2011

The Best Available Technology Standard

Lital Helman
Columbia University

Gideon Parchomovsky
University of Pennsylvania

THE BEST AVAILABLE TECHNOLOGY STANDARD

Lital Helman and Gideon Parchomovsky***

Copyright liability for webhosting will be a key determinant of the evolution of the Internet in years to come. Depending on their design, the legal rules that shape the liability of webhosts can stunt the development of the Internet as a medium of expression or enhance it. Hence, adopting the optimal liability regime is a matter of crucial importance.

This Article proposes a radical change in webhosts' copyright liability for illegal content posted by users. Our main thesis is that webhosts' liability should be guided by the "Best Available Technology" principle, according to which webhosts that employ the best filtering technology available on the market will be immune from liability for copyright infringement.

Adoption of our proposed liability regime would offer several key advantages relative to the extant regime. First, it would provide webhosts with the certainty they need to continue to operate and grow. Second, it would result in superior and more balanced enforcement of the rights of copyright holders and would achieve this at a much lower cost. Third, it would spur competition in the market for filtering technology and induce constant improvement in enforcement technology. Fourth, and finally, it would dramatically reduce the rate of copyright infringement suits against website operators and the cost of adjudicating them.

We further demonstrate that the analytical framework we construct represents a superior approach not only to the liability of webhosts, but also to those of file sharing and possibly Internet Service Provider liability.

* Fellow, the Kernochan Center For Law and the Arts, Columbia Law School.

** Professor, University of Pennsylvania Law School and Bar-Ilan University Faculty of Law. This Article greatly benefited from comments and criticisms by Jane Ginsburg, Peter Siegelman, Tim Wu, John Duffy, Scott Kieff, Michael Abramowicz, Robert Brauneis, John Whealan, Peter Menell, James Grimmelmann, Felix Wu, Brett Frischmann, Edward Lee, Eva Subotnik, and Martin Mois, as well as the participants of the 2011 Intellectual Property Scholars Conference, seminar participants at George Washington School of Law, the participants of the IP and Information Law Colloquium at Cardozo School of Law, the participants of the Associates' and Fellows' Workshop at Columbia Law School, and the participants of the Faculty Workshop at Fordham. We are grateful to Sam Hartzell for excellent research assistance.

Introduction

Copyright liability for webhosting will likely determine the evolution of the Internet as a medium of expression in years to come. As user generated content (UGC) continues to grow, so do the implications of imposing liability on host sites, social networks, and various online forums and blog platforms.[1] Depending on their design, the legal rules that shape the liability of webhosts can stunt the development of the Internet as a medium of expression or facilitate it. Hence, adopting the optimal liability regime is a matter of crucial importance.

This challenge did not escape the attention of Congress. In 1998, Congress addressed this issue in the Digital Millennium Copyright Act (DMCA).[2] Not surprisingly, the congressional solution represented a compromise between the demands of the content industries to impose liability on internet intermediaries and the pleas of the internet indus-

1. The popularity of UGC on websites has increased tremendously in recent years. This includes blogs, social networks (such as Facebook and Twitter), photo and video sharing sites (such as Flickr and YouTube), news sites (such as Slashdot, Groklaw, and network news sites that seek viewer input, such as CNN.com), collaborative content sites, or "wikis" (such as Wikipedia), games and virtual worlds (such as Second Life), and consumer review websites (such as Yelp and Angie's List). UGC has become so popular that companies use UGC for advertising, see, for example, Doritos and Pepsi MAX Present: Crash the Super Bowl, http://www.crashthesuperbowl.com/#/contest-info (on file with the *Columbia Law Review*) (last visited Sept. 10, 2011), and commercial websites solicit customer reviews (such as Amazon.com and TripAdvisor).

2. Pub. L. No. 105-304, 112 Stat. 2860 (1998) (codified as amended in scattered sections of 5, 17, 28, and 35 U.S.C.).

tries to afford them sufficient breathing room to operate and grow.[3] Congress chose to reconcile the conflicting demands by adopting a two-pronged approach consisting of both ex ante and ex post measures. Ex ante, Congress exempted qualifying webhosts from the duty to detect infringements, save in the most egregious cases in which a "red flag" hangs over content,[4] and placed the burden of detecting infringements on the shoulders of copyright owners.[5] Ex post, after infringing content was identified by copyright owners and upon receipt of notice to this effect, webhosts were required by Congress to remove the offensive content expeditiously.[6] This approach is widely known as the "notice and takedown" procedure.[7]

Congress formalized this arrangement in section 512(c) of the Copyright Act. The section established a safe harbor that shields webhosts from liability for monetary damages resulting from infringement committed by users, provided that they do not interfere with the content posted by users, and as long as they adopt and reasonably implement a system for removing infringing content at the request of copyright owners.[8] Webhosts that do not qualify for the section 512(c) safe harbor are exposed to copyright liability for hosting the infringing content and to the full array of copyright damages.[9]

In this Article, we propose a radically different approach to websites' liability for infringing content posted by users, which we term "the best

---

3. In re Charter Commc'ns, Inc., Subpoena Enforcement Matter, 393 F.3d 771, 774 (8th Cir. 2005) ("It was designed to strike a balance between the interests of ISPs in avoiding liability for infringing use of their services and the interest of copyright owners in protecting their intellectual property and minimizing online piracy."); see also Niva Elkin-Koren, Making Technology Visible: Liability of Internet Service Providers for Peer-to-Peer Traffic, 9 N.Y.U. J. Legis. & Pub. Pol'y 15, 28 (2005) (noting DMCA regime "reflects a compromise between the demands of copyright owners . . . and the concerns of the Internet industry"); Jessica Litman, The Politics of Intellectual Property, 27 Cardozo Arts & Ent. L.J. 313, 314 (2009) (describing how Congress relied on lobbyists for copyright-affected industries "to get together and negotiate the language for any new copyright legislation").

4. 17 U.S.C. § 512(c)(1)(A)(ii) (2006).

5. See, e.g., Perfect 10, Inc. v. CCBill LLC, 488 F.3d 1102, 1113 (9th Cir. 2007) ("The DMCA notification procedures place the burden of policing copyright infringement—identifying the potentially infringing material and adequately documenting infringement—squarely on the owners of the copyright."); UMG Recordings, Inc. v. Veoh Networks Inc., 665 F. Supp. 2d 1099, 1108 (C.D. Cal. 2009) ("[T]he burden is on the copyright holder to provide notice of allegedly infringing material . . . .").

6. 17 U.S.C. § 512(c)(1)(C).

7. For the original, judicially created takedown regime, see Religious Tech. Ctr. v. Netcom On-Line Commc'n Servs., Inc., 907 F. Supp. 1361, 1375 (N.D. Cal. 1995) (holding webhost faces contributory liability where it continues to publicly distribute infringing material after being notified of infringement).

8. 17 U.S.C. § 512(c)(1)(C). Qualifications include § 512(c)(1)(B) (stating that webhost "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") and § 512(c)(1)(C) (describing the notice and takedown policy for infringing content).

9. See id. §§ 502–505 (stating remedies for copyright infringement).

available technology safe harbor." Under our proposal, websites will be immune from copyright liability—both direct and indirect—as long as they can show that they employed the best filtering technology available on the market at the time the alleged infringement occurred.[10]

The technological safe harbor we envision is a dynamic one; it will change with technological progress, creating opportunities for technology companies to develop superior filtering technologies and spurring websites to adopt them. Because the best available technology is an adaptive standard, we expect it to underwrite constant competition in the market for filtering technologies by guaranteeing successful improvers of existing standards that their improvements will be adopted by webhosts who seek shelter under the wings of the technological safe harbor.

Hence, our legal standard would have beneficial effects on two levels: First, it would serve as a constant spur for technology companies to improve existing filtering technologies toward more accurate identification and removal of infringing content on the one hand, and retention of noninfringing content on the other. Second, it would induce websites to implement the technologies, for failure to do so might expose them to liability.

Finally, to clarify, the best available technology standard is not intended to supplant the notice and takedown procedure, ex post. Content owners who learn of the presence of infringing content on a particular website will continue to have the option of notifying the website's operator and demanding its removal. Likewise, our proposal retains the "reverse notice and put-back procedure," which enables users to challenge unwarranted takedowns and demand that their content be restored.[11]

This Article proceeds as follows. In Part I, we explore the inefficiencies of the current liability standard for webhosts under the DMCA and the common law doctrines for direct and secondary liability. In Part II, we present our proposal for a new technological safe harbor and explain its virtues. In Part III, we deal with potential challenges and objections to our proposal. In Part IV, we examine the possibility of expanding our proposed safe harbor to areas other than webhost liability, particularly peer-to-peer services and access providers. A short conclusion ensues.

