https://archive.org/web/> (last visited May 21, 2020).

¹¹³See *supra* n. 31 and 40.

¹¹⁴See *Yen & Liu*, *supra* n. 42, at 6–11. See also *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 352–354 (1991) (rejecting "sweat of the brow" doctrine).

¹¹⁵17 U.S.C.A. § 107(4) (Westlaw through Pub. L. No. 116-152).

¹¹⁶*Id.*

¹¹⁷*Stewart v. Abend*, 495 U.S. 207, 238 (1990) (quoting 3 Nimmer on Copyright, *supra* n. 22, at § 13.05[A]). See also *Rogers v. Koons*, 960 F.2d 301, 311 (2d Cir. 1992).

¹¹⁸4 Nimmer on Copyright, *supra* n. 22, at § 13.05[A][4] (citations omitted). See *Robinson v. Random House, Inc.*, 877 F. Supp. 830, 842 & n.4, 843 (S.D.N.Y. 1995) ("most important of the four," "vital fourth factor"); *Sandoval v. New Line Cinema Corp.*, 973 F. Supp. 409, 414 (S.D.N.Y. 1997) (This [fourth] factor is arguably the most important of the four enumerated factors of the fair use analysis), *aff'd on other grounds*, 147 F.3d 215 (2d Cir. 1998); *Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1275 (11th Cir. 2014) ("Because Defendants' use is nontransformative and fulfills the educational purposes that Plaintiffs, at least in part, market their works for, the threat of market substitution here is great and thus the fourth factor looms large in the overall fair use

Works that were once freely available on the Web may move behind paywalls, move to a subscription news archive or service,¹¹⁹ or be published as books or ebooks. Perma.cc may have cited to and preserved early drafts of such works that were once freely available on the Web but which have been removed. Perhaps such movement or change in status from freely accessible to unavailable does not seem realistic, but I remind the reader of my experience with ABA Cloud-Based Ethics Opinions,¹²⁰ which were once freely accessible and are now behind a paywall. In such situations, there may be real economic loss from readers of scholarly articles who choose to access Perma.cc links rather than pay up at paywalls (this is “market substitution”),¹²¹ but there is more to the analysis than the prospective plaintiff’s losses. There are potential markets here for archived news and blogs. There is even a potential market for archiving garden-variety Web sites—Weather.com on a given day. We must look at the overall market (including net positive effects, perhaps offered by Perma.cc).¹²²

Again, Nimmer is important to our analysis. He points out the danger of circularity in evaluating potential markets:

A danger of circularity is posed here—a potential market, no matter how unlikely, has always been supplanted in every fair use case, to the extent that the defendant, by definition, has made some actual use of plaintiff’s work, which use could in turn be defined in terms of the relevant potential market. In other words, it is a given in every fair use case that plaintiff suffers a loss of a potential market if that potential is defined as the theoretical market for licensing the very use at bar. For example, if the plaintiff complains that snippets of her rock-and-roll song lyrics have been appropriated by defendant for a quiz in its book of 1960’s trivia, one could define the supplanted potential market as the possibility of licensing rock song lyrics for quiz books.¹²³

analysis.”). Nimmer comments on an empirical study affirming the importance of the fourth factor. See 4 Nimmer on Copyright, *supra* n. 22, at § 13.05[A][4], fn. 217.4 (commenting on Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions, 1978–2005*, 156 U. Pa. L. Rev. 549, 584 (2008)).

¹¹⁹The movement of “teaser” news stories on the Web from free viewing to subscription only is discussed in *supra* n. 29 and 40 and accompanying text.

¹²⁰See *supra* n. 28 and accompanying text.

¹²¹See *Cambridge Univ. Press*, 769 F.3d at 1275. Market substitution affects the first, third, and fourth factors. The reference to market substitution ensnares the third factor in the fourth. The reference in third factor analysis to purpose and character likewise implicates the first fair use factor. Copying a work verbatim not only weighs against fair use under the third factor; it also “may reveal a dearth of transformative character” inclining the first factor against fair use, plus show more likely market harm by superseding the original, thus disfavoring fair use under the fourth factor. 4 Nimmer on Copyright, *supra* n. 22, at § 13.05[C][2].

¹²²See 4 Patry on Copyright, *supra* n. 47, at § 10:55. “For example, a search engine’s retrieval and display of excerpts from a copyrighted book will likely lead to increased sales, as well as providing valuable information to the individual who initiated the query. Under these facts, the fourth factor should weigh in defendant’s favor.” *Id.* But such use of search engines is transformative, and such is not the case here with Perma.cc. See also *Rogers v. Koons*, 960 F.2d 301, 311–312 (2d Cir. 1992) (“Under this factor a balance must be struck between the benefit gained by the copyright owner when the copying is found an unfair use and the benefit gained by the public when the use is held to be fair. The less adverse impact on the owner, the less public benefit need be shown to sustain non-commercial fair use.”).

¹²³4 Nimmer on Copyright, *supra* n. 22, § 13.05[A][4] (citations omitted).

We might suppose there is a potential market for archived, cited sources. Owners of copyrighted works are not entitled to dominion over every “theoretical” market. Perma.cc creates a single market source for a solution. Without it, the market of solutions is fractured, incomplete, and with diverse solutions for some, but not all, Web citations in an article. Arguably, owners of Web content do not have an interest in or capacity of occupying the niche of a comprehensive solution that Perma.cc offers. But moreover, do they have the right through copyright law to see that the market for archived Web materials stays fractured, and multisourced—a world without Perma.cc? Nimmer’s commentary suggests that there are limits to copyright dominion in potential markets. Copyright owners might be interested in preservation archiving of works they own, but they probably lack the will to work with others, even competitors, for a single solution like Perma.cc. Without Perma.cc, there is a *market failure* to provide a comprehensive answer to link rot in the context of scholarly citations.

Professor Wendy Gordon claims that *market failure* is a rationale for limiting the rights of copyright owners.¹²⁴ Gordon proposes: “Fair use should be awarded to the defendant in a copyright infringement action when (1) market failure is present; (2) transfer of the use to defendant is socially desirable; and (3) an award of fair use would not cause substantial injury to the incentives of the plaintiff copyright owner.”¹²⁵ As further elaboration of the principle of market failure, “Market failure should be found only when the defendant can prove that the copyright owner would refuse to license out of a desire unrelated to the goals of copyright—notably, a desire to keep certain information from the public.”¹²⁶ Parody and criticism are exhibit A of an owner having an interest to keep certain information from the public.¹²⁷ But Perma.cc does not easily fit within these examples, nor is it necessarily true that copyright owners would refuse to license access to materials otherwise archived by Perma.cc. What copyright owners might not be willing to do is license their works with their competitors in a comprehensive, long-term archival solution to link rot like Perma.cc. They may prefer their own individual solutions to price, market, brand, and control of their own works. Even if we apply the market failure theory to a licensing of a comprehensive solution, there may be “substantial injury” to some copyright owners if they had moved their works behind a

¹²⁴Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of The Betamax Case and Its Predecessors*, 82 Columbia L. Rev. 1600, 1614 (1982).

¹²⁵*Id.*

¹²⁶*Id.* at 1634.

¹²⁷See Fisher v. Dees, 794 F.2d 432, 437 (9th Cir. 1986) (“The parody defense to copyright infringement exists precisely to make possible a use that generally cannot be bought.”) (citing Gordon, *supra* n. 127, at 1633).

paywall or submitted for formal publication. Furthermore, market failure as a theory does not have wide-scale adoption.

To illustrate the limited application of market failure as a theory, consider *Clean Flicks of Colo., LLC v. Soderbergh*.¹²⁸ Clean Flicks purchased legal copies of DVD movies or obtained them from customers, it then replaced the DVD with a DVD redacted for sexuality, violence, and offensive language according to a menu of specifications selected by the customer. There were other defendants, using different redaction techniques. Plaintiffs from the motion picture industry sued for infringement. Among other things, defendants “argued lack of market harm”:

The counterclaim defendants contend that there is no adverse effect from their use of the movies on the value of the copyrighted work to the Studios. They suggest that the Studios benefit because they are selling more copies of their movies as a result of the editing parties’ practice of maintaining a one-to-one ratio of the original and edited versions. It is assumed that the consumers of the edited versions would not have themselves purchased the authorized versions because of the objectionable content and the Studios do not compete in this alternative market.¹²⁹

The District Court was unsympathetic, instead championing “the intrinsic value of the right to control the content of the copyrighted work which is the essence of the law of copyright.”¹³⁰ Nimmer is critical:

By framing the matter in terms of “a question of what audience the copyright owner wants to reach,” the decision ignores the logic of prior cases that construe this fourth factor as favoring fair use to the extent defendant’s work *fills a market niche* that plaintiff has no interest in occupying.¹³¹

Clean Flicks and codefendants filled a “market niche,” but the copyright owners’ preference to serve only certain markets took precedence.

Under the theory of “market failure” propounded by Professor Gordon,¹³² Clean Flicks and codefendants occupied a niche that movie studios were not filling. Market failure was present, transfer to the defendants was “socially desirable,” and the movie studios actually benefited from increased sales of their films. Furthermore, the refusal to give permission was motivated by other than traditional objectives of copyright to incentivize new works, but the refusals controlled the uses and presentation of film works. All the elements of Gordon’s theory were satisfied, but the Colorado District Court ruled otherwise, giving the studio unfettered control over the

¹²⁸433 F. Supp. 2d 1236 (D. Colo. 2006).

¹²⁹*Id.* at 1241–1242.

¹³⁰*Id.* at 1242. For a critical review of the Clean Flicks decision, see 4 Nimmer on Copyright, *supra* n. 22, at § 13.05 [A][4].

¹³¹*Id.* at § 13.05 [A][4] (emphasis added) (citing *Clean Flicks*, 433 F. Supp. 2d, at 1242).

¹³²See *supra* n. 127 through 129 and accompanying text.

work. Thus, we cannot entirely rely on theories of market failure to support Perma.cc archiving.

Even assuming the theory of “market failure” were to be widely adopted by courts in the future, it might not squarely apply to the facts Perma.cc presents. First, the vast array of copyright owners whose works are on the Web and who are cited and archived by Perma.cc’s end-user scholars may be unable to fill a niche of a comprehensive solution to scholarly link rot, but some of them (at least a few) will want to provide their own archival solutions into the future.¹³³ The issue is how the market is defined—a comprehensive archival solution for scholarly link rot or a fragmented archival solution in favor of a few copyright owners. Secondly, if that hurdle is passed, it is probably true that having a comprehensive scholarly link rot solution like Perma.cc is socially desirable, but thirdly, there will be adverse financial incentives for copyright owners of the Web material to include the content in free archives.

Just to illustrate, a scholar places a draft of a book or article on the Web prior to publication, but then she withdraws it once a publisher has accepted the manuscript—only to find that her book or article is still available through Perma.cc.¹³⁴ The incentive to place drafts on the Web (perhaps for comment and criticism) is diminished if archives like Perma.cc decrease sales of the final book or journal for the article. Copyright owners, perhaps scholars in their own right, may want to remove papers or materials that they no longer stand behind. If Perma.cc keeps them in perpetuity, copyright owners’ incentive to make their draft or early versions of works available on the Web is diminished. News sites may become reluctant to put up “teaser” news stories,¹³⁵ outside of their subscription services, if their owners know they may be preserved in perpetuity in free archives like Perma.cc. Perhaps the damage is small and the risks slight in the relatively esoteric domain of legal scholarship, but Perma.cc could be used or implemented for all kinds of scholarship, publishing, and even used in news reporting itself (well beyond the original use for legal scholarship and court documents). The damages would then be quite far reaching if the scale of adoption grows into new segments of scholarship, publishing, and news.¹³⁶ In any event, the final element of Gordon’s market failure theory is problematic for Perma.cc.