I. The Current Liability Standard for Webhosts

The last two decades saw a substantial growth in online platforms. In particular, the development of new technologies and business models allowed members of the general public to post content on websites and share it with others.[12] While the emergence of platforms dedicated to user content has generated an upsurge in creativity and originality, the

10. The best available technology would be one that minimizes the total number of false positives and false negatives. For elaboration and discussion, see infra Part II.B.

11. 17 U.S.C. § 512(g)(2).

12. See supra note 1 (discussing types of user generated content).

materials posted by users are not always created by them, and in many cases, the posting occurs without permission from the copyright owner.[13] In the absence of fair use or some other defense, these cases almost certainly constitute a copyright infringement.[14]

The more interesting question is, who is liable for the infringement? It is clear as a matter of positive law that the posting individuals bear direct liability for their actions. Direct liability arises whenever an individual reproduces, adapts, distributes, publicly displays, or performs a copyrighted work, or digitally performs a sound recording, without authorization from the copyright owner.[15] Consequently, the unauthorized posting of copyrighted content constitutes an infringement of the exclusive rights to reproduce, distribute, publicly display, or perform the content, and may also amount to a violation of the owner's adaptation right.[16]

A more challenging question is whether liability also attaches to the operator of the hosting site or platform. In principle, webhosts may be subject to either direct or indirect liability in these cases. As far as direct liability is concerned, webhosts may bear liability for distributing, performing, reproducing, and displaying content, and in some cases, for adapting it.[17] In addition, website operators are exposed to indirect liability. Indirect liability comes in two varieties: contributory liability and vicarious liability.[18] Contributory liability arises when one, with knowledge, induces, enables, facilitates, or materially contributes to an infringing activity by another.[19] Vicarious liability attaches when one has the ability to

---

13. See, e.g., Viacom Int'l Inc. v. YouTube, Inc., 718 F. Supp. 2d 514, 518–19 (S.D.N.Y. 2010) (quoting plaintiff's claim that "tens of thousands of videos on YouTube, resulting in hundreds of millions of views, were taken unlawfully from Viacom's copyrighted works without authorization").

14. Posting of works online for wide access typically comes under one or more of the exclusive rights of copyright owners enshrined in 17 U.S.C. § 106. Thus, for example, § 106(1) covers reproduction, § 106(2) covers derivative works, § 106(3) covers distribution, § 106(4) and § 106(6) cover public performance of several types of works, and § 106(5) covers the right to display several types of works publicly.

15. 17 U.S.C. § 106. Liability is of course subject to defenses, such as fair use. Id. § 107.

16. Id. § 106.

17. Id.

18. Some commentators view inducement theory as a third, independent secondary liability doctrine. Whether inducement is an independent source for derivative liability or part of contributory liability is unclear. Compare Charles W. Adams, Indirect Infringement from a Tort Law Perspective, 42 U. Rich. L. Rev. 635, 636 (2008) ("Copyright law has three separate doctrines for third-party liability: vicarious infringement; contributory infringement; and inducing infringement."), with Perfect 10, Inc. v. Amazon.com, Inc., 487 F.3d 701, 725 (9th Cir. 2007) (describing inducement as one of two categories of contributory liability (citing Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 545 U.S. 913, 942 (2005) (Ginsburg, J., concurring))).

19. Gershwin Publ'g Corp. v. Columbia Artists Mgmt., 443 F.2d 1159, 1162 (2d Cir. 1971).

control or oversee an infringement and has a financial interest in it.[20] The scope of these doctrines is so broad that in principle indirect liability could attach to most active websites.[21]

The conflict between the interest of copyright owners in enhancing enforcement by extending liability to operators of online services, and the interest of the internet industries in minimizing their exposure to liability prompted Congress to intervene.[22] In 1998, Congress enacted the DMCA, which represented a compromise between copyright owners and the internet industries.[23] For our purpose, the most important change effected by the new legislation was embodied in Title II of the DMCA, subsequently codified as section 512 of the Copyright Act.[24] Section 512 establishes four safe harbors that shelter four central activities of online service providers (OSPs)[25] from monetary liability,[26] provided that they comply with certain threshold criteria.[27] Most notably, to avail themselves of the safe harbors, OSPs must meet two general conditions: They must accommodate standard technical measures content owners use to protect

---

20. See Shapiro, Bernstein & Co. v. H.L. Green Co., 316 F.2d 304, 307 (2d Cir. 1963) ("When the right and ability to supervise coalesce with an obvious and direct financial interest in the exploitation of copyrighted materials . . . the purposes of copyright law may be best effectuated by the imposition of liability upon the beneficiary of that exploitation.").

21. See Lital Helman, Pull Too Hard and the Rope May Break: On the Secondary Liability of Technology Providers for Copyright Infringement, 19 Tex. Intell. Prop. L.J. 111, 114–19 (2010) [hereinafter Helman, Secondary Liability] (discussing how doctrines of indirect liability have expanded over time). Courts also often obscure the differences between the doctrines, thus increasing the risk of liability. See id. at 116–17 ("The danger in this confusion . . . is that lawmakers will bypass the phase of proving all the elements of a particular doctrine, and will subject defendants to liability based on some assortment of standards from these doctrines, thus subjecting a broader class of activities to liability than originally intended."); see also Douglas Lichtman & William Landes, Indirect Liability for Copyright Infringement: An Economic Perspective, 16 Harv. J.L. & Tech. 395, 404 (2003) ("[The *Napster* court's] analysis seems to blur the line between the requirement under contributory infringement that a culpable party have knowledge of the direct infringement and the requirement under vicarious liability that a culpable party have control over the specific infringer." (discussing A&M Records, Inc. v. Napster, Inc., 284 F.3d 1091 (9th Cir. 2002))).

22. The telecommunication industry was in fact active in lobbying against copyright owners in this debate. For a description of the creation of the DMCA, see, for example, Timothy Wu, Copyright's Communications Policy, 103 Mich. L. Rev. 278, 350–56 (2004).

23. See supra note 3 (describing compromise in DMCA).

24. 17 U.S.C. § 512 (2006).

25. As defined in § 512(k)(1).

26. Section 512(a) protects services which provide internet access. Section 512(b) shields against liability for temporarily storing infringing materials. Section 512(c) protects hosting services, and section 512(d) applies to "information location tools," such as links to content on other sites. Qualifying OSPs are exposed to a limited injunctive relief under section 512(j).

27. The threshold requirements are fixed in 17 U.S.C. § 512(i) and in the specific safe harbors, § 512 (a)–(d).

their materials, and they must adopt and reasonably implement a policy for terminating infringing users.[28]

Of the four safe harbors, the one pertinent to our analysis is section 512(c), which covers hosting activities "at the direction of a user."[29] The safe harbor specified in section 512(c) purports to relieve webhosts of monetary damages for both direct and indirect liability.[30]

Effectively, section 512 expressly exempts webhosts from monitoring user generated content ex ante;[31] such monitoring is limited to "red flag" cases.[32] The main obligation the section imposes on webhosts is to remove offending content expeditiously, ex post, upon receiving notification from a copyright owner.[33]

The design of the current safe harbor for webhosting has several shortcomings. First, the current regime provides webhosts with almost no incentive to participate in copyright enforcement. Regardless of one's view on the optimal level of copyright enforcement, this lack of incentive creates an inefficient and wasteful enforcement scheme. Second, the present legal regime does not afford sufficient certainty for webhosts. Third, and finally, the current enforcement model runs the risk of curtailing too much lawful speech, because copyright owners control enforcement and because of private ordering that has emerged in the shadow of the law. In the proceeding discussion we elaborate on each of these problems.

A. *Lack of Incentives for Webhosts to Participate in Copyright Enforcement*

As we noted earlier, with the exception of few cases, section 512 does not require webhosts to monitor content on their site ex ante as a prerequisite for enjoying the safe harbor.[34] The section puts a premium on the ex post phase: Upon being notified of the presence of infringing content on their system, webhosts are required to remove the offending content

---

28. 17 U.S.C. § 512(i)(1).

29. Id. § 512(c). The limitation of liability includes storage and allied functions, such as providing access to the materials. See, e.g., Io Grp., Inc. v. Veoh Networks, Inc., 586 F. Supp. 2d 1132, 1146–47 (N.D. Cal. 2008) (allowing safe harbor protection for automated functions that facilitate access to content).

30. See H.R. Rep. No. 105-551, pt. 2, at 50 (1998) ("The limitations in subsections (a) through (d) protect qualifying service providers from liability for all monetary relief for direct, vicarious and contributory infringement."); S. Rep. No. 105-190, at 40 (1998) (same). But see infra note 66 and accompanying text (noting some commentators have suggested section 512's safe harbor is effective only against direct copyright infringement).

31. 17 U.S.C. § 512(m) ("Nothing in this section shall be construed to condition the applicability of subsections (a) through (d) on (1) a service provider monitoring its service or affirmatively seeking facts indicating infringing activity, except to the extent consistent with a standard technical measure complying with the provisions of subsection (i) . . . .").

32. See infra notes 37–45 and accompanying text.

33. See 17 U.S.C. § 512(c)(1)(C) (creating safe harbor for service providers that remove purportedly infringing content).

34. Id. § 512(m).

and, in appropriate cases, terminate the user who uploaded the content to the system.[35]

Absent notification from a content owner, webhosts are expected to take independent action *only* if they have *actual* knowledge of infringement,[36] or are "aware of facts or circumstances from which infringing activity is apparent,"[37] known as the "red flag" test.[38] Otherwise, webhosts' best strategy is to remain passive and abstain from taking any affirmative measures not required by the section. Indeed, it would be irrational for profit-maximizing enterprises, such as many webhosts, to incur the extra cost of deploying affirmative measures to protect third-party rights when immunity is guaranteed irrespective of such measures.[39]

Importantly, "red flag" cases have been construed very narrowly by the courts.[40] For example, the Ninth Circuit opined that provision of services to websites with names such as "illegal.net" and "stolencelebritypics.com" is insufficient to raise a "red flag."[41] Other courts concluded that third party takedown notices do not amount to a "red flag,"[42] nor do imprecise notices of infringement suffice for that purpose.[43] Most recently, a court held in *Viacom v. YouTube* that "knowledge of prevalence of [infringing] activity in general is not enough," and that generally, "if investigation of 'facts and circumstances' is required to identify materials as infringing, then those facts and circumstances are not

---

35. Id. § 512(c)(1)(C) (removal of content); id. § 512(i)(1)(A) (termination of repeat infringers). Abstaining from removing the infringing files would cost the webhost its protected status under section 512. See id. § 512(c)(1)(C).

36. Id. § 512(c)(1)(A)(i).

37. Id. § 512(c)(1)(A)(ii).

38. H.R. Rep. No. 105-551, pt. 2, at 53, 57 (1998) ("[A] service provider need not monitor its service or affirmatively seek facts indicating infringing activity . . . to claim this limitation on liability . . . . However, if the service provider becomes aware of a 'red flag' from which infringing activity is apparent, it will lose the limitation of liability if it takes no action."); see also Perfect 10, Inc. v. CCBill LLC, 481 F.3d 751, 763 (9th Cir.) (describing "red flag" test), modified, 488 F.3d 1102, 1113–14 (9th Cir. 2007).

39. See infra Part I.C for agreements between webhosts and content owners where webhosts are willing to perform extra measures in return for legal certainty.

40. Jane C. Ginsburg, User-Generated Content Sites and Section 512 of the US Copyright Act, *in* Copyright Enforcement and the Internet 183, 190 (Irini A. Stamatoudi ed., 2010) ("[T]he flag may need to be an immense crimson banner before the service provider's obligation to intervene comes into play . . . ."). Trademark case law took a similar approach, requiring a high level of knowledge before a duty to cease infringements arises. See Tiffany (NJ) Inc. v. eBay, Inc., 576 F. Supp. 2d 463, 508–10 (S.D.N.Y. 2008) (noting specificity of knowledge required in trademark infringement cases), aff'd in part and rev'd in part on other grounds, 600 F.3d 93 (2d Cir. 2010).

41. *Perfect 10*, 481 F.3d at 763–64.

42. Corbis Corp. v. Amazon.com, Inc., 351 F. Supp. 2d 1090, 1108 (W.D. Wash. 2004).

43. Hendrickson v. eBay Inc., 165 F. Supp. 2d 1082, 1092–93 (C.D. Cal. 2001) (citing 17 U.S.C. § 512(c)(3)(B) (2006) and discussing inadequacy of imprecise infringement notices).

'red flags.'"[44] Therefore, as a matter of legal reality, a webhost that adheres to the threshold requirements of the DMCA and to its ex post procedure is most likely to be off the hook.[45]

Not only does the law provide no incentive for webhosts to filter out copyrighted materials, but it actually provides them with a *disincentive* to filter or monitor content outside the very narrow "red flag" bounds. This disincentive is created by the second exception to the section 512(c) safe harbor, which provides that a webhost will be ineligible for the safe harbor if it derives a financial benefit from an infringement that it can control.[46] While the exact effect of this exception is unclear,[47] at least one court ruled that the ability to filter copyrighted materials is a factor that can satisfy the "control" prong of this exception.[48] This interpretation discourages webhosts from utilizing filters for fear of losing eligibility for the section 512(c) safe harbor.[49]

At the end of the day, then, the message Congress conveyed to webhosts is very simple: They are expected to remain passive in the ex ante phase. An attempt by webhosts to screen may even run the risk of exposure to considerable monetary liability.[50]

The congressional choice to exempt webhosts from detecting infringements is puzzling. Webhosts' active participation in copyright enforcement could be most valuable. Specifically, active monitoring and filtering by webhosts can effectively complement the enforcement efforts of content owners. Webhosts are best situated to disrupt infringing activities ex ante, at relatively low cost.[51] Between content owners and webhosts, the latter have better ability to deploy automatic filters and to identify

---

44. Viacom Int'l Inc. v. YouTube, Inc., 718 F. Supp. 2d 514, 523–24 (S.D.N.Y. 2010) (quoting UMG Recordings, Inc. v. Veoh Networks Inc., 665 F. Supp. 2d 1099, 1108 (C.D. Cal. 2009)).

45. Id. at 525 ("[The DMCA's] safe harbor is clear and practical: if a service provider knows (from notice from the owner, or a 'red flag') of specific instances of infringement, the provider must promptly remove the infringing material. If not, the burden is on the owner to identify the infringement.").

46. 17 U.S.C. § 512(c)(1)(B). This exception closely resembles the doctrine of vicarious liability. See infra notes 67–68 and accompanying text. For the definition of vicarious liability under copyright law, see supra note 20 and accompanying text.

47. See infra notes 67–71 and accompanying text.

48. Tur v. YouTube, Inc., No. CV064436 FMC AJWX, 2007 WL 1893635, at *3 (C.D. Cal. June 20, 2007) ("[T]he requirement presupposes some antecedent ability to limit or filter copyrighted material."), appeal dismissed per curiam, 562 F.3d 1212 (9th Cir. 2009); see also Io Grp., Inc. v. Veoh Networks, Inc., 586 F. Supp. 2d 1132, 1151 (N.D. Cal. 2008) ("[T]he statute presupposes a service provider's control of its system or network.").

49. On the other hand, in other cases a webhost was not penalized for using a filter. See *UMG*, 665 F. Supp. 2d at 1103 (noting Veoh took more measures than required under DMCA); *Io Grp.*, 586 F. Supp. 2d at 1155 (noting Veoh took "active steps to limit incidents of infringement on its website and works diligently to keep unauthorized works off its website").

50. See supra notes 46, 48–49 and accompanying text.

51. See discussion infra Part II.A.1.

and block materials as they are loaded, prior to their posting.[52] Copyright owners, by contrast, cannot apply automatic filters to block content from being uploaded. Nor can they limit their search to newly uploaded content. Thus, content owners must constantly monitor the entire repertoire of every site on the Internet, in every file format, in order to locate infringing materials. Furthermore, monitoring and filtering by webhosts would obviate the need for duplicative enforcement efforts by content owners. Under the current regime, each copyright owner must deploy enforcement measures in order to enforce her rights. In contrast, webhosts can enforce the rights of multiple content owners at once by running the same technological system. Indeed, each webhost could perform more cost-effectively the activities that are currently performed by multiple content owners.