¹³³See *supra* n. 29.

¹³⁴Of course, this is already the case with scholarly archives offered by the Social Science Research Network (SSRN), BePress, LawArXiv, and ResearchGate. Many preprints are uploaded and, depending on the publisher, they may or may not be removed (or removed for an “embargo” period) after formal publication.

¹³⁵See *supra* n. 29 and 40 and accompanying text.

¹³⁶Perma.cc is encouraging its links to be used in court documents. See Perma.cc, *About Perma.cc*, <https://perma.cc/about#accounts> (“Academic institutions and courts can become registrars of Perma.cc for free, and can provide accounts to their users for free as well”) (last visited May 20, 2020). The Michigan Reporter of Decisions, Michigan Supreme Court is a Perma.cc “Partner” or registrar. *See id.*

On the other hand, scholarship taking advantage of Perma.cc has enhanced value because its citations are stable and free of link rot. Furthermore, there is greater incentive to use sources on the Internet in scholarship because Perma.cc is giving scholarly materials on the Web a permanency that the Web previously lacked. The incentives would be there for Web sites and materials to be placed on the Web because they might be picked up as part of the scholarly discourse. It is both the benefits to society and detriments to copyright holders that must be considered when weighing the fourth factor.¹³⁷ This goes directly to the balance that the Copyright Section of the Constitution strikes: “To 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.”¹³⁸

How the courts evaluate the market benefits to society and benefits and detriments to Web content creators may be decisive in resolving the matter. One of the advantages to copyright owners is that they may be able to point to real or more concrete damages in instances where earlier, no-longer-authorized drafts of their work circulate on the Web by using Perma.cc after publication of a final product behind a paywall or in book form. The benefits to society are more nebulous, considering the value of Perma.cc long-term preservation.

The nature of the work

Copyright scholar William Patry noted in his copyright treatise: “As the Second Circuit observed, citing this treatise, the second factor of the fair use test has rarely played a significant role in the determination of a fair use dispute.”¹³⁹ However, we must consider the nature of the works as part of a thorough analysis.

For our purposes, the works in question may be Web site pages, online government reports, news and blog articles, articles in preprints, unpublished book chapters, PowerPoints, audio/visual files, and in fact any type of material found on the web.¹⁴⁰ Traditionally, factor two favors copyright owners if the works are unpublished or if the works are more creative than nonfictional.

¹³⁷Rogers v. Koons, 960 F.2d 301, 311-12 (2d Cir. 1992) (“Under this factor a balance must be struck between the benefit gained by the copyright owner when the copying is found an unfair use and the benefit gained by the public when the use is held to be fair. The less adverse impact on the owner, the less public benefit need be shown to sustain non-commercial fair use.”).

¹³⁸U.S. Const. art. 1, § 8, cl. 8.

¹³⁹4 Patry on Copyright, *supra* n. 47, at § 10:138 (citing Authors Guild v. Google Inc., 804 F.3d 202, 220 (2d Cir. 2015)).

¹⁴⁰See *supra* n. 6.

Publication

Can a work be available on the Web and unpublished? Nimmer summarizes the law:

“Publication” was a term of art under the 1909 Act. The relevant decisions under this enactment indicated that publication occurred when, by consent of the copyright owner, the original or tangible copies of a work are sold, leased, loaned, given away, or otherwise made available to the general public, or when an authorized offer is made to dispose of the work in any such manner, even if a sale or other such disposition does not in fact occur.¹⁴¹

This definition is traditionally used in cases dealing with the issue of whether general publication has occurred under the 1909 Copyright Act, thereby forcing some works into the public domain. It is not required after January 1, 1978.¹⁴² However, the definition does give us some insight in how courts might rule on the question of whether materials on the Web are published—they are made available to the general public and are not placed in a private archive as in the famous Salinger case (which involved infringement of copyrighted unpublished letters of Salinger, held at a library and used in his biography without his permission).¹⁴³

For Web materials that vanish altogether, with no way to obtain them, the Senate Report to the 1976 Copyright Act provides guidance:

Availability of the work.—A key, though not necessarily determinative, factor in fair use is whether or not the work is available to the potential user. If the work is “out of print” and unavailable for purchase through normal channels, the user may have more justification for reproducing it than in the ordinary case, but the existence of organizations licensed to provide photocopies of out-of-print works at reasonable cost is a factor to be considered. The applicability of the fair use doctrine to unpublished works is narrowly limited since although the work is unavailable, this is the deliberate choice on the part of the copyright owner. Under ordinary circumstances the copyright owner’s “right of first publication” would outweigh any needs of reproduction for classroom purposes.¹⁴⁴

On the one hand, the Senate Report seems to support findings of fair use for “out of print” works, which by analogy would seem to suggest that materials from the Web that have become completely unavailable are fair

¹⁴¹1 Nimmer on Copyright, *supra* n. 22, at § 4.03[A] (emphasis added), citing American Vitagraph, Inc. v. Levy, 659 F.2d 1023 (9th Cir. 1981); Testa v. Janssen, 492 F. Supp. 198 (W.D. Pa. 1980); Brown v. Tabb, 714 F.2d 1088 (11th Cir. 1983); Bartok v. Boosey & Hawkes, Inc., 523 F.2d 941 (2d Cir. 1975); Bell v. Combined Registry Co., 397 F. Supp. 1241 (N.D. Ill. 1975), aff’d, 536 F.2d 164 (7th Cir. 1976), cert. denied, 429 U.S. 1001, 97 S. Ct. 530, 50 L. Ed. 2d 612 (1976); Dowdley v. Phoenix Films, Inc., 199 U.S.P.Q. 579 (S.D.N.Y. 1978); Kramer v. Newman, 749 F. Supp. 542, 549 (S.D.N.Y. 1990).

¹⁴²See Yen & Liu, *supra* n. 42, at 181–182.

¹⁴³Salinger v. Random House, Inc., 811 F.2d 90 (2d Cir. 1987), superseded by statute, Fair Use of Unpublished Works, Pub. L. 102-492, 106 Stat 3145, as recognized in Authors Guild v. Google, Inc., 804 F.3d 202, 212 n.13 (2d Cir. 2015).

¹⁴⁴S. Rep. No. 94-473, at 64 (1975).

use for reproduction. However, the Senate Report goes on to recognize that this unavailability may be due to “deliberate choice” of the copyright owner. Then the report further bolsters the rights of copyright owners, preferring their “right of first publication” over reproduction in classroom use. Now, as I have argued, that placement on the Web is a form of publication. But if not, then works initially placed on the Web that have gone on to be published by a publisher must constitute “first publication.” Then the law might frown upon archived Perma.cc documents that had first circumvented that publication.

Where copyright law fails us is in adequately dealing with works that have “gone out of publication,” for lack of a better term, by no longer being accessible. Nimmer opines, “the fact that a work is out of print cannot mean that its copyright becomes vitiated.”¹⁴⁵ Some works, such as serial magazines, are designed to go out of print.¹⁴⁶ But the Web has presented a whole new scenario. It is an ever-changing typeface and, from a historical perspective, has the same durability as writing words in sand. Preserving those moments of publication is a unique problem. Authors may have perverse incentives to indefinitely preserve works in permanent, freely accessible form on the Web. For scholars, who should be incentivized to rely on such works, the lack of stability is deeply problematic for their writings. Some balance of technological solutions (such as Perma.cc and the Internet Archive)¹⁴⁷ and law needs to be achieved. Perhaps scholars should be given special license for preserving Web materials in archives such as Perma.cc. On the other hand, perhaps no more should be preserved than is necessary to substantiate the citation—something which Perma.cc does not currently facilitate.¹⁴⁸

Informational versus creative works

Informational works may be more freely published under case law. For example, in a Second Circuit case, the Court found: “The work in question was avowedly informational, and such works may be more freely published under Section 107 than those of a creative nature.”¹⁴⁹ Many works on the Web may be thought of as informational, but the analysis is more subtle.¹⁵⁰

¹⁴⁵ See *infra* n. 22, at § 13.05[A][2][b][i].

¹⁴⁶ See *id.*

¹⁴⁷ See *supra* n. 91 through 93 and accompanying text (discussion of Internet Archive).

¹⁴⁸ See *infra* section “Amount and Substantiality of Copying” (discussing the amount of use). For instance, a screenshot of just an applicable passage could be retained instead of the entire work.

¹⁴⁹ Diamond v. Am-Law Pub. Corp., 745 F.2d 142, 148 (2d Cir. 1984) (citing MCA, Inc. v. Wilson, 677 F.2d 180 (2d Cir. 1981)), *overruled in part*, Lesley v. Spike TV, No. CV 04-2758 DT (PLAX), 2005 U.S. Dist. LEXIS 51305, at *22 n.9 (C.D. Cal. July 26, 2005) (overruled over issue of awarding attorney’s fees).

¹⁵⁰ See Harper & Row, Publrs. v. Nation Enters., 471 U.S. 539, 563, 105 S. Ct. 2218, 2232 (1985).

Courts have recognized that copying the expression of otherwise informational works may not favor the defendant under fair use analysis.¹⁵¹ Informational works containing “evaluative, analytical, or subjectively descriptive material that surpasses the bare facts necessary to communicate information, or [which] derives from the author’s experiences or opinions” may actually favor protection and the copyright owners under fair use analysis of the second factor.¹⁵² Thus, even if Web materials archived by Perma.cc are informational, there are many aspects of such works that courts may rule as favoring the plaintiffs pertaining to the second factor of fair use.

Only on an ad hoc basis can this factor be adequately evaluated. Hence, coming up with a general observation about its application to potential cases involving Perma.cc is not possible.

Amount and substantiality of copying

As to the amount of copying, *de minimis* amounts may excuse the defendant from liability, but outside that realm, the courts do not accord an exacting standard:

Like all the fair use factors, it has no precise threshold below which the factor is accorded decisive significance. If the amount copied is very slight in relation to the work as a whole, the third factor might strongly favor the alleged infringer, but that will not always be the case.¹⁵³

Perma.cc archives entire Web pages, documents, and other file types. It is unlikely that this factor will favor Perma.cc or its users during fair use analysis.

In fact, the recent Second Circuit opinion in *Fox News Network v. TVEyes* recognizes that there is no categorical rule against copying entire works:

This factor clearly favors Fox because TVEyes makes available virtually the entirety of the Fox programming that TVEyes users want to see and hear. While “courts have rejected any categorical rule that a copying of the entirety cannot be a fair use,” “a finding of fair use is [less] likely … when the copying is extensive, or encompasses the most important parts of the original.” … In this respect, the TVEyes Watch function is radically dissimilar to the service at issue in Google Books.¹⁵⁴

[Even] within the field of fact works, there are gradations as to the relative proportion of fact and fancy. One may move from sparsely embellished maps and directories to elegantly written biography. The extent to which one must permit expressive language to be copied, in order to assure dissemination of the underlying facts, will thus vary from case to case.

Id.

¹⁵¹See *Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1270 (11th Cir. 2014), remanded to Cambridge Univ. Press v. Becker, 371 F. Supp. 3d 1218 (N.D. Ga. 2016), aff’d in part and vacated in part, Cambridge Univ. Press v. Albert, 906 F.3d 1290 (11th Cir. 2018), remanded to Cambridge Univ. Press v. Becker, No. 1:08-CV-1425-ODE, 2020 U.S. Dist. LEXIS 35134 (N.D. Ga. Mar. 2, 2020).

¹⁵²*Cambridge Univ. Press*, 769 F.3d, at 1232.

¹⁵³*Ringgold v. Black Entertainment Television, Inc.*, 126 F.3d 70, 76 (2nd Cir. 1997).

¹⁵⁴*Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 179 (2d Cir. 2018).

The opinion then goes on through extensive analysis of the Google Book “snippet function” to show that no more than is necessary is copied.¹⁵⁵ In contrast to Google Books, there is no “snippet function” in Perma.cc to limit the Web material to what is necessary for purposes of the citation.