Monitoring by webhosts would also allow enforcement of the rights of some content owners whose rights may not be enforced *at all* under the current law. The high expenditures involved in policing the Internet on a regular basis and issuing takedown requests may be prohibitive for many content owners, particularly individual authors, and independent studios and publishers.[53] Cooperation with webhosts may be the only feasible way to enforce the rights of these content owners.

A corollary, less obvious, cost of the current regime is its effect on the market for enforcement technologies. Currently, technology companies have only partial motivation to come up with superior enforcement technologies.[54] Given that at present only few websites engage in some sort of filtering, the lion's share of demand for enforcement technologies comes from large commercial content owners.[55] While the demand of such entities underwrites a certain level of competition in the market for enforcement technologies, it falls short of inducing full competition. Rightsholders have no access to webhosts' codes to install their desired filtering tools, and webhosts have minimal incentive to allow them to do so.[56] Indeed, filtering companies offer products that scan the entire Internet for infringements, for the use of copyright owners.[57] But this method is obviously far less cost-effective than if webhosts were locating

52. We are omitting from the framework the option of third-party filtering or ISP monitoring. See discussion infra Parts II.A.2, IV.B.

53. See Tim Wu, Tolerated Use, 31 Colum. J.L. & Arts 617, 619 (2008) [hereinafter Wu, Tolerated Use] (noting enforcement costs might discourage copyright owners from enforcing their rights). In fact, filtering can be quite pricy. See infra note 121 (discussing cost of particular filtering service).

54. See infra notes 57–58 and accompanying text.

55. See infra note 58 and accompanying text.

56. As explained above, webhosts are entitled to the DMCA safe harbor regardless of their use of enforcement technologies. See 17 U.S.C. § 512(m)(1) (2006) (showing DMCA safe harbor does not require "a service provider monitoring its service or affirmatively seeking facts indicating infringing activity"); see also supra notes 34–49 and accompanying text. But see infra Part I.C (discussing agreements between webhosts and rightsholders).

57. These products include, for example, Gracenote, Advestigo, Auditude, Vobile, and Attributor.

infringements ex ante. What is more, the market is skewed toward the needs of content owners, incentivizing companies to develop enforcement technologies that identify as many infringements as possible with little concern for false positives.[58]

The cost of the extant regime is exacerbated by the time-sensitive value of intellectual property. As a general rule, copyright content is most valuable immediately after its release.[59] With the passage of time, the value of content diminishes. As a result, by the time offending content is finally removed, it may be virtually valueless.

What is more, current law does not mandate a "take-down, stay-down" policy, which would prevent infringing content from being reposted by another user.[60] As a result, content owners must send and resend notices for each and every unauthorized upload even on the same platform.[61] Accordingly, in many instances, enforcement under the current regime is a Sisyphean task with no real potential for curing infringements.

Notably, the problem we discuss exists regardless of one's view of the optimal level of copyright enforcement. The current enforcement scheme is problematic even under the view that copyright enforcement is excessive. First, as we show, enforcement under the current regime is socially wasteful. Second, the claim of overenforcement is usually concerned with erroneous removal of permissive content; few would argue against enforcement of blatant, obvious infringements.[62]

---

58. For example, BayTSP, touting the accuracy of its system, has successfully marketed its Content Authentication Platform to major movie studios, music labels, and sports leagues, but it makes no mention of protecting fair uses of copyrighted digital content. See BayTSP Announces 15 Customers Using its Content Authentication Platform for Copyright, Business Intelligence and Monetization, All Business (Mar. 2, 2009), http://www.allbusiness.com/media-telecommunications/movies-sound-recording/11799800-1.html (on file with the *Columbia Law Review*) (announcing that fifteen companies are using BayTSP's platform and describing underlying technologies as "best of breed").

59. See, e.g., Nat'l Basketball Ass'n v. Motorola, Inc., 105 F.3d 841, 853 (2d Cir. 1997) (recognizing importance of protecting "property rights in time-sensitive information" regarding "hot news").

60. Under a "take-down, stay-down" policy, the webhost creates a database of material it had removed in the past per a request of a copyright owner. If any user tries to upload the same content that has been removed, the webhost would recognize the material and block the attempt.

61. See Jane C. Ginsburg, Separating the *Sony* Sheep from the *Grokster* Goats: Reckoning the Future Business Plans of Copyright-Dependent Technology Entrepreneurs, 50 Ariz. L. Rev. 577, 591 (2008) [hereinafter Ginsburg, Separating] ("[A]bsent a pre-upload clearance requirement, one may anticipate that at least some of the content the notified service provider takes down will promptly reappear, hydra-like, on other hosts' sites.").

62. As we discuss in Part I.B below, we agree that erroneous takedowns are a problem under the current regime, and believe our regime addresses this concern as well. See also infra Part I.C.

B. *Legal Uncertainty Under the Section 512(c) Safe Harbor*

One of Congress's goals in enacting the section 512(c) safe harbor was to increase legal certainty for webhosts.[63] Yet section 512(c) as it currently stands falls short of this mark.[64]

A key uncertainty concerns whether section 512(c) immunizes webhosts against *indirect* liability as a general matter. While the statutory history appears fairly settled that it does,[65] some commentators have suggested that section 512 is effective only against direct copyright infringement.[66]

A related and better-founded doubt arises as to whether section 512(c) protects against vicarious liability.[67] The language of the section states that when webhosts have the right and ability to control an infringement and derive direct financial benefit from it, they will be excluded from the protection of the safe harbor.[68] This statutory language appears to indicate that Congress carved out vicarious liability from the scope of the section 512(c) safe harbor. Several scholars have suggested that Congress intended this result,[69] and, indeed, this view has some support in the case law.[70] Other commentators, however, opine that the safe har-

63. H.R. Rep. No. 105-551, pt. 2, at 49–50 (1998) (noting goal to "provide[ ] greater certainty to service providers concerning their legal exposure for infringements that may occur in the course of their activities"); S. Rep. No. 105-190, at 40 (1998) (same).

64. See, e.g., Edward Lee, Decoding the DMCA Safe Harbors, 32 Colum. J.L. & Arts 233, 234 (2009) [hereinafter Lee, Decoding] ("Despite the importance of the DMCA safe harbors for ISPs, basic aspects of them remain unclear."); see also Helman, Secondary Liability, supra note 21, at 135–36 (describing uncertainty in relationship between section 512 and subsequent legislation and case law).

65. See supra note 30.

66. Mark Lemley & R. Anthony Reese, Reducing Digital Copyright Infringement Without Restricting Innovation, 56 Stan. L. Rev. 1345, 1369–72 (2004) ("This safe harbor largely preserves the availability of relief against service providers on the basis of secondary liability for infringement committed using the service, though the safe harbors may somewhat heighten the requirements for holding the provider secondarily liable."). But see A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004, 1025 (9th Cir. 2001) (rejecting district court's conclusion that "subsection 512(d) [does not] shelter[ ] contributory infringers").

67. For the definition of vicarious liability under copyright law, see supra note 20 and accompanying text.

68. 17 U.S.C. § 512(c)(1)(B) (2006).

69. 3 Melville B. Nimmer & David Nimmer, Nimmer on Copyright § 12B.04[A][2][b], at 12B-57 (2010); Mark A. Lemley, Rationalizing Internet Safe Harbors, 6 J. on Telecomm. & High Tech. L. 101, 104 & n.23, 113–14 & n.53 (2007); Alfred C. Yen, Internet Service Provider Liability for Subscriber Copyright Infringement, Enterprise Liability, and the First Amendment, 88 Geo. L.J. 1833, 1882 (2000) ("Subsection (B), therefore, does no more than condition nonliability on the nonexistence of vicarious liability.").

70. CoStar Grp. Inc. v. LoopNet, Inc., 164 F. Supp. 2d 688, 704 (D. Md. 2001) ("[T]he DMCA provides no safe harbor for vicarious infringement because it codifies both elements of vicarious liability."), aff'd, 373 F.3d 544 (4th Cir. 2004). The appellate court, however, held that if an ISP is liable under common law, it "could still look to the DMCA for a safe harbor if it fulfilled the conditions therein." *CoStar*, 373 F.3d at 555; see also

bor *does* protect against vicarious liability and have interpreted the statute to reflect this result.[71]

Another source of uncertainty concerns the eligibility of webhosts that induce copyright infringement for the section 512(c) safe harbor. Under one interpretation, such webhosts are ineligible for the 512(c) safe harbor.[72] Under this interpretation, and in light of the broad construction of inducement the Supreme Court espoused in *Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd.*, any webhost whose business model is predicated on large volumes of traffic runs the risk of falling outside the safe harbor.[73] In *Grokster* the Court ruled that the presence of infringing content often serves as a draw that increases traffic to websites.[74] Hence, the availability of pirated content on a site may constitute prima facie evidence of inducement.

The *Viacom v. YouTube* court, attending to this issue for the first time, did not adopt this interpretation, holding instead that questions of "intent" or "neutrality" towards infringements are irrelevant to eligibility under the DMCA.[75] Some district courts, however, appear to lean towards the view that if webhosts "encouraged or fostered . . . infringement, they would be ineligible for the DMCA's safe harbor provisions."[76] This ques-

---

Perfect 10, Inc. v. CCBill LLC, 488 F.3d 1102, 1117 (9th Cir. 2007) (holding "financial benefit" under section 512(c) should be interpreted consistent with the similar language used to define vicarious infringement); Io Grp., Inc. v. Veoh Networks, Inc., 586 F. Supp. 2d 1132, 1150 (N.D. Cal. 2008) (same).

71. See Jonathan Band & Matthew Schruers, Safe Harbors Against the Liability Hurricane: The Communications Decency Act and the Digital Millennium Copyright Act, 20 Cardozo Arts & Ent. L.J. 295, 305 (2002) (describing how courts have "collapsed some of the distinctions between traditional secondary liability and the DMCA" though Congress did not intend that result); Charles S. Wright, Actual Versus Legal Control: Reading Vicarious Liability for Copyright Infringement into the Digital Millennium Copyright Act of 1998, 75 Wash. L. Rev. 1005, 1028–31 (2000) ("[T]he committee reports leave no doubt that Congress intended to provide some relief from vicarious liability."). To reach this result, scholars argued that the term "control" for purposes of liability under section 512(c) should be interpreted as requiring more than just "capability to control" as is accepted under the common law test of vicarious liability. This interpretation is consistent with the fact that the capability to block infringements is mandated by the notice and takedown procedure under section 512(c)(1)(C). See Hendrickson v. eBay Inc., 165 F. Supp. 2d 1082, 1093–94 (C.D. Cal. 2001) ("Congress could not have intended for courts to hold that a service provider loses immunity under the safe harbor provision of the DMCA because it engages in acts that are specifically required by the DMCA."). Likewise, commentators contended that "'receiv[ing]' a financial 'benefit'" under section 512(c) is a heightened requirement compared with merely "'having' a financial 'interest'" under vicarious liability. Lee, Decoding, supra note 64, at 241.

72. See Ginsburg, Separating, supra note 61, at 591–92 (describing neutrality prerequisite to safe harbor eligibility).

73. Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 545 U.S. 913, 940 (2005).

74. Id.

75. 718 F. Supp. 2d 514, 525–26 (S.D.N.Y. 2010).

76. Arista Records LLC v. USENET.com, Inc., 633 F. Supp. 2d 124, 142 (S.D.N.Y. 2009); see also Columbia Pictures Indus. v. Fung, No. CV 06-5578 SVW(JCx), 2009 WL 6355911, at *16 n.26 (C.D. Cal. Dec. 21, 2009) (noting as one of the grounds of ineligibility

tion is contested in the appeal in *Viacom v. YouTube* which is pending in the Second Circuit.[77]

As a result of the ambiguity in section 512(c), services can never be certain that they are really protected from vicarious liability or from claims of inducement—a vague doctrine in and of itself[78]—even if they comply with all the statutory criteria that qualify one for the safe harbor.

Uncertainty comes at a cost, especially when safe harbors are concerned. First, an unclear safe harbor is largely self-defeating; safe harbors, by their very nature, are supposed to provide actors with certainty.[79] The point and purpose of a safe harbor is to mitigate the results of ambiguous or excessively broad liability doctrines. In such areas of the law, where liability is predicated on broad, vague standards that cannot be easily clarified, it is desirable to establish liability free zones as long as the safe harbors' conditions are clearly specified. Otherwise, lawmakers substitute one vague doctrine for another.[80] A vague safe harbor is of little practical use.[81]

Second, and relatedly, the uncertainty shrouding the section 512(c) safe harbor creates an incentive for webhosts to adopt a risk-averse disposition. As John Calfee and Richard Craswell famously demonstrated, vague standards induce actors to overinvest in precautions since failing to make the necessary marginal investment in precautions exposes them to *full* liability.[82] In other words, insufficient investment in precautions

---

of defendant to section 512(c) that "[p]laintiffs' claims are premised on active inducement of infringement, not passive transmission or storage of infringing materials").

77. Notice of Appeal, Viacom Int'l Inc. v. YouTube Inc., No. 1:07-cv-02103 (LLS) (2d Cir. Aug. 11, 2010). Part of the case has now settled, with the National Music Publishers' Association (NMPA) and Google reaching a settlement. See Stipulation of Dismissal with Prejudice as to Certain Appellants, Football Ass'n Premier League Ltd. v. YouTube, Inc., No. 1:07-CV-03582 (2d Cir. Aug. 18, 2011), available at http://docs.justia.com/cases/federal/appellate-courts/ca2/10-3342/372/0.pdf?1313810498 (on file with the *Columbia Law Review*); see also Press Release, Nat'l Music Publishers' Ass'n, NMPA Reaches Resolution of Copyright Infringement Lawsuit Against YouTube (Aug, 17, 2011), available at http://www.nmpa.org/media/showwhatsnew.asp?id=57 (on file with the *Columbia Law Review*) (describing settlement between YouTube and NMPA); YouTube, Creating New Opportunities for Publishers and Songwriters (Aug. 11, 2011), http://youtube-global.blogspot.com/2011/08/creating-new-opportunities-for.html (on file with the *Columbia Law Review*) (same).

78. See, e.g., Jane C. Ginsburg & Sam Ricketson, Inducers and Authorisers: A Comparison of the US Supreme Court's *Grokster* Decision and the Australian Federal Court's *KaZaa* Ruling, 11 Media & Arts L. Rev. 1, 5–7 (2006) (describing uncertainty regarding inducement doctrine after *Sony* and *Grokster*).

79. Legal certainty was one of the explicit goals Congress had in mind while enacting section 512(c). See supra note 63.

80. See Gideon Parchomovsky & Kevin A. Goldman, Fair Use Harbors, 93 Va. L. Rev. 1483, 1502–03, 1510–11 (2007) (considering benefits of crafting clear and certain safe harbor rules).

81. Id.

82. John E. Calfee & Richard Craswell, Some Effects of Uncertainty on Compliance with Legal Standards, 70 Va. L. Rev. 965, 986 (1984) ("[E]xcessive damage awards will tend to increase any incentives to overcomply . . . ."); see also A. Mitchell Polinsky & Steven

comes at the double cost of paying for the (insufficient) precautions and incurring liability. Overinvesting in precautions, by contrast, raises the probability that no liability will attach. The adoption of excessive precautions is socially wasteful, and in this case, it is likely to produce negative externalities as it leads to unnecessary removal of content and censoring too much speech for fear of liability.[83]

Third, vague legal doctrines give rise to strike suits.[84] Powerful industry participants can take advantage of ambiguous doctrines to threaten legal action against defendants who do not have the financial wherewithal to engage in lengthy and expensive legal battles. Perhaps the best recent illustration of this concern can be found in *Io Group, Inc. v. Veoh Networks, Inc.*[85] and *UMG Recordings, Inc. v. Veoh Networks Inc.*,[86] where two different plaintiffs sued the same video webhost for copyright infringement. In both cases, Veoh won summary judgment based on the section 512(c) safe harbor.[87] Despite its victories in those cases,[88] the legal fees Veoh incurred drove the company to bankruptcy.[89] In light of Veoh's fate, other small companies may simply choose to settle meritless suits even though the law is on their side.[90]

C. *Harms to Speech and Users' Interests*

An additional concern emanating from the present regime, which places enforcement in the hands of copyright owners, is that too much speech will be curtailed. Naturally, the interests of commercial content owners are not perfectly aligned with those of society at large. Commercial content owners, who are most likely to bring infringement suits, are interested in maximizing their revenues; society at large has a much

---

Shavell, Punitive Damages: An Economic Analysis, 111 Harv. L. Rev. 869, 873 (1998) (observing "if injurers are made to pay more than for the harm they cause, wasteful precautions may be taken . . . and risky but socially beneficial activities may be undesirably curtailed").