This third factor also considers the substantiality of the copying, the most famous case being *Harper & Row Publishers v. Nation Enterprises*.¹⁵⁶ The Supreme Court agreed with the District Court that the Nation took the “heart” of President Ford’s memoirs in a “news scoop” even though it took only 300 words from the 200,000 word book.¹⁵⁷ It is even harder to argue that archiving the whole Web page, news article, document, or file does not also include the heart of the work. This factor would weigh against Perma.cc and its users.

Summation of fair use factors

As often is the case, analysis of fair use is unsatisfactory. The myriad of balancing factors, opposing assumptions, and subtle interpretations in case law make this an uncertain task. However, assuming that the fourth factor, “effect upon the potential market,” must weigh most heavily and affect our analysis of the other factors, we must look to it for guidance. Unfortunately, the elements of the best theory—market failure—for justifying archival preservation as a solution to link rot do not all favor Perma.cc. In support of Perma.cc is that it is an answer to market failure of the Web to provide a comprehensive solution to link rot for citations to the Web materials in scholarly articles. Perma.cc is socially desirable. Of course, many materials on the Web are in the public domain, and some copyright owners may welcome Perma.cc archiving, which would otherwise burden them. But, even assuming most Web materials are of little economic value, there may be substantial economic harm to a few copyright owners of Web materials that seek to license materials once made freely available. Finally, copyright owners of Web materials that have migrated behind paywalls are willing to license their works, but on the other hand, it is doubtful that such owners would work with others to provide a comprehensive, licensed solution for anything on the Web like Perma.cc offers. How are we to judge Perma.cc based on this fourth factor, the most weighty of all factors? In the end, we must conclude that there is no clear path to fair use in consideration of the fourth factor, except as may be found on an *ad hoc* basis—we must subject each item in the Perma.cc archive to analysis. It is no wonder that Perma.cc’s license agreement shifts copyright

¹⁵⁵*Id.*

¹⁵⁶Harper & Row, Publrs. v. Nation Enters., 471 U.S. 539, 564-65 (1985).

¹⁵⁷*Id.* at 544-545.

compliance responsibilities to registrars and end users.¹⁵⁸ This is a system in which mistakes will be made—there will be infringement.

The first factor, the nature of the use, at first blush is educational (although in some cases, Perma.cc-linked materials may be used in court documents). However, the questions about “payment of the customary price” and exploitation of copyrighted material do not easily dismiss themselves. Law reviews, journals, and court papers are giving value to their readers through the use of Perma.cc links. Legal scholars, judges, and lawyers are benefiting. If there were limitations on the types of materials being used—federal government reports—the uses would be noncommercial, but such sources constitute a fraction of the Web. One issue is whether the use of a Web site, which otherwise has no value, for scholarly purposes gives it economic value for which a price must be paid. But this is exactly the circularity Nimmer warned us against. Another is whether the use of materials that migrate behind paywalls is a commercial use for which the customary price must be paid. Again, we rapidly find ourselves in the fourth factor’s analysis of effects on potential markets. That analysis gave us no clear path to fair use.

We have already introduced the nature of the work into our analysis. It is safe to assume that works on the Web satisfy the publication requirement, so violating a right of first publication is not an issue. We may also assume that most of the works are informational—hence, their value for scholarship. But informational resources are not devoid of “evaluative, analytical, or subjectively descriptive material” and may contain authors’ opinions as well. There is such a wide range of material on the Web, it is impossible to make generalizations, and we should only apply the test on an ad hoc basis, which we cannot do here.

Perma.cc archives entire works—Web pages, news articles, government reports, and other materials, including a variety of file types. It is not set up to glean the snippet of information that is necessary to justify the scholarly citation in a footnote. It captures the whole page because that is what the technology does—in both Web ARChive (WARC) or screenshot formats. The information is useful for checking law review and journal citations to Web sources, but in many instances, more information than is actually necessary may be captured.

A reader of an earlier draft of this work criticized my analysis from finding liability for a lack of fair use as a “tower of suppositions.” The critic pointed out that to begin with, there must be a plaintiff crazy enough to sue (and not just request that material be taken down).¹⁵⁹ Even if someone

¹⁵⁸See *supra* section “Perma.cc and Terms of Service.”

¹⁵⁹For most owners of works cited by Perma.cc, if indeed unhappy with the archiving, the ease of the takedown process should satisfy them. However, in the recent Internet Archive litigation, publishers (Hachette Book Group, Inc., HarperCollins, LLC, Wiley & Sons, Inc., and Penguin Random House LLC) were not satisfied with the option

did sue, the critic found a finding of lack of fair use to be one supposition after another—in the whole, being unlikely. It is argued that risk of copyright liability had to be minimal under these circumstances. But this is the nature of fair use analysis. Fair use analysis is always about factors, each with their own suppositions, which like tumblers in a lock must align to either support or deny fair use. Even one case involving a law school journal or author using Perma.cc would be one too many given the fragile finances of many law schools and the relative marginal benefit of income from law reviews and journals.

In summation, only on an ad hoc basis can any firm conclusions be drawn. But there is no clearly demarcated route to establishing fair use for general use of Perma.cc. Consequently, by turning over policy making to registrars, Perma.cc has shed the laborious task of deciding what can go into its archives,¹⁶⁰ but registrars, end users, and ultimately, the readers of scholarly works may be saddled with an onerous, case-by-case task. Individual end users are in the best position to judge the facts and issues for themselves, but they (even with law backgrounds) are less likely to be adept at navigating the specialized legal knowledge of copyright required without considerable education in the subject—a role registrars must take on.

Library exemptions from liability under 17 U.S.C. 108

At its heart, Perma.cc is run by libraries—not just the Library Innovation Library at Harvard Law School but a network of library partners across the world that act as registrars of users.¹⁶¹ Responsibilities for library registrars include educating users and even setting policies for use of the archive.¹⁶² Thus, Harvard is transferring the responsibility for determining appropriate use of the archive (including copyright) to its network registrars. This should give library registrars pause when considering potential for copyright abuses.

It must be noted that a specific section of the copyright code deals with exemptions for libraries and their users—17 U.S.C. § 108.¹⁶³ The general problem for libraries, under the code section, is that they are only allowed to make

of takedown requests. See *supra* n. 96 through 97 and accompanying text. If Perma.cc “scales up” to more journals and more disciplines, the risk from litigious publishers increases. The risk also increases with pending legislation to create a copyright “small claims court.” See *infra* n. 242 through 250 and accompanying text.

¹⁶⁰That said, Perma.cc may be directly liable for infringing because it, not its users, makes copies of Web materials (although at the instruction of users). See *infra* n. 220. It thus violates the reproduction right of copyright owners under 17 U.S.C.A. § 106(1) (Westlaw through Pub. L. No. 116-152).

¹⁶¹Perma.cc, *About Perma.cc*, <https://perma.cc/about#perma-partners> (last visited May 22, 2020).

¹⁶²Perma.cc, *Perma.cc User Guide/For Registrars*, <https://perma.cc/docs/libraries#libraries-authority> (last visited May 22, 2020).

¹⁶³17 U.S.C.A. § 108 (Westlaw through Pub. L. No. 116-152).

archival copies of what they actually have in their collections.¹⁶⁴ Neither Harvard's law library nor registrant libraries are likely to own the Web materials that users contribute to the Perma.cc archive (although it is possible). Also problematic is several of § 108's provisions require that the "copy ... becomes the property of the user... ."¹⁶⁵ Perma.cc's terms of service do not specify who owns the copy of the original Web material that is in its archive, although the user has control over it (and other users at the same institution may see "Perma Records").¹⁶⁶ Nimmer observes, "Section 108 provides that the exemption does not apply to musical works, to pictorial, graphic or sculptural works, or to motion pictures or other audiovisual works."¹⁶⁷ Users might archive an assortment of file types (and consequently media) with Perma.cc. All in all, § 108 offers no escape from infringement for the libraries and users associated with Perma.cc, nor is it of help to Harvard Law Library and Perma.cc.

Liability for Perma.cc and registrars under 17 U.S.C. 512(c), contributory infringement and vicarious liability

At some point readers may think that the specter of Perma.cc and its end users being hauled into court is unlikely because of the Digital Millennium Copyright Act's notice-and-takedown provisions of 17 U.S.C. § 512(c)(1)(C), which serve as a defense from liability if followed.¹⁶⁸ However effective and inexpensive for copyright owners of Web materials

¹⁶⁴See 2 Nimmer on Copyright, *supra* n. 22, § 8.03 [E][2][e] ("In order for the reproduction exemption to apply, a copy or phonorecord of the work of which a user has requested reproduction must already be in the collection of the library that received the request, or in the collection of another library from whom it may obtain the reproduction.") (citing 17 U.S.C. § 108(d)(e)). The same is true with respect to copies made of unpublished works under 17 U.S.C. § 108(b)(1). See *Association of Am. Medical Colleges v. Carey*, 728 F. Supp. 873 (N.D.N.Y. 1990), rev'd, 928 F.2d 519 (2d Cir. 1991) ("Though there is no case law remotely on point, it is clear that this exception would only apply to unpublished works which are properly in the possession of an archive in the first place.") (cited in 2 Nimmer on Copyright, *supra* n. 22, § 8.03[E][1][a], n. 61).

¹⁶⁵17 U.S.C.A. § 108(d), (e).

¹⁶⁶Perma.cc, *Perma.cc User Guide/Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#organizing-links> (last visited May 4, 2020).

¹⁶⁷2 Nimmer on Copyright, *supra* n. 22, § 8.03.

¹⁶⁸An email from Adam Ziegler of Harvard's Library Innovation Lab confirms that they believe Perma.cc issues with copyrighted, infringing works will be handled through DMCA takedown requests, but end users might have "an issue" as well:

So taking your hypothetical, if you were to direct Perma to try to preserve the book chapter PDFs publicly available at the links you supplied, the software would try to do what you wanted it to do. If you didn't have a valid basis for doing that, or you subsequently abused the fact that you had preserved the book chapters, then Perma might get a DMCA takedown request from the author, and we would respond appropriately. *You might have a potential issue with the author as well.* The same would be true if you took the author's PDFs and made them available on Google Drive, Dropbox, Github, AWS, or any other service, or even if you just emailed the PDFs out to 100 friends. One difference between Perma and these other services is that with Perma you are acting under the auspices of your library (the Perma Registrar), and the library is in a position to do training, set policies (such as "do not preserve PDFs of book chapters" or "do not preserve visual works"), vet those able to use the service for scholarly, educational and research purposes, monitor for possible abuse and, if necessary, suspend privileges.

to find remedy under the notice-and-takedown provisions,¹⁶⁹ they have other potential remedies against Perma.cc grounded in the theories of contributory infringement and vicarious liability when defendants fail to comply with § 512(c)(1)(A) and (B). Admittedly, this section may tax the reader after considering transformative and fair use in detail in previous sections. Nonetheless, § 512 provides cover for third-party service providers, like Perma.cc, and probably library registrars, who provide online achieving services for their users. However, under § 512, that cover is limited. This section addresses theories of contributory infringement and vicarious liability, which may sweep up Perma.cc and its registrars, requiring defense. Section 512 does so without even mentioning the historic defense articulated by the Supreme Court in *Sony v. Universal City Studios*, and referred to as the “staple doctrine” or “substantial non-infringing uses.”¹⁷⁰ We will also review this historic defense.

Perma.cc and contributory infringement under 512(c)

Various provisions operate in concert to offer protection for contributory infringement and vicarious liability under § 512(c). This section will explore this code provision, as well as report on litigation in the Second Circuit that may be key to its interpretation, while at the same time undermining its effective implementation.

How the Digital Millennium Copyright Act (DMCA) works to protect Internet service providers

Section 512(c)(1)(A) is based on contributory infringement. That theory requires contribution to the infringement and knowledge, actual or constructive, of infringement of end users of products or services. Section 512(c)(1)(A) is based on that theory and allows escape for “service providers”¹⁷¹ like Perma.cc if it:

Email from Adam Ziegler, Director, Library Innovation Lab, Harvard Law School Library, to author (Apr. 17, 2020, 12:58 pm CST) (emphasis added) (on file with author). Note the importance of registrar libraries to mitigating risk and ensuring copyright compliance.

¹⁶⁹Actually, copyright owners are not happy with the DMCA, including the notice-and-takedown provisions. See *infra* n. 195 and accompanying text.

¹⁷⁰See *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 441-42 (1984) (Blackmun dissenting), superseded by statute Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2877, as recognized in *Realnetworks, Inc. v. DVD Copy Control Ass'n*, 641 F. Supp. 2d 913, 941 (N.D. Cal. 2009) (superseded by the anticircumvention and antitrafficking sections of the DMCA, which are entirely different than § 512(c)).

¹⁷¹A service provider would include Perma.cc.

(1) Service provider.—

(A) As used in subsection (a), the term “service provider” means an entity offering the transmission, routing, or providing of connections for digital online communications, between or among points

- i. does not have actual knowledge that the material or an activity using the material on the system or network is infringing;
- ii. in the absence of such actual knowledge, is not aware of facts or circumstances from which infringing activity is apparent; or
- iii. upon obtaining such knowledge or awareness, acts expeditiously to remove, or disable access to, the material¹⁷²

Clause (iii) is not the takedown provision that is generally spoken of with respect to the DMCA. That is found in § 512(c)(1)(C) and (c)(3). What is important to note is that Perma.cc has a duty to expeditiously remove items that it knows is infringing under clause (i), or if it has *red flag* knowledge under clause (ii). The plain reading of the statute is that this duty obligates Perma.cc regardless of whether there is a notice from copyright owners.

The Second Circuit in *Viacom Int'l v. YouTube, Inc.* interprets the knowledge requirement of clause (i) to require knowledge of “specific infringing activity”:

In particular, we are persuaded that the basic operation of § 512(c) requires *knowledge or awareness of specific infringing activity*. Under § 512(c)(1)(A), knowledge or awareness alone does not disqualify the service provider; rather, the provider that gains knowledge or awareness of infringing activity retains safe-harbor protection if it “acts expeditiously to remove, or disable access to, the material.” 17 U.S.C. § 512(c)(1)(A)(iii). Thus, the nature of the removal obligation itself contemplates knowledge or awareness of specific infringing material because expeditious removal is possible only if the service provider knows with particularity which items to remove. Indeed, to require expeditious removal in the absence of specific knowledge or awareness would be to mandate an amorphous obligation to “take commercially reasonable steps” in response to a generalized awareness of infringement. Such a view cannot be reconciled with the language of the statute, which requires “expeditious []” action to remove or disable “the material” at issue. 17 U.S.C. § 512(c)(1)(A)(iii).¹⁷³

Thus, the knowledge Perma.cc must have is of “specific infringing activity.” At the same time, the Second Circuit ruled that under 512(m), Perma.cc has no duty to monitor for infringing activity.¹⁷⁴

specified by a user, *of material of the user's choosing*, without modification to the content of the material as sent or received.

(B) As used in this section, other than subsection (a), the term “service provider” means *a provider of online services or network access, or the operator of facilities therefor, and includes an entity described in subparagraph (A)*.

17 U.S.C.A. § 512(k)(1) (Westlaw through Pub. L. No. 116-152).

¹⁷²17 U.S.C.A. § 512(c)(1)(C)(A) (Westlaw through Pub. L. No. 116-152).

¹⁷³*Viacom Int'l, Inc. v. YouTube, Inc.*, 676 F.3d 19, 30-31 (2d Cir. 2012) (emphasis added). Viacom is cited extensively herein, but it has never been cited in the First Circuit (a place where litigation involving Perma.cc is most likely to be brought).

¹⁷⁴“The DMCA provision most relevant to the abrogation inquiry [of common law willful blindness] is § 512(m), which provides that safe harbor protection shall not be conditioned on a service provider monitoring its service or affirmatively seeking facts indicating infringing activity.” *Viacom*, 676 F.3d, at 25. The court left the common

With regard to “red flag knowledge” in § 512(c)(1)(C)(ii), it also requires specific knowledge, but that depends on “whether the provider was subjectively aware of facts that would have made the specific infringement ‘objectively’ obvious to a reasonable person.”¹⁷⁵

Without having a duty to monitor its archive, what might operatives of Perma.cc be aware of or know that would subject it to liability or a duty to “takedown” under § 512(c)(1)(A)(i) or (ii)? Aggrieved parties (whose content was infringed by end users of Perma.cc), rather than resorting to the notice-and-takedown scheme of § 512(c)(1)(C), could bring an action in line with § 512(c)(1)(A)(i) or (ii), which would permit discovery to determine the specific knowledge (either actual or “red flag”) of Perma.cc’s operators. Most plaintiffs may be content with “takedown” procedures under § 512(c)(1)(C), but determined content owners might allege third-party liability without protection of § 512(c)(1)(A) (and consequently without any cover of § 512(c)) because Perma.cc operators had the requisite level of knowledge.

The failure of Viacom’s litigation and its consequences

An example of such a case is *Viacom International v. YouTube*. In that case, YouTube principals were aware of numerous instances of infringing content on their site but took no action. YouTube employees had conducted Web site surveys showing that 75%–80% of all content was copyrighted material.¹⁷⁶ Credit Suisse, advisor to Google, “estimated … more than 60% of copyrighted premium content” of which only 10% was authorized.¹⁷⁷ Internal YouTube communications showed that a YouTube executive requested that infringing material from “Premier League” soccer be taken down ahead of meetings with sports leagues to bid for Premier League rights. Internal YouTube reports by cofounder Jawed Karim disclosed that the shows *Family Guy*, *South Park*, *MTV cribs*, *Daily Show*, *Reno 911*, and *Dave Chapelle* were available on YouTube. The Second Circuit cited this in its opinion that jurors could find the requisite knowledge under § 512(c)(1)(A):

Karim further opined that, “although YouTube is not legally required to monitor content … and complies with DMCA takedown requests, we would benefit from preemptively removing content that is blatantly illegal and likely to attract criticism.” He also noted that “a more thorough analysis” of the issue would be required. At least some of the TV shows to which Karim referred are owned by Viacom. A reasonable juror could conclude from the March 2006 report that Karim knew of the presence of *Viacom-owned material*

¹⁷⁵ law cause of action for willful blindness intact, but without obligating YouTube to a duty to monitor. See *id.* at 39–40.

¹⁷⁶ *Id.* at 31.

¹⁷⁷ *Id.* at 33.

¹⁷⁷ *Id.*

on YouTube, since he presumably located specific clips of the shows in question before he could announce that YouTube hosted the content “[a]s of today.” A reasonable juror could also conclude that Karim believed the clips he located to be infringing (since he refers to them as “blatantly illegal”), and that YouTube did not remove the content from the Web site until conducting “a more thorough analysis,” thus exposing the company to liability in the interim.¹⁷⁸

YouTube executives also had knowledge of infringing Bud Light commercials, but cofounder Stephen Chen argued, “can we please leave these in a bit longer? another [sic] week or two can’t hurt.”¹⁷⁹

What appears to have happened is that YouTube executives waited for a *takedown* notice under § 512(c)(1)(C)¹⁸⁰ when they already, arguably, had the requisite knowledge obliging them to take down the material on their own initiative under § 512(c)(1)(A)(iii), at least if they wanted to preserve their exemption from liability. Consequently, the Second Circuit vacated the order granting summary judgment from below and instructed the “District Court to determine on remand whether any specific infringements of which YouTube had knowledge or awareness correspond to the *clips-in-suit* in these actions.”¹⁸¹

On remand, the New York Southern District Court again ruled for summary judgment in favor of the defendants. How is this possible? The District court ordered the parties to answer for each *clip-in-suit*: “what precise information was given to or reasonably apparent to YouTube identifying the location or site of the infringing matter?”¹⁸² From that, the District Court interpreted the knowledge of “specific infringements” to depend upon identification of the *clips-in-suit*, with the burden on the plaintiffs. Because the Web location of *clips-in-suit* was never identified by evidence uncovered from YouTube,¹⁸³ in part, as a function of the “volume of

¹⁷⁸*Id.* (emphasis added).

¹⁷⁹*Id.* at 33–34.

¹⁸⁰See *Viacom Int'l, Inc. v. Youtube, Inc.*, 940 F. Supp. 2d 110, 119 (S.D.N.Y. April 18, 2013).

YouTube's founders decided to “take down whole movies,” “entire TV shows, like an entire family guy episode” (*id.*), “South Park, and full length anime episodes,” “nudity/porn and any death videos,” but to leave up “music videos,” “news programs,” (email from Brent Hurley to Cuong Do dated Nov. 24, 2005), “sports, commercials” (emails between Jawed Karim, Steve Chen, and Chad Hurley dated Sept. 3, 2005), and “comedy clips (Conan, Leno, etc.)” (email from Jawed Karim to Steve Chen dated Sept. 1, 2005). YouTube then “disabled community flagging for infringement” (*Viacom Opp.* at 41), declined to develop a feature “to send automated email alerts to copyright owners when illegal content was uploaded” (*Viacom 2010 Br.* at 11), and eventually stopped regularly monitoring its site for infringements, deciding instead “to keep substantially all infringing videos on the site as a draw to users, unless and until YouTube received a ‘takedown notice’ from the actual copyright owner identifying a specific infringing clip by URL and demanding its removal from the site.”

Id. (emphasis added).

¹⁸¹*Viacom*, 676 F.3d, at 34.

¹⁸²*Viacom*, 940 F. Supp. 2d, 113 (emphasis added).

¹⁸³See *id.* Pursuant to the first item, I requested the parties to report, for each clip-in-suit, “what precise information was given to or reasonably apparent to YouTube identifying the location or site of the infringing matter?” (Tr. Oct. 12, 2012, p. 29) YouTube submitted a list of 63,060 clips-in-suit, claimed it never received adequate notices of any of those infringements, and challenged plaintiffs to fill in the blanks specifying how they claim such notice was given. *Id.*

material,” the court found that Viacom lacked sufficient proof and after considering other issues,¹⁸⁴ ruled for summary judgment in favor of YouTube.¹⁸⁵

In effect, the District Court eviscerated the requirement for YouTube to take down infringing material on its own initiative under § 512(c)(1)(A) and conflated it with § 512(c)(1)(C). It suggested as much stating:

If, as plaintiffs assert, neither side can determine the presence or absence of specific infringements because of the volume of material, that merely demonstrates the wisdom of the legislative requirement that it be the owner of the copyright, or his agent, *who identifies the infringement by giving the service provider notice.* 17 U.S.C. § 512 (c)(3)(A).¹⁸⁶

What is wrong with this statement is that the notice requirement comes from § 512(c)(1)(C) and, given the plain reading of the statute, is not required in § 512(c)(3)(A). The District Court got it wrong: The code does obligate Internet service providers to take down infringing material in certain instances without the notice from the aggrieved party. The litigation trail ends when the District Court denied plaintiff’s motion to certify a very large class against YouTube.¹⁸⁷

The Southern District of New York’s penultimate decision in *Viacom* may turn out to be an outlier, but did the court have any choice? Given the instructions from the Second Circuit to apply knowledge of “specific infringing activity”¹⁸⁸ to § 512(c)(1)(A)(i) and (ii), the centrality of “clips-in-suit,”¹⁸⁹ and the sheer volume of evidence (none of which tied clips-in-suit with identifying and location information to Viacom and YouTube), the court’s application of a notice requirement to § 512(c)(1)(A)(i) and

¹⁸⁴Willful blindness is the other issue taken up on remand by the District Court. It applies the “specific infringement” standard to willful blindness as well. “As shown by the Court of Appeals’ discussion of ‘red flags,’ under the DMCA, what disqualifies the service provider from the DMCA’s protection is blindness to ‘specific and identifiable instances of infringement.’” *Id.* at 116.