83. See infra Part I.C.

84. See, e.g., Lucian Ayre Bebchuk, Suing Solely to Extract a Settlement Offer, 17 J. Legal Stud. 437, 440 (1988) (noting imperfect information as a primary incentive to bringing frivolous lawsuits); Avery Katz, The Effect of Frivolous Lawsuits on the Settlement of Litigation, 10 Int'l Rev. L. & Econ. 3, 4 (1990) (presenting "a model that explains strike suits as a result of defendant uncertainty regarding the merit of plaintiffs' claims").

85. 586 F. Supp. 2d 1132 (N.D. Cal. 2008).

86. 665 F. Supp. 2d 1099 (C.D. Cal. 2009).

87. *UMG*, 665 F. Supp. 2d. at 1118; *Io Grp.*, 586 F. Supp. 2d at 1154.

88. *UMG* is on appeal to the Ninth Circuit. Appellants' Brief, UMG Recordings, Inc. v. Veoh Networks, Inc., No. 09-56777 (9th Cir. June 17, 2010), 2010 WL 3706518.

89. Joe Mullin, Uh-oh Veoh: Big Copyright Win Can't Save Online Video-Sharing Company, Corporate Counsel (Mar. 4, 2010), http://www.law.com/jsp/cc/PubArticleCC.jsp?id=1202445517277 (on file with the *Columbia Law Review*). The court refused to grant Veoh its request for legal fees. UMG Recordings, Inc. v. Veoh Networks, Inc., No. CV 07-5744 AHM (AJWx), 2010 WL 1407316, at *1 (C.D. Cal. Apr. 6, 2010).

90. Besides the fear that companies would settle meritless suits, uncertainty may lead them to seek agreements with content owners in order to avoid litigation in the first place. See infra Part I.C.

broader interest in creativity and free speech. This misalignment has several adverse consequences.

First, under the current framework, the initial decision as to whether content is infringing rests solely with copyright owners, who can order the removal of posts at their sole discretion.[91] As interested parties, copyright owners are prone to err on the side of more protection and will thus regard fewer cases than optimal as noninfringing or fair uses.[92] Indeed, copyright holders rarely paused to contemplate the possibility of fair use before requesting the removal of a post until the rulings in *Online Policy Group v. Diebold, Inc.* and *Lenz v. Universal Music Corp.*, where the courts held that a demand to remove content without considering whether it qualifies as fair use may amount to a copyright misuse.[93] Tellingly, a study of the behavior of copyright owners found excessive and even manipulative use of the takedown process.[94]

Website operators, for their part, expeditiously comply with the takedown requests of content owners. As Edward Lee has explained, "it would be foolish, if not a breach of corporate fiduciary duty, for [webhosts] not to do so."[95] The structure of section 512(c) and the risk of becoming ineligible for the safe harbor spur webhosts to side with content owners, without providing adequate protection for users. The net result is that more speech than necessary is being removed from websites.

Online speech is further diminished as a consequence of private arrangements between content owners and webhosts.[96] The most well known private agreement in this context is the "User Generated Content Principles" (UGC Principles) that was established in 2007 by leading con-

---

91. See 17 U.S.C. § 512(c)(3)(A)(i) (2006) (providing person who notifies webhost of infringing material must be copyright holder or holder's agent). Users can contest this decision and request the material to be reposted within fourteen days unless the copyright holder notifies the webhost that it has filed an action seeking a court order. Id. § 512(g).

92. See Sonia K. Katyal, Filtering, Piracy Surveillance and Disobedience, 32 Colum. J.L. & Arts 401, 411 (2009) (arguing that copyright owners may "label all [debatable] detected behavior as 'illegal,' and refer to their adversaries as outlaws or pirates"); see also Annemarie Bridy, Graduated Response and the Turn to Private Ordering in Online Copyright Enforcement, 89 Or. L. Rev. 81, 127 (2010) (noting that "[i]n the project of online surveillance, rights owners have historically put zeal before accuracy").

93. Lenz v. Universal Music Corp., 572 F. Supp. 2d 1150, 1151–52 (N.D. Cal. 2008); Online Policy Grp. v. Diebold, Inc., 337 F. Supp. 2d 1195, 1204 (N.D. Cal. 2004).

94. Pan C. Lee, Daniel S. Park, Allen W. Wang & Jennifer M. Urban, Public Knowledge, Copyright Abuse and Notice 3 (2010), available at http://www.law.berkeley.edu/files/CRAabuseandnotice_final_posted.pdf (on file with the *Columbia Law Review*); see also Greg Sandoval, YouTube Users Caught in Warner Music Spat, CNET News (Jan. 27, 2009, 4:00 AM), http://news.cnet.com/8301-1023_3-10150588-93.html (on file with the *Columbia Law Review*) (describing Warner Music Group's amplified DMCA notices following a licensing quarrel with YouTube).

95. Lee, Decoding, supra note 64, at 234.

96. The emergence of private ordering in this area demonstrates the shortcomings of the extant regime. It shows that webhosts are willing to expend considerable amounts to purchase the peace of mind that the section 512(c) safe harbor was supposed to grant them.

tent producers such as Disney and Viacom, and service providers, including Microsoft and MySpace.[97] Such agreements provide webhosts with semicontractual protection against liability[98] in exchange for their agreement to take various measures to curb copyright infringements by users, including employment of filtering systems.[99]

One may argue that there is nothing wrong with private ordering that leads to a more efficient allocation of filtering obligations.[100] But private ordering in this context suffers from several flaws. First, private arrangements are underinclusive. The majority of content owners and webhosts, let alone newcomers in both industries, are not covered by contracts of this sort.[101] Undercoverage has adverse implications for competition. The enhanced protection of major parties increases the exposure to litigation of nonparty webhosts,[102] and simultaneously creates considerable surplus revenue for major parties. This reality sets hurdles for nonparty rivals to compete with parties to these agreements, which are typically major parties to begin with.[103] For the same reason, agreements between major industry players may constitute a barrier to entry for newcomers. The surplus resources of major parties render it difficult for nascent rivals to compete with them. Moreover, content owners may refuse to enter agreements with webhosts that they would prefer to see vanish

---

97. Principles for User Generated Content Services [hereinafter UGC Principles], http://www.ugcprinciples.com (on file with the *Columbia Law Review*) (last visited Sept. 14, 2011). Avoiding liability is probably not the only reason to enter these agreements. Cooperation may result in business opportunities, including the opportunity for webhosts to derive financial benefit from copyrighted content, which would otherwise expose webhosts to liability under 17 U.S.C. § 512(c)(1)(B).

98. UGC Principles, supra note 97, § 14.

99. Id. § 3.

100. See supra Part I.A for discussion of the inefficiency of enforcement obligations under the current regime.

101. The signatories of the UGC Principles have intended the Principles to become "best practices" and apply beyond the immediate signatories. By shaping best practices, the UGC Principles may deem a site that fails to comply an outlier in litigation processes, rather than a law-abiding, compliance-oriented company. See 3 Ian C. Ballon, E-Commerce and Internet Law 49.05[2] (2010–2011 update). While there is still no case law on the status of the UGC Principles, there are currently no signs—beyond the parties' intention—that the Principles apply beyond the signed parties. The UGC Principles—or another agreement—might, however, acquire a stronger status in the future.

102. It is generally more cost-effective for rightsholders to focus their litigation efforts on major, commercial webhosts that arguably profit from infringements. If copyright owners enter agreements with major webhosts, however, their resources are freed to pursue litigation against nonparty webhosts.

103. Indeed, the UGC Principles are open to the participation of new technology providers. Yet, webhosts may elect not to participate in it for various reasons. In that case, the decision would have an undercoverage effect. Moreover, agreements between content owners and technology companies are, obviously, not open to the participation of any webhosts that wish to sign them. In these cases, the problem of undercoverage is exacerbated.

from the market.[104] Alternatively, rightsholders can wage unjustified legal claims against nonparty webhosts, intended to compel them to enter such agreements. This, in turn, may cause some of them to exit the market and reduce competition.[105]

More importantly for our purposes, collective private ordering to determine the nature of online filtering may injure users' freedom of expression. Webhosts comprise a central venue for speech, but no provision in the UGC Principles assures that webhosts would filter posts sparingly to avoid curtailing more speech than necessary.[106] In fact, filtering mechanisms that webhosts use are often unable to distinguish a verbatim copy from a highly transformative use that contains a fraction of a copyrighted work for parody, criticism, or educational purposes.[107]

That webhosts and content owners choose to externalize costs on nonmembers is not utterly surprising. Private arrangements produce an incentive to choose norms that externalize costs on nonmembers of the arrangement. As Eric Posner observed, "Groups have a strong[er] incentive to adopt or develop norms that externalize costs than those that merely maximize joint welfare without producing negative externalities."[108]

Worse yet, as opposed to a law, which can be challenged by anyone who is injured by it, collective private ordering to determine online filtering may injure users' freedom of expression without any meaningful controls.[109] Withholding relevant information from nonmembers of the agreements further impedes their ability to effectively challenge the arrangement and maintain their rights.