¹⁸⁵See *id.* at 114–115.

¹⁸⁶*Id.* at 115 (emphasis added).

¹⁸⁷*Football Ass’n Premier League v. Youtube, Inc.*, 297 F.R.D. 64, 65 (S.D.N.Y. 2013). The District Court opens its opinion skeptically:

The putative class consists of every person and entity in the world who own infringed copyrighted works, who have or will register them with the U.S. Copyright Office as required, whose works fall into either of two categories: they were the subject of prior infringement which was blocked by YouTube after notice, but suffered additional infringement through subsequent uploads (the “repeat infringement class”), or are musical compositions which defendants tracked, monetized or identified and allowed to be used without proper authorization (the “music publisher class”). Plaintiffs assert that there are “at least thousands of class members” in the Repeat Infringement Class, and “hundreds” in the Music Publisher Class.

Id. (citations omitted).

¹⁸⁸See *Viacom Int’l, Inc. v. YouTube, Inc.*, 676 F.3d 19, 30 (2d Cir. 2012).

¹⁸⁹See *Viacom*, 940 F. Supp. at 113.

(ii)¹⁹⁰ on Viacom (as well as placing the burden of proof on Viacom)¹⁹¹ may be inescapable as a matter of judicial expediency. Other courts may follow. However, even without specifically identifying Viacom, copyright owners' frustration with § 512(c) is palpable and has been documented by a recent Copyright Office Report criticizing the shift in the statute's interpretation to favor online service providers.¹⁹² There is an inconsistency in the way § 512(c) is written and how it has been applied. That dissonance favors archives like Perma.cc.

At least in the Federal Southern District of New York, the lesson for Perma.cc from the *Viacom* litigation is that Web archives, in their role as Internet service providers, can elect a policy to wait for takedown notices without losing safe-harbor protection, regardless of what they may know about infringing content. The knowledge requirement of § 512(c)(1)(A) has been conflated with § 512(c)(1)(C). Whether all of the notice requirements of the latter apply to the former is an issue for another day, except we may be certain that the location of the offending content continues to be a requirement.¹⁹³

As with *Viacom v. YouTube*, for Perma.cc volume is already an issue making knowledge of specific infringing activity, ownership, and location less likely. As of this writing, Permalink has preserved 1,685,273 links with 47,936 end users.¹⁹⁴ Between Perma.cc's volume of links and number of end users, there are certainly infringing uses, but of which does Perma.cc have any specific knowledge? If it gets a takedown notice, Perma.cc acts under the terms of its copyright policy to remove the material.¹⁹⁵ But more on that issue later.¹⁹⁶

The parties that are most likely to have knowledge of specific infringing activities are Perma.cc's registrars, since they set policy for their end users, with respect to archiving content.¹⁹⁷ They also can see what is archived by Perma.cc at the direction of users.¹⁹⁸ Potential liability for registrars is discussed below.¹⁹⁹

¹⁹⁰See *id.* at 115.

¹⁹¹See *id.*

¹⁹²United States Copyright Office, Section 512 of Title 17, at 77–83 (May 2020), <https://www.copyright.gov/policy/section512/section-512-full-report.pdf> (finds § 512 unbalanced and in favor of online service providers).

¹⁹³See Viacom, 940 F. Supp. 2d, at 113.

¹⁹⁴Perma.cc, <https://perma.cc> (last visited June 9, 2020). On a similar scale, *Viacom v. YouTube* in the penultimate case before the District Court involved “63,060 clips-in-suit” for which YouTube had never received “adequate notice” of infringement. Viacom, 940 F. Supp. at 113.

¹⁹⁵Perma.cc, *Copyright Policy*, <https://perma.cc/copyright-policy> (last visited June 9, 2020).

¹⁹⁶See *infra* section “Perma.cc and Vicarious Liability under § 512(c).”

¹⁹⁷See Perma.cc, *Perma.cc User Guide/For Registrars*, <https://perma.cc/docs/libraries#libraries-authority> (last visited May 28, 2020).

¹⁹⁸Supporting this is the fact that Registrars have access to “Private Records” (records of archived Web materials that have a “noarchive” metatag) created by Perma.cc. See Perma.cc, *Perma.cc User Guide/Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#private-records> (last visited May 28, 2020).

¹⁹⁹See *infra* V.D.

In addition, although lacking a duty to monitor, Perma.cc may still be liable for common law willful blindness within the parameters set forth by the Second Circuit in *Viacom*.²⁰⁰

The staple doctrine of “substantial non-infringing uses”

An additional issue related to contributory infringement is whether the staple doctrine of exemption from liability for technologies with “substantial non-infringing uses” from *Sony Corp. of Am. v. Universal City Studios* applies,²⁰¹ or whether it has been abrogated by § 512(c)(1)(A), at least with respect to online service providers. Per *Viacom*, “As a general matter, we interpret a statute to abrogate a common law principle only if the statute ‘speak[s] directly to the question addressed by the common law.’”²⁰² Furthermore, the court in *Viacom* was considering the common law doctrine of willful blindness, and not liability for contributory infringement (which is clearly addressed by § 512(c)(1)(A)), when it discussed abrogation.²⁰³

In *MGM Studios v. Grokster*, in 2005, well after the adoption of the DMCA in 1998, the Supreme Court dismissed the staple exemption because of the defendants’ knowledge of infringement (which rose to the level of inducement).²⁰⁴ However, it did not dismiss the doctrine as abrogated by §512(c). Several other courts have considered § 512(c) safe harbors while considering defendant’s plea for exemptions under the staple doctrine but some with only minimal analysis of “substantial non-infringing uses” except to cite *Sony* or *Grokster*.²⁰⁵ Nimmer writes, “Section 512 limits liability for that

²⁰⁰See *supra* n. 177. On remand to District Court for the Southern District Court of New York, the court found no liability for willful blindness for the specific “clips-in-suit,” no evidence of inducement, and that YouTube was protected under 17 U.S.C.A. 512 (c). *Viacom Int'l, Inc. v. YouTube, Inc.*, 940 F. Supp. 2d 110, 117, 121, 123 (S.D.N.Y. 2013).

²⁰¹*Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 449, (1984) (Blackmun dissenting), superseded by statute, Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2877, as recognized in *Realnetworks, Inc. v. DVD Copy Control Ass'n*, 641 F. Supp. 2d 913, 941 (N.D. Cal. 2009).

²⁰²*Viacom Int'l, Inc. v. YouTube, Inc.*, 676 F.3d 19, 35 (2d Cir. 2012) (citing *Matar v. Dichter*, 563 F.3d 9, 14 (2d Cir. 2009)). See also *United States v. Texas*, 507 U.S. 529, 534, 113 S. Ct. 1631, 1634 (1993).

²⁰³*Viacom*, 676 F.3d at 35.

²⁰⁴*MGM Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 935 (2005) (“Thus, where evidence goes beyond a product’s characteristics or the knowledge that it may be put to infringing uses, and shows statements or actions directed to promoting infringement, Sony’s staple-article rule will not preclude liability.”).

²⁰⁵*EMI Christian Music Grp., Inc. v. MP3tunes, LLC*, 844 F.3d 79, 92 (2d Cir. 2016) (jury’s finding of “red flag” knowledge under § 512(c) was proper—reversing trial court, and “substantial non-infringing uses” not applied as a defense because of inducement), cert. denied *Robertson v. EMI Christian Music Grp., Inc.*, 137 S. Ct. 2269 (2017); *Columbia Pictures Indus. v. Gary Fung*, 710 F.3d 1020, 1031-32, 1043, 1049 (9th Cir. 2013) (affirming summary judgment on behalf of the plaintiffs based on § 512(c) analysis, and only referencing “substantial non-infringing uses” with respect to *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 442, (1984) (Blackmun dissenting), superseded by statute, Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2877, as recognized in *Realnetworks, Inc. v. DVD Copy Control Ass'n*, 641 F. Supp. 2d 913, 941 (N.D. Cal. 2009)), cert. dismissed *Fung v. Columbia Pictures Indus.*, 571 U.S. 1007 (2013); *UMG Recordings, Inc. v. Shelter Capital Partners Ltd. Liab. Co.*, 718 F.3d 1006, 1021-22, 1036 (9th Cir. 2013) (defendant entitled to safe harbor protection of § 512(c), and referenced “substantial non-infringing uses” with respect to *Sony* 464 U.S. at 442 and

conduct to such an extent that it becomes scarcely relevant how the affected conduct would fare under Sony.”²⁰⁶ Regardless of Nimmer’s concerns about the relevancy of *Sony*, the Senate Report accompanying the DMCA makes clear that existing defenses from infringement in the law apply:

Even if a service provider’s activities fall outside the limitations on liability specified in the bill, the service provider is not necessarily an infringer; liability in these circumstances would be adjudicated based on the doctrines of direct, vicarious or contributory liability for infringement as they are articulated in the Copyright Act and in the court decisions interpreting and applying that statute, which are unchanged by section 512. In the event that a service provider does not qualify for the limitation on liability, *it still may claim all of the defenses available to it under current law*. New section 512 simply defines the circumstances under which a service provider, as defined in this Section, may enjoy a limitation on liability for copyright infringement.²⁰⁷

The weight of a Senate Report, which was not the Conference report, on judicial interpretation is always subject to question, but this statement clearly supports the continued application of the staple doctrine and substantial noninfringing uses defense. Thus, in the event of suit, Perma.cc can still claim the exemption of substantial noninfringing uses as a defense to traditional claims brought for contributory infringement, although perhaps with diminished relevance per the operation of § 512(c).

Perma.cc and vicarious liability under 512(c)

Besides liability under § 512(c)(1)(A)(i) and (ii), the Second Circuit considered liability under § 512(c)(1)(B), which is essentially vicarious liability.

A&M Records v. Napster, Inc., 239 F.3d 1004, 1020-21 (9th Cir. 2001); Capitol Records, LLC v. Vimeo, LLC, 972 F. Supp. 2d 500, 519-521, 531 (S.D.N.Y. 2013) (applied 512(c) analysis, but referenced “substantial non-infringing uses” only with respect to Grokster, 545 U.S. at 933), *reconsideration granted* Capitol Records, LLC v. Vimeo, LLC, 972 F. Supp. 2d 537 (S.D.N.Y. 2013), *aff’d in part and vacated in part by remand*, Capitol Records, Ltd. Liab. Co. v. Vimeo, Ltd. Liab. Co., 826 F.3d 78 (2d Cir. 2016), *cert. denied* Capitol Records, Ltd. Liab. Co. v. Vimeo, Ltd. Liab. Co., 826 F.3d 78 (2d Cir. 2016); Disney Enters. v. Hotfile Corp., No. 11-20427-CIV-WILLIAMS, 2013 U.S. Dist. LEXIS 172339, at *99, *110, *123, *158 (S.D. Fla. Aug. 28, 2013) (motion granted for Plaintiff with regards to issue of Defendant’s § 512 defense, but summary judgment not granted with respect to defense based on “substantial non-infringing uses and inducement”) (“The fact that these questions remain makes summary judgment [on contributory infringement] inappropriate on the theories of inducement and contributory infringement liability. And while Hotfile may have a difficult time explaining its “innocence” to a jury, the genuine issues of material fact must be resolved by a jury at trial.”); Arista Records LLC v. Myxer Inc., No. CV 08-03935 GAF (JCx), 2011 U.S. Dist. LEXIS 109668, at *92, *96-*98, *137-*38 (C.D. Cal. Apr. 1, 2011) (Plaintiff’s summary judgment motion to disqualify defendant’s 501(c) defense denied, and motion denied for contributory infringement based on defendant’s evidence of “substantial non-infringing use”); Capitol Records, Inc. v. MP3tunes, LLC, 821 F. Supp. 2d 627, 643-646, 649, 651 (S.D.N.Y. 2011) (defendant’s motion for summary judgment on its defense under §512 (c) granted, except for failure to respond to takedown notices, but defense based on substantial noninfringing uses failed); Arista Records LLC v. USENET.com, 633 F. Supp. 2d 124, 156 (S.D.N.Y. 2009) (defense based on substantial noninfringing uses fails when defendant had continuing contact with users, and § 512(c) not applied as defense because of evidence lost or destroyed that would have illustrated defendant’s state of mind).