Having explained the shortcomings of the extant regime, in the next Part we will introduce an alternative approach for webhost liability, which

---

104. Claims in this spirit have often been raised against copyright owners by digital services in litigation. See, e.g., Arista Records LLC v. Lime Grp. LLC, 532 F. Supp. 2d 556, 565 (S.D.N.Y. 2007) (involving defendant, who was found to be intentional copyright infringer, claiming that music labels engaged in antitrust violations, including, inter alia, refusing access to hash-based filtering without obtaining a license). Similarly, webhosts may decline to filter out content of some content owners.

105. See supra notes 84–90 and accompanying text (discussing risk of strike suits).

106. The most concrete obligation pertaining to fair use in the UGC Principles is in section 6, which reads: "When sending notices and making claims of infringement, Copyright Owners should accommodate fair use." UGC Principles, supra note 97, § 6.

107. See, e.g., Corynne McSherry, Everyone Who's Made a Hitler Parody Video, Leave the Room, Electronic Frontier Foundation Deeplinks Blog (Apr. 20, 2010), http://www.eff.org/deeplinks/2010/04/everyone-who-s-made-hitler-parody-leave-room (on file with the *Columbia Law Review*) (criticizing YouTube's filter for identifying as infringing posts that include mere fractions of copyrighted works, including parodies).

108. Eric A. Posner, Law, Economics, and Inefficient Norms, 144 U. Pa. L. Rev. 1697, 1723 (1996).

109. There is in fact very low use of the "counter notification" procedure under the DMCA. See Jennifer M. Urban & Laura Quilter, Efficient Process or "Chilling Effects"? Takedown Notices Under Section 512 of the Digital Millennium Copyright Act, 22 Santa Clara Computer & High Tech. L.J. 621, 679–80 (2006) ("Only seven counternotices are included in the Chilling Effects dataset . . . .").

we dub "the technological safe harbor." As we will show, the adoption of this model will alleviate many of the problems that plague the current legal standard.

II. The Best Available Technology Safe Harbor

In this Part we advance our technological safe harbor regime. We submit that the technological safe harbor will not only mitigate many of the problems discussed in Part I, but will also improve the overall utility of the copyright enforcement system. Part II begins by describing the conceptual framework for the technological safe harbor model. We then move to discussing the best available technology standard and practical aspects of the filtering process. Finally, we address the benefits of the model. We show that the technological safe harbor model is welfare enhancing on two levels. First, it would create an incentive for the relevant parties to filter content in a manner that respects both copyright and fair use. Second, it would spur dynamic competition in the market of filtering mechanisms by third parties.

A. *The Conceptual Framework*

In his seminal work, *The Cost of Accidents*, Guido Calabresi introduced the concept of "the cheapest cost avoider."[110] Calabresi persuasively argued that the burden of minimizing the harm from accidents should be imposed on the party that can achieve this result at the least possible cost.[111] Although Calabresi's idea is very powerful, it is also a bit misleading. The real challenge for lawmakers is not to *allocate liability* to the party who can abate costs more cheaply, but rather to identify the *mechanism* by which costs may be minimized. The difference is not a matter of semantics. It is often the case that the best way to minimize the cost of accidents—or, in our case, infringements—is via collaboration between the parties involved.[112] Furthermore, sometimes the cost of avoidance may be minimized by involving third parties.[113] The challenge of online infringements by websites is precisely such a case.

The challenge of online infringements involves two dimensions. The first is temporal. Online infringements may be addressed ex ante, before

---

110. See, e.g., Guido Calabresi, The Costs of Accidents: A Legal and Economic Analysis 139 (1970) (introducing notion that liability should be imposed on cheapest, or least, cost avoider in context of tort law).

111. Id.; see also Kyle D. Logue & Joel Slemrod, Of Coase, Calabresi, and Optimal Tax Liability, 63 Tax L. Rev. 797, 819 (2010) (noting rationale of least harm avoider, together with considerations of judgment-proof defendants, is standard economic rationale for vicarious liability in tort law).

112. Avihay Dorfman & Assaf Jacob, Copyright as Tort, 12 Theoretical Inquiries L. 59, 70–72 (2011) (noting that "the normative question (concerning the scope that ownership's authority should have) . . . [is] one of coordination in a world entertaining an owner and a mass of non-owners").

113. See infra Part II.A.2.

content is posted, through careful screening of the materials posted onto websites, or ex post, after content is uploaded, via a notice and takedown procedure. The second dimension requires policymakers to decide on which party to impose the burden of preventing infringement. The burden can either be imposed on content owners or on webhosts. Of course, combinations of the above options are also possible. But we commence our discussion with the four basic options and add the combinations later.

1. *Two-Party Analysis.* — Consider first the option of ex ante avoidance through filtering. The duty to screen all content and filter out infringing material may be imposed either on content owners or on webhosts. Imposing the duty on content owners makes very little sense as they have no access to the code of hosting sites, and website operators will rightly be reluctant to grant content owners access privileges to their proprietary code or otherwise intervene in the core operation of their business.[114] Without access, content owners will not be able to install and update their filtering systems.[115] But even if access were to be mandated, filtering by content owners raises the obvious problem of duplicative expenditures that could be easily avoided. Given that detection and filtering are characterized by economies of scale, it makes no sense to require each and every content owner to engage in filtering on each and every internet site; this would inject serious delays into the system and would be overall welfare-reducing.

What about the alternative option of imposing the responsibility to filter on webhosts? This option is clearly more sensible. Webhosts can easily implement filters and will realize the economies of scale inherent in detection. Each website can enforce the rights of multiple content owners by running the same technological system at a relatively small increase in the marginal cost. Webhosts, however, face a serious information problem. In many cases, they have no way of knowing which materials are copyrighted by someone other than the users who have uploaded them.[116] Nor can they know if the upload was authorized. It must be borne in mind that the quantity of content is growing at a dizzying rate and most works, although protected, are not registered by their au-

---

114. Unauthorized access to computers is illegal under federal law as well as under various state laws. See, e.g., 18 U.S.C. §§ 1030, 2511 (2006) (prohibiting, respectively, computer fraud and interception of electronic communications); see also Max Stul Oppenheimer, Internet Cookies: When is Permission Consent?, 85 Neb. L. Rev. 383, 396–98 (2006) (discussing federal and state laws concerning unauthorized access to computers).

115. Because of the lack of access to webhosts' codes, anti-infringement tools that are currently on the market typically constantly scan the entire Internet for infringements. See supra note 57 and accompanying text.

116. See, e.g., Pamela Samuelson et al., The Copyright Principles Project: Directions for Reform, 25 Berkeley Tech. L.J. 1175, 1186 (2010) ("[I]nadequacies in notice about copyright claims and reduced incentives to register copyright claims have contributed to substantial difficulties in tracking down who owns which rights in which works.").

thors.[117] Webhosts have no realistic way to overcome this problem. And they should not be expected to do so.

This leaves us with the ex post solutions. Ex post intervention falls prey to the same problems we just discussed. Ex post, after the offending materials are posted, the relevant remedial step is removal. The duty to remove infringing content may be imposed on either content owners or webhosts. The option of requiring content owners to remove offending materials can be readily rejected. Without full access to the proprietary code of websites, content owners will not be able to perform this task, and webhosts are not likely to grant them the requisite access. Imposing the duty to remove on content owners will likewise reintroduce the problem of massive duplicative expenditures, as it will require content owners to search for infringing content before they can proceed to remove it. This problematic scenario is what section 512 currently offers.