²⁰⁶4 Nimmer on Copyright, *supra* n. 22, at § 12A.19[B].

²⁰⁷S. Rep. No. 105-190, at 44-45 (1998). See also H. Rep. No. 105-796 (1998) (Conf. Rep.) (not addressing the issue).

The provision requires that the service provider “does not receive a financial benefit directly attributable to the infringing activity, in a case in which the service provider has the right and ability to control such activity.”²⁰⁸ Perma.cc is receiving a financial benefit from subscription fees of submitting authors who have individual memberships (lacking a library or institutional registrar) and from nonacademic registrars.²⁰⁹ In many cases, these fees may just serve as a stimulus to encourage scholars to pressure their own libraries and institutions to register with Perma.cc, but the fees are a benefit. Furthermore, nonacademic institutions are paying “full freight.” Thus, Perma.cc is financially benefitting, at least at first glance.

The Senate Report (from the Judiciary Committee) accompanying the original DMCA makes clear that in many instances, setup and periodic fees for the online service are acceptable and do not negate the limitation of liability:

In determining whether the financial benefit criterion is satisfied, courts should take a common-sense, fact-based approach, not a formalistic one. In general, a service provider conducting a legitimate business would not be considered to receive a “financial benefit directly attributable to the infringing activity” *where the infringer makes the same kind of payment as non-infringing users of the provider’s service*. Thus, receiving a one-time set-up fee and flat periodic payments for service from a person engaging in infringing activities would not constitute receiving a “financial benefit directly attributable to the infringing activity.” Nor is subparagraph (B) intended to cover fees based on the length of the message (per number of bytes, for example) or by connect time. *It would however, include any such fees where the value of the service lies in providing access to infringing material.*²¹⁰

At first glance, because Perma.cc’s users, whether infringing or noninfringing, would pay the same subscription fee (as individual subscribers or nonacademic institutions), the payment of periodic subscription fees is not an issue. However, the real test is whether the value of Perma.cc “lies in providing access to infringing material.” But Perma.cc’s real value is stopping link rot and letting scholars check one another’s citations to Web materials, rather than serving as a “back channel” for infringing material. If the Senate Report has any interpretive weight (the Conference Report does not address the issue),²¹¹ Perma.cc’s monetary benefits can be ignored under § 512(c)(1)(B).

²⁰⁸17 U.S.C.A. § 512(c)(1)(B) (Westlaw through Pub. L. No. 116-152).

²⁰⁹Courts are also excluded from paying subscription fees. See Perma.cc, *About Perma.cc*, <https://perma.cc/about#accounts> (last visited May 22, 2020). Personal subscriptions are offered at \$10 per month for 10 links, \$25 for 100 links, and \$100 for 500 links. Perma.cc, *Settings, Subscription*, <https://perma.cc/settings/subscription> (last visited June 9, 2020).

²¹⁰S. Rep. No. 105-190, at 44–45 (1998).

²¹¹See generally, H.R. Rep. No. 105-796 (1998) (Conf. Rep.).

The next question is the right and ability to control under § 512(c)(1)(B). The Second Circuit, in assessing the right and ability to control, states that it “requires something more than the ability to remove or block access to materials posted on a service provider’s Web site.”²¹² There is no knowledge requirement in the subsection.²¹³ The Ninth Circuit, in a very similar video streaming case, *UMG Recordings, Inc. v. Shelter Capital Partners LLC*, found that the defendant service provider’s use of filtering did not constitute control.²¹⁴ What then is that ability to control—that “something more” than simply the right and ability to block or take down material? The Second Circuit in *Viacom v. YouTube* clarified:

To date, only one court has found that a service provider had the right and ability to control infringing activity under § 512(c)(1)(B). In *Perfect 10, Inc. v. Cybersett Ventures, Inc.*, 213 F. Supp. 2d 1146 (C.D. Cal. 2002), the court found control where the service provider instituted a monitoring program by which user Web sites received “detailed instructions regard[ing] issues of layout, appearance, and content.” *Id.* at 1173. The service provider also forbade certain types of content and refused access to users who failed to comply with its instructions. *Id.* Similarly, inducement of copyright infringement under *Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 125 S. Ct. 2764, 162 L. Ed. 2d 781 (2005), which “premises liability on purposeful, culpable expression and conduct,” *id.* at 937, might also rise to the level of control under § 512(c)(1)(B). Both of these examples involve a service provider exerting substantial influence on the activities of users, without necessarily—or even frequently—acquiring knowledge of specific infringing activity.²¹⁵

²¹²*Viacom Int’l, Inc. v. YouTube, Inc.*, 676 F.3d 19, 38 (2d Cir. 2012) (quoting *Capitol Records, Inc. v. MP3tunes, LLC*, 821 F. Supp. 2d 627, 645 (S.D.N.Y. 2011)). On remand, the District Court for the Southern District of New York found that YouTube did not have the right and ability to control infringing activity. *Viacom Int’l, Inc. v. YouTube, Inc.*, 940 F. Supp. 2d 110, 122 (S.D.N.Y. 2013).

²¹³See *Viacom*, 676 F.3d at 36.

²¹⁴*UMG Recordings, Inc. v. Shelter Capital Partners Ltd. Liab. Co.*, 718 F.3d 1006, 1030 (9th Cir. 2013).

We agree with the Second Circuit and hold that, in order to have the “right and ability to control,” the service provider must “exert[] substantial influence on the activities of users.” “Substantial influence” may include, as the Second Circuit suggested, high levels of control over activities of users, as in *Cybersett*. Or it may include purposeful conduct, as in *Grokster*. In this case, Veoh’s interactions with and conduct toward its users did not rise to such a level. As Judge Matz recognized, “(a) the allegedly infringing material resided on Veoh’s system; (b) Veoh had the ability to remove such material; (c) Veoh could have implemented, and did implement, filtering systems; and (d) Veoh could have searched for potentially infringing content.” Such circumstances are not equivalent to the activities found to constitute substantial influence in *Cybersett* and *Grokster*. Nor has UMG, in its initial or supplemental briefing to this court, pointed to other evidence raising a genuine issue of material fact as to whether Veoh’s activities involved “something more than the ability to remove or block access to materials posted on a service provider’s Web site.”

Id. at 1030 (citations omitted). Like Perma.cc, the defendant changed the format of the content (video) that it was uploading. “Veoh’s software also automatically converts, or ‘transcodes,’ the video file into Flash 7 format. This is done because ‘the vast majority of internet users have software that can play videos’ in this format. Veoh presets the requisite settings for the Flash conversion.” *Id.* at 1012. Thus, changing formats for storage is not an issue.

²¹⁵*Viacom*, 676 F.3d at 47–48.

The facts in the *Perfect 10* case do not square with Perma.cc, which is not monitoring the “layout, appearance, and content” of submitted content, nor are there any facts to suggest inducement of infringement.

However, there are four possible reasons why Perma.cc might meet this *Viacom* test “of something more” than the right and ability to takedown material, thus demonstrating control. First, it should be remembered that it is Perma.cc that harvests Web materials after being directed to do so by the end user.²¹⁶ The user never uploads the content from his or her computer; instead, Perma.cc “grabs” the content from the Web. Second, Perma.cc makes available the material both in the “capture” format and as a screenshot.²¹⁷ Third, Perma.cc has determined to archive pages with “noarchive” metatags, although with restricted access to users and their organizations.²¹⁸ Fourth, Perma.cc has set in place a system of registrars whose duty is to enforce policies and practices, “ensuring that those in [their] network[s] do not abuse their accounts.”²¹⁹ Certainly, copyright compliance is one of those issues for abuse.²²⁰ Perma.cc is performing actions that look like control. The question is whether in doing so it fails to qualify for general exemption as a service provider (for “information residing on systems or networks at directions of users”)²²¹ from liability by operation of § 512(c)(1)(B).

The facts of Perma.cc’s operation run uncomfortably close to a pre-DMCA case, *Polygram Int’l Publ’g, Inc. v. Nev./TIG, Inc.*, which also turned on the issue of control.²²² *Polygram* can be distinguished as not being an Internet case—specifically, not involving Internet service providers that would fall under § 512(c). However, the *Polygram* case puts to the forefront the central issue: Who should be at risk for copyright violations when an organization, benefiting from infringing copyright, fails to enforce its own rules? The current answer may be that the DMCA §512(c), at least as interpreted in *Viacom*, has fundamentally changed the balance in favor

²¹⁶Perma.cc, *See About Perma.cc*, <https://perma.cc/about#how-perma-works> (“Users go to the Perma.cc Web site and input a URL. Perma.cc downloads the material at that URL and gives back a new URL (a “Perma.cc link”) that can then be inserted in a paper, article, blog or whatever the author needs.”) (last visited May 20, 2020).

²¹⁷See Perma.cc, *Perma.cc User Guide Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#preservation-formats> (last visited May 22, 2020). The “capture” format uses Web ARChive (or “warc”). See Wikipedia, *Web ARChive*, https://en.wikipedia.org/wiki/Web_ARChive (last visited May 22, 2022). It is widely used in Web harvesting and “Web crawls.” *Id.*

²¹⁸See Perma.cc, *Perma.cc User Guide: Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#private-records> (last visited June 9, 2020).

²¹⁹Perma.cc, *Perma.cc User Guide/For Registrars*, <https://perma.cc/docs/libraries#libraries-authority> (last visited June 9, 2020). Registrars can also set their own policies and best practices. *Id.* Registrars and end users can change access of permalinks to private viewing only. See Perma.cc, *Perma.cc User Guide: Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#private-records> (last visited June 9, 2020).

²²⁰See *supra* n. 8.

²²¹See heading to 17 U.S.C. § 512(c) (Westlaw through Pub. L. No. 116-152) and text of (c)(1) (exempting liability).

²²²855 F. Supp. 1314 (D. Mass. 1994).

of online service providers and placed the burden of risk on copyright owners. It should not surprise us then that copyright owners demand changes in the DMCA.²²³

The issue of control is vital to determining Perma.cc's liability under the theory of contributory infringement. It is uncertain at best. However, Perma.cc's ability to remove or block content, by itself, is not enough to establish control and resulting liability.

512(c) and (d) protection from liability and registrars

Perma.cc has been prudent in placing the responsibility on registrars or partner organizations to set policies for its users.²²⁴ The registrars stand between Perma.cc and its end users. However, this arrangement creates questions as to whether registrars, assuming suit under theories of contributory or vicarious liability, can benefit from the protections of § 512(c). For instance, the takedown notices are handled by Perma.cc, not the registrars or end users.²²⁵ Consequently, why would § 512(c) apply to registrars? While registrars may not receive takedown notices, they can act preemptively to preserve protection under § 512(c)(1)(A)(iii) and remove or block access to material they know to be infringing.²²⁶ The question is whether such action is necessary prior to a takedown notice given the results of the *Viacom* litigation.²²⁷

The issue of registrars' status under § 512, and whether they could find shelter under § 512(c), was raised in the Introduction.²²⁸ To conclude the analysis, we must look at § 512(k). Regardless of their role in notice-and-takedown procedures, there is also some question about whether library

²²³See *supra* n. 195 and accompanying text. The EU has sought to rebalance the relationship between online service providers and copyright owners by mandating filtering of uploaded content. See Directive 2019/790 of the European Parliament and the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC, 2019 O.J. (L130) 92, 119, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32019L0790> (hereinafter, "EU Copyright Directive"). However, Nonprofit educational repositories are generally exempt as online content-service providers.

Providers of services, such as not-for-profit online encyclopedias, *not-for-profit educational and scientific repositories*, open source software-developing and-sharing platforms, providers of electronic communications services as defined in Directive (EU) 2018/1972, online marketplaces, business-to-business cloud services and cloud services that *allow users to upload content for their own use*, are not 'online content-sharing service providers' within the meaning of this Directive. *Id.* at 113 (emphasis added).

²²⁴Perma.cc, *Perma.cc User Guide/For Registrars*, <https://perma.cc/docs/libraries#libraries-authority> (last visited May 22, 2020) ("As a registrar, you'll help train and support Perma.cc users within your network and you'll be the main point of contact between your network and the Perma team. That means you are the main point of contact for questions from your Perma users regarding policies and practices.").

²²⁵Perma.cc, *Copyright Policy*, <https://perma.cc/copyright-policy>.

²²⁶Registrars have the capability to set permalinks to "private" access. See Perma.cc, *Perma.cc User Guide: Perma Records & Links*, <https://perma.cc/docs/perma-link-creation#private-records> (last visited June 12, 2020).

²²⁷See *supra* n. 184 through 189 and accompanying text.

²²⁸See *supra* n. 38 through 39 and accompanying text.

registrars are “service providers” that may benefit from the notice-and-take-down provisions of § 512(c):

(k) Definitions—

(1) Service provider.—

(A) As used in subsection (a), the term “service provider” means an entity offering the transmission, routing, or *providing of connections* for digital online communications, between or among points specified by a user, of material of the user’s choosing, without modification to the content of the material as sent or received.

(B) As used in this section, other than subsection (a), the term “service provider” means *a provider of online services or network access*, or the operator of facilities therefor, and includes an entity described in subparagraph (A).²²⁹

Arguably, library registrars partner with and provide Perma.cc “network” access to its Perma.cc users and are thus “service providers,” but it is Perma.cc that is the real network, and whether registrars of networks are treated as service providers under § 512(c) is untested in litigation.

Even assuming registrars qualify, they do not have a role in the notice-and-takedown process. For example, where does a law library registrar place its DMCA takedown notice, as required under § 512 (c)(3)? Is the general university DMCA notice that universities put on all Web pages, including their law library’s pages, sufficient to pick up Perma.cc links in law reviews and journals? Or should such notices be placed in the law reviews and journals themselves (especially electronic formats)? What about the requirement of having a designated DMCA agent under § 512(c)(2)? Do law reviews and journals using Perma.cc links need to list this information for their registrars to gain protection? Running parallel to § 512(c) is subsection (d), which covers *linked* information under the same standards as subparagraph (c), which deals with archived information. Finally, shouldn’t registrar libraries and institutions be required to abide by § 512(i) requirements to “adopt[] and reasonably implement[], and inform[] subscribers and account holders of the service provider’s system or network of, a policy that provides for the termination in appropriate circumstances of subscribers and account holders of the service provider’s system or network who are repeat infringers”?²³⁰

Simply relying on the steps Perma.cc takes may be insufficient to protect registrars. It is unlikely that law libraries, acting as registrars, would even receive a takedown notice under § 512(c)(1)(C), (g). How can they claim shelter under § 512(c) by Perma.cc’s actions, even though they are registrars and listed as “Perma.cc’s Partners”?²³¹ Section 512 simply doesn’t cover the situation. Beyond

²²⁹17 U.S.C.A. § 512(k)(1) (Westlaw through Pub. L. No. 116-152) (emphasis added).

²³⁰17 U.S.C.A. § 512(i) (Westlaw through Pub. L. No. 116-152).

²³¹Perma.cc, *About Perma.cc*, <https://perma.cc/about#perma-partners> (last visited May 22, 2020).

academic law libraries, do courts or other institutional registrars (whether or not academic) meet DMCA § 512 requirements? It may not even be on their radar.

Examining § 512(d), which covers *linking* instead of archiving or storing information, our analysis needs to consider whether linking to Perma.cc (and possibly institutions hosting digital law reviews and journals) may be protected under the same schema as § 512(c). For this analysis we need to apply the following:

(d) Information location tools.—*A service provider* shall not be liable for monetary relief, or, except as provided in subsection (j), for injunctive or other equitable relief, for infringement of copyright by reason of the provider referring or linking users to an online location containing infringing material or infringing activity, by using information location tools, including a directory, index, reference, pointer, or *hypertext link*, if the service provider—[repeat of exceptions for (c)].²³²

Again the definition of *service provider* is key,²³³ but especially within the context of “referring” or “hypertext linking.” It is hard to imagine a publisher of a print law journal with Perma.cc links trying to take shelter under § 512(d) and fitting within that provision. However, many journals are now published in PDF formats online, hosted at their school’s Web site and elsewhere.²³⁴ Some have pledged to publish exclusively in open access, *digital* formats.²³⁵ Returning to Perma.cc, the perhaps unlikely issue is: By making available law reviews and journals on their Web sites and using Perma.cc, are host institutions exposing themselves to liability for failing to respond to takedown notices for Perma.cc links to infringing copies of archived materials? Furthermore, do host institutions risk liability when they have actual or “red flag” knowledge of links to infringing material within these law reviews and journals?

These may seem remote risks, but concurrent with writing this article, I had another article being published by AALL’s *Law Library Journal*. In the page proofs the editors, without my asking, added 50 Perma.cc links and made Perma.cc archival copies all of my Web-linked citations. These citations were almost entirely journal articles and book chapters from other fields on natural language processing and artificial intelligence. I found this use of Perma.cc to be risky. In the end, all of the Perma.cc links were removed, but had we gone to press with the Perma.cc links, would AALL be subject to takedown notices under § 512(d)(3) because it publishes its journal electronically on its Web site (and by doing so becomes a *service provider* under § 512(d))? In connection with takedown notices, would the electronic journal be required to appoint a

²³²17 U.S.C.A. § 512(d) (Westlaw through Pub. L. No. 116-152) (emphasis added).

²³³See *supra* n. 232 and accompanying text.

²³⁴See, e.g., Law Review Commons, <https://lawreviewcommons.com/> (last visited Aug. 13, 2020) (listing schools with law reviews and journals in digital PDF formats).

²³⁵Berkman Klein Center for Internet & Society at Harvard University, Durham Statement on Open Access to Legal Scholarship, Feb. 11, 2009, <https://cyber.harvard.edu/publications/durhamstatement>.

designated agent for receipt of such notices?²³⁶ Furthermore, regardless of answering takedown notices, would AALL's *Law Library Journal* be responsible for the actual or red flag knowledge²³⁷ of its editors who by linking to archived copies of other journal articles and book chapters invoked actual or red flag knowledge? Indeed, by doing so, they were in fact²³⁸ linking to infringing copyrighted articles, and hence they would be unable to find shelter under § 512 (d) due to (d)(1)(A) and (B). Would they be responsible under a theory of contributory infringement for controlling the editing process that added the links and for profiting from journal subscriptions without shelter under § 512 due to (d)(2)? All of this is untested in court and speculative, but I am describing risks that find their basis in existing law and in plausible and occurrent facts.

Assuming law libraries and other registrars lack cover under § 512 (for instance, if courts find that registrars are not *service providers*), they may need to address case law doctrines of vicarious liability, contributory infringement, and even willful blindness.²³⁹ Without § 512(m), it will be harder to argue they have no duty to monitor, or that monitoring by libraries does not trigger liability as a kind of control under vicarious liability or knowledge under contributory infringement. Even the staple doctrine of "substantial non-infringing uses" is not of much use because Web material is copied while such content is still available on the chance that someday it might be unavailable, at least at its original location. This is not exactly the "time-shifting" or authorized use justifying the doctrinal defense in *Sony v. Universal City Studios*.²⁴⁰ It is time-shifting on the contingency that someday the Perma.cc-linked material might one day not be available. The specter of *Polygram* raises its head,²⁴¹ with the possibility of courts assigning risk to those who failed to regulate usage of copyrighted works. This lack of solid protection under *Sony* should concern end users and registrars alike.

²³⁶See 17 U.S.C.A. § 512(c)(2) (Westlaw through Pub. L. No. 116-152). Perhaps the journal could publish Perma.cc's designated agent notice. See Perma.cc, *Copyright Policy*, (Sept. 23, 2013), <https://perma.cc/copyright-policy>.

The undersigned believe that it will benefit legal education and improve the dissemination of legal scholarly information if law schools commit to making the legal scholarship they publish available in stable, open, digital formats in place of print. To accomplish this end, law schools should commit to making agreed-upon stable, open, digital formats, rather than print, the preferable formats for legal scholarship.

Id. The original signatories included University of Chicago, Columbia University, Cornell University, Duke University, Georgetown University, Harvard University, New York University, Northwestern University, the University of Pennsylvania, Stanford University, the University of Texas, and Yale University. *Id.*

²³⁷For review of actual and red flag knowledge under the analogous § 512 (c), see *supra* n. 174–178 and accompanying text.

²³⁸See 17 U.S.C.A. § 512 (d)(1)(A) and (B) (Westlaw through Pub. L. No. 116-152).

²³⁹For discussion of willful blindness, see *supra* n. 177.

²⁴⁰See *supra* n. 52 through 53 and accompanying text.

²⁴¹See *Polygram Int'l Publ'g, Inc. v. Nevada/TIG, Inc.*, 855 F. Supp. 1314 (D. Mass. June 20, 1994). See also *supra* n. 225 and accompanying text.

Frustration with 512 and the CASE Act

Copyright owners were quite discontented with the DMCA takedown system²⁴² and wanted, and eventually got, reforms. A bill known as the “CASE Act” was amalgamated into the *Consolidated Appropriations Act, 2021*.²⁴³ The legislation sets up a small claims board for copyright infringement. The purpose of the CASE Act is to

ensure[] that copyright interests without high expected damages have some mode of enforceability so that individual creators, “many of whom rely upon the promise of exclusive rights associated with the grant of copyright to earn a living and provide for their families[,] ... have a realistic ability to enforce those rights when they have a comparatively modest claim for damages.”²⁴⁴

The CASE Act might spawn many more claims to be brought for infringing content in Web archives because the cost of litigation would be significantly less.

However, under § 1506(aa) of the Act, libraries and archives may, without charge and following yet-to-be promulgated Register of Copyright regulations, “preemptively opt out of” proceedings before the Board by getting on a list with the Register of Copyrights.²⁴⁵ Just when

²⁴²See *supra* n. 195 and accompanying text.

²⁴³See Consolidated Appropriations Act, 2021, H.R. 133, 116th Cong. (2020) (enacted), <https://www.congress.gov/bill/116th-congress/house-bill/133/>. Title II, Subchapter A, Chapter 15—Copyright Small Claims—of the Act, including §§ 1501–1511 covers the creation, claims, and procedures for a Copyright Claims Board.

²⁴⁴H. Rep. No. 116-152, at 19 (2019) (citing Letter from Lamar Smith, Chairman, H. Comm. on the Judiciary, to Maria A. Pallante, Register of Copyrights and Director, U.S. Copyright Office (Oct. 11, 2011), included in U.S. Copyright Office, Copyright Small Claims: A Report of the Register of Copyrights (2013), <https://www.copyright.gov/docs/smallclaims/usco-smallcopyrightclaims.pdf>). Apparently, the ABA supports the legislation, which will help “small creators.”

On the one hand, you have the notice and takedown process that can be inefficient, cumbersome, and, as many small creators will tell you, often pointless. On the other hand, there is the Article III Federal court system that can be expensive, time-consuming, and often out of reach for many working-class and middle-class creators.

For instance, the average cost of litigating an infringement case in Federal court is approximately \$350,000, but the total amount of damages that can be awarded, for instance, in a CASE Act-eligible matter cannot exceed \$30,000. In that instance, the cost of litigating a case could be more than 10 times the damages that are at issue.

According to a survey by the American Bar Association, which supports this legislation, most lawyers will not take infringement cases with damages at or lower than \$30,000. As a result, many petitioners are functionally unable to vindicate their rights under law. In other words, these creators are given a right without a remedy.

The CASE Act will provide a viable alternative. This legislation would establish a voluntary forum for small copyright claims housed within the Copyright Office. Disputes would be heard by a new entity called the Copyright Claims Board made up of intellectual property experts with experience representing both creators and the users of copyrighted material.

¹⁶⁵Cong. Rec. H8352 (daily ed. Oct. 22, 2019) (statement of Rep. Hakeem Jeffries), <https://www.congress.gov/116/crec/2019/10/22/CREC-2019-10-22-house.pdf>.

²⁴⁵See Consolidated Appropriations Act, 2021, H.R. 133, 116th Cong., Title II, Subtitle A, § 1506(aa) (2020) (enacted), <https://www.congress.gov/bill/116th-congress/house-bill/133/>. “Library” and ‘archives’ mean any library or archives respectively that qualifies for limitation on exclusive rights under section 108.” *Id.* at § 1506 (aa)(4). The preemption option for libraries was not in the earlier bills of the CASE Act. See Copyright Alternative in Small-Claims Enforcement Act of 2019, H.R. 2426, 116th Cong. (2019) and

such preemption has to take place is undefined, but it appears that since the Register of Copyrights is to “maintain a publicly available list” of the libraries and archives that have “preempted out,” then the pre-emption might be something that has to be done on a one-time or periodic basis in advance of proceedings.²⁴⁶ Librarians and archivists need to think carefully (perhaps with aid of counsel) before electing such a preemption for their libraries and archives about whether the long-term risks make the preemption worthwhile. Essentially, there is a tradeoff: The Copyright Small Claims Board makes it less costly to sue (encouraging suits), but it also limits liability. Even if a library or archive does not “preempt out” via the Registrar’s list, it would be entitled to opt out of suit (participation is voluntary)²⁴⁷ and have the suit dismissed without prejudice within 60 days of notice of filing.²⁴⁸ In addition to library and archive preemptions, claims against federal and state agencies are also excluded.²⁴⁹ Participation is voluntary, and rights to pursue claims in other forums are preserved. Total damages are limited to \$30,000 (exclusive of attorney’s fees).²⁵⁰ Certainly, the adoption of this Act is worth considering for those participating as Perma.cc registrars because it lowers the economic hurdles to suit while at the same time granting legal procedures for exclusion for libraries and archives. This may stimulate reform of practices and policies pertaining to Perma.cc.

Factors registrars and end users might want to consider to protect themselves from copyright liability

Because we may not apply § 512(c) with any certainty to protect registrars, and because they are contractually bound to indemnify Perma.cc for liability from use by their end users of the system, thoughtful consideration of policy and practices for administering Perma.cc usage among law journals, faculty, and students is imperative.

The most important step is to convey the sense of responsibility for copyright compliance that students, librarians, and faculty have in the use of Perma.cc. Even if an academic law library does not subscribe to

²⁴⁶Copyright Alternative in Small-Claims Enforcement Act of 2019, S. 1273, 116th Cong. (2019) (lacking any reference to libraries or archives).

²⁴⁷See *id.*, at § 1506(i) (2020) (enacted), <https://www.congress.gov/bill/116th-congress/house-bill/133>.

²⁴⁸See *id.*, at § 1504(a).

²⁴⁹See *id.*, at § 1504(e)(1)(D). Damages are limited to \$15,000 for each infringed work (assuming the registration is “timely filed”). See *id.*, at § 1504(e)(1)(A)(ii)(I).

²⁵⁰See *id.*, at § 1504(e)(1)(B). However, a successful claim brought against Perma.cc for infringement may trigger indemnification obligations for a law library of a state university. See *supra* section “User and Library Obligations and Indemnification.”

Perma.cc, its law reviews and journals may wish to publish works of authors who use them. Consequently, there is a need for broad discussion on this topic throughout the profession. Whether there should be a wholesale prescriptive ban on classes of works is another question. Indeed, it might be easier to start with what is safe—federal government Web sites and documents and instances where the end user has received permission. However, for anyone who has ever asked for it, permission is difficult, time consuming, and will subject publications with Perma.cc links to delay (that is probably why Perma.cc does not insist upon it). Such a policy would significantly hamper the utility of Perma.cc as a service. Essentially, users or registrars need to complete a fair use analysis to fully evaluate compliance with copyright law²⁵¹—not something that should lightly be entrusted to the disparate members of an organization, including student editors, when the registrar may be liable on theories of contributory infringement or indirect liability (or under Perma.cc's indemnification clause). On the other hand, the specter of infringement may result in little more than a takedown notice, which library registrars may deem an acceptable risk. However, this is an odd position for *law* libraries to take—we will accept infringement because the risk is only a rare takedown notice. This seems to be at odds with what should be our fundamental values and the message we convey to student editors. Assuming a library wanted to avoid infringement (and attendant risks, rare though they may be), then for the sake of a few bright line rules to ensure copyright compliance, whole categories of materials—for instance, Web blogs, news sites, movies, academic articles, and audio/visual materials—might be delegated to preservation in Perma.cc's *private* setting. On the other hand, electing *private* archiving diminishes the value of Perma.cc for scholars.

Library registrars should determine if the role of Perma.cc citations in footnotes is to provide complete access to all readers to Web materials or if it is to affirm that the law review (or author) has done its job in verifying the content cited and consequently *privately* archiving the item on Perma.cc is enough. At least that way, journal editors or authors could verify the accuracy of a citation.²⁵² Anyone wishing to access the cited content in its entirety would have to check subscription databases or archives for the material. Some material may just not be made available to the public, and it may have to be enough to know that it was once accessed by editors and authors to verify a proposition—something journals may continue to do by accessing a Web material archived under a *private* setting.

²⁵¹See *supra* sections “Fair Use and ‘Transformative Use’” and “Fair Use Analysis without Transformative Use.”

²⁵²This is similar to the Harvard Blue Book Rule 17.2.1 allowing the parenthetical “(on file with author”).

In the final analysis, some might return to point to the efficacy of the notice-and-takedown system of § 512(c)(1)(C) as the ultimate barrier between registrars and liability and as the reality of the risks associated with using Perma.cc. But as pointed out in a previous section,²⁵³ it is Perma.cc that fits squarely within this section—not registrars or end users. Nonetheless, to preserve the argument that §512(c) protects them, registrars may want to preemptively deal with infringing content by switching links to the *private* archival setting. Alternatively, § 512(d) pertaining to links may apply (especially to online journals). To preserve protections under that subparagraph, registrars need to plan to deal with infringing content. Indeed, § 512(i) obliges Internet service providers to adopt and notify end users of certain policies, including cancelation of accounts for abuse.²⁵⁴ Online journals using Perma.cc may also need to provide notice of a designated agent for takedown notice (perhaps the same as Perma.cc's).²⁵⁵ Finally, future legislation may make suits against Perma.cc or its registrars more plausible.²⁵⁶

Conclusion

When I commenced writing this article, I presumed that, despite an obligation to objectivity, as a librarian and law professor, my own inclinations would be to justify preservation activities of Perma.cc and its network of mostly library registrars to stop link rot. However, I now wonder whether the scholarly footnote's function (even in using Perma.cc) is simply limited to supporting scholarly propositions, a task that can be done by checking citations, and in most cases, preserving only *private* versions of Web content in Perma.cc archives. I realize this may run counter to our basic instincts as librarians and scholars—to provide access and to document the ever-shifting sands of materials on the Web.

Without Congressional mandate to do so—without amending 17. U.S.C. § 108 to permit libraries and archives to capture that which they do not own²⁵⁷—I have come to believe that we need to respect, however small or

²⁵³See *supra* section “§ 512(c) and (d) Protection from Liability and Registrars.”

²⁵⁴See *supra* n. 233 and accompanying text.

²⁵⁵See *supra* n. 239.

²⁵⁶See *supra* section “Frustration with “§ 512 and the CASE Act.”

²⁵⁷Although I cannot document it, as Chair of the Copyright Committee of the American Association of Law Libraries in 2005–2006, I participated in a Copyright Office sponsored round table to discuss amendment of § 108 to permit libraries and archives to archive the web. The proposal was not well received by content-creating industries, who wanted concessions of their own—that for-profit libraries no longer receive the benefits of § 108. Ultimately, no progress was made. Certainly, some will point to the continued success of the Internet Archive in archival efforts to prove such changes turned out to be unnecessary, but I am more skeptical. See *supra* n. 91 through 93 and accompanying text.

In 2011, the Copyright Office produced a report that remarked that “Section 108 exception does not contemplate mass digitization.” U.S. Copyright Office, *Legal Issues in Mass Digitization: A Preliminary Analysis and Discussion Document* 19-20 (Oct. 2011), https://www.copyright.gov/docs/massdigitization/USCOMassDigitization_

even uncontested, the rights of the Web's diverse copyright owners absent their permission to do otherwise. In general, there is no clear path to fair use in Perma.cc's application to Web material²⁵⁸—and in particular there is no transformative use.²⁵⁹ It is not just a question of risk (which right now appears to be slight) and escaping liability. It is about recognizing that current law provides no cover for Perma.cc archiving outside of works in the public domain and authorized uses. Certainly, I would approve of changes in the law to support Web archiving of Web materials, but a proper limiting principle (so that all copying in the name of archiving is not fair use) is elusive. I also encourage technological advances that would permit users of Perma.cc to *select just the portion they need to support a scholarly proposition or argument*—for instance, selective screen capture.

Because legislation, like the CASE Act—as a part of the Consolidated Appropriation Act, 2021²⁶⁰—makes it easier for authors and creators to sue instead of simply issuing takedown requests, libraries and archives need to become familiar with and make decisions about the yet-to-be promulgated regulations of the Registrar of Copyrights to preempt themselves from jurisdiction of the Small Claims Board under the Act. However, even with such an exemption, there is dissonance between online archiving practices and copyright law.

Unfortunately, I probably have reached what will be an unpopular conclusion in the community of my respected colleagues, but fortunately I have not uncovered huge copyright risk. Rather, I have uncovered a fundamental lack of accord between copyright law and online archiving, as in the instance of Perma.cc.

Acknowledgments

The author wishes to give special thanks to Kim Kirschenbaum for her editing assistance. This article was prepared in anticipation of the Yale Citation and Law Conference to be held in April of 2021. I am indebted to The 2020 All Zoom Boulder Conference on Legal Information: Scholarship and Teaching for their constructive comments and criticisms. I am also grateful to the Little Boulder Citations Workshop, held in January 2021, for reviewing a revised draft of this article. I am especially beholden to Sarah Lamdan for leading the review of my work at both conferences and for her numerous helpful edits and suggestions.

e context, however, was the mass digitization of books, and in particular the Google books digitization litigation. See *id.* at i, 19–20. See also, Copyright.gov, *Mass Digitization Pilot Program*, <https://www.copyright.gov/policy/massdigitization/> (last visited June 6, 2020).

²⁵⁸See *supra* section “Summation of Fair Use Factors.”

²⁵⁹See *supra* section “Fair Use and ‘Transformative Use.’”

²⁶⁰See *supra* n. 250 through 255 and accompanying text.