Imposing the duty to remove on webhosts *without more* will also accomplish little. In principle, webhosts can easily remove content from their websites, but once again they will lack the requisite information to perform this task. Except perhaps in the most obvious cases of infringement that involve high-profile commercial works, webhosts will have no way of identifying infringing content.[118]

Our discussion of the discrete solutions makes it clear that the challenge of online infringement calls for a composite solution predicated on collaboration between content owners, webhosts, and technology providers. Put differently, the optimal solution in this case involves a certain apportionment of responsibilities among the actors involved. Begin with the content owners. They are best situated to deal with the informational dimension of the problem. After all, they possess all the relevant information about their protected works. Hence, it makes sense to impose on content owners the duty to inform webhosts about the legal status of their works. Concretely, content owners should be required to submit information on their protected works to website operators so that they can include these works in the database of copyrighted works. Once this information is received, webhosts can employ filtering systems to detect infringements.

Webhosts, for their part, should be entrusted with the tasks of screening for infringing material and preventing it from being posted if it matches copyrighted works in the database.[119] Webhosts should also be responsible for removing infringing content upon receiving notice from copyright owners or whenever the technological systems they employ discover infringing content that they initially missed.

---

117. See infra notes 136–148 and accompanying text.

118. See Samuelson et al., supra note 116, at 1186 ("[I]nadequacies in notice about copyright claims and reduced incentives to register copyright claims have contributed to substantial difficulties in tracking down who owns which rights in which works.").

119. For a discussion of the degree of match needed for blocking the material, see infra Part III.A.

Once webhosts employ software solutions capable of content matching, they will be able to search for infringing material on a continuous basis and remove offending content whenever it is found. Furthermore, in this technological reality, content owners would be able to submit new copyrighted works to webhosts on an ongoing basis and have the webhosts enforce their rights for them. Of course, if content owners somehow detect infringing content that has been missed by the technology, they should be able to notify the webhosts and ask that the putatively infringing content be removed.

2. *Multi-Party Analysis.* — The addition of third parties to the analysis suggests an even more cost-effective solution.[120] Instead of allocating the filtering responsibility to every webhost, it is possible to entrust all filtering duties to independent providers who will provide filtering services to all webhosts. In fact, doing so will result in a dramatic cost reduction. To see why, consider the cost structure of filtering. The fixed cost associated with filtering is quite significant. A filtering system needs to be purchased, installed, updated, and maintained. The marginal cost of filtering is constant and comparatively low. Once a filtering system is up and running, the cost of screening content does not vary within very broad limits. Nor is the quality of the screening compromised. Hence, the average cost of screening diminishes with scale.

The fact that filtering displays economies of scale suggests that it makes no sense to require each webhost to conduct its own filtering. Imposing the responsibility to filter on each and every webhost would mean that hundreds of thousands of webhosts will need to incur the relatively high fixed cost of setting up and operationalizing their own filtering systems. Worse yet, such responsibility might pose a risk that small, noncommercial webhosts will be deprived of the proposed safe harbor and be exposed to liability merely because they cannot afford to install the filter.[121] This would raise entry barriers for newcomers and have an anticompetitive effect on the market for webhosts.[122]

A dramatic cost saving could be effected in this case if webhosts were to outsource filtering to a third party that would carry out this activity for them. The establishment of a central filtering clearinghouse that would vet content for all webhosts would eliminate duplicative expenditures on enforcement technology, without scarifying its efficiency.

Naturally, a solution of a single clearinghouse raises other concerns. Such a clearinghouse would constitute a monopoly that would be likely to

120. We are grateful to Peter Siegelman for suggesting this solution to us.

121. This can be quite a steep price. As of 2007, Audible Magic was charging sites a monthly fee that could add up to about $1 million annually for its filtering services. Kevin J. Delaney et al., Policing Web Video with 'Fingerprints', Wall St. J., Apr. 23, 2007, at B1.

122. Competition in the filtering market might partially alleviate this concern; yet it is likely that competition in the market for filtering technologies will result in multiple filtering technologies with varying prices, with the better screening technologies being the more expensive ones.

overcharge webhosts for its services and might provide inferior service. The monopoly problem may be easily sidestepped, however. Instead of setting up a single clearinghouse, it is possible to establish several clearinghouses—say, five or six—which would compete among themselves to lure webhosts.[123] Competition among clearinghouses over filtering services will ensure competitive prices to webhosts, as well as quality of service.

The creation of several filtering clearinghouses, as opposed to one, would generate another important advantage. It is likely to create competition in the market for filtering technologies. The filtering centers we envision are essentially service providers. To perform their task they will need filtering systems, which they can either make themselves or procure from other technology companies. The presence of several centers that compete for clientele would mean that each of them would have an incentive to adopt superior filtering systems as they are produced. This, in turn, would give technology companies that produce filtering systems a strong financial motivation to come up with improved filtering products.

In light of the many advantages of this model, in terms of both dynamic and static efficiency, we believe that it should inform the legal policy toward online infringement. In the proceeding section, we explain how current law should be changed to enable implementation of the technological enforcement model we propose.[124]

---

123. Today, most webhosts utilize third-party software, with the palpable exception of Google, which utilizes its own system on YouTube. See What is YouTube's Content ID Tool?, YouTube Help, Google, http://www.google.com/support/youtube/bin/answer.py?hl=en&answer=83766 (on file with the *Columbia Law Review*) (last updated Apr. 27, 2011); see also Tony Bates, The Perils of YouTube Filtering: Part I, Mich. Telecomm. & Tech. L. Rev. Blog (Dec. 7, 2007, 3:43 AM), http://www.mttlrblog.org/2007/12/07/the-perils-of-youtube-filtering-part-1 (on file with the *Columbia Law Review*) (commenting on YouTube policy); Tony Bates, The Perils of YouTube Filtering: Part 2, Mich. Telecomm. & Tech. L. Rev. Blog (Dec. 7, 2007, 1:10 PM), http://www.mttlrblog.org/2007/12/07/the-perils-of-youtube-filtering-part-2 (on file with the *Columbia Law Review*) (same).

124. Cf. Peter S. Menell & David Nimmer, Legal Realism in Action: Indirect Copyright Liability's Continuing Tort Framework and *Sony*'s De Facto Demise, 55 UCLA L. Rev. 143, 149 (2007) ("We believe the traditional tort framework offers a balanced and dynamic mechanism for addressing the many challenges of adapting copyright law to new technology."). Menell and Nimmer argue that the law on secondary liability of technology providers reflects (and should reflect) the tort principle of Reasonable Alternative Design (RAD), rather than the Supreme Court's "staple article of commerce" defense, announced in Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417 (1984). Under the RAD principle, "a product is defective in design when its foreseeable risks of harm could have been reduced or avoided though the adoption of a RAD . . . and the omission of the alternative design renders the product not reasonably safe." Menell & Nimmer, supra, at 154. Our Best Available Technology Standard may resemble the RAD principle in this respect. Under this analysis, the best available technology is a RAD: If adopted, it could shield webhosts from copyright liability. If not, the webhost's technology might appear "defective" and expose the webhost to copyright liability. Note, however, that our standard is not limited to indirect copyright liability and covers direct liability as well. Also, as discussed in Part II, the mechanism of our proposal is substantially different from the ex post mechanism of the RAD principle.

B. *The Best Available Technology Standard*

The optimal approach to the challenge of online infringement appears to suggest that webhosts should employ efficient technology to screen copyrighted works from their websites based on information provided to them by content owners. In order to implement this idea in reality, we propose a fundamental change in the liability regime: to substitute the safe harbor in section 512 with a regime that we term "the technological safe harbor." Under our proposal, webhosts will be exempt from monetary liability if they can show that they employed the "best" filtering technology available on the market when the alleged infringement occurred.

To get a handle on what we mean by the best available technology, it is useful to use the reference point of a "perfect screening technology." A perfect screening technology is the Platonic form of a filtering technology, one that is capable of blocking *all* infringing materials, yet leaves *all* noninfringing materials intact. A perfect screening technology is also perfectly cost-effective: It strikes an equilibrium, where the maximum enforcement measures that do not inflict undue harm are being taken.

Alas, a perfect screening technology is unattainable. Even the most accurate filter, which is based on the most updated reference database, would have few tools to evaluate the legality of posts. Posts may be licensed or authorized, or may qualify as fair use.[125] It is also possible that the copyright of the underlying work that owners inserted into the database has expired[126] or perhaps was without merit in the first place.[127] Given that seminal concepts of copyright law such as the idea/expression dichotomy[128] and fair use[129] are vague standards[130] (as opposed to bright-line rules),[131] technological systems, sophisticated though they

125. 17 U.S.C. § 107 (2006) ("[T]he fair use of a copyrighted work . . . is not an infringement of copyright.").

126. Id. § 106A(d) (setting duration of copyrights).

127. For example, the work might be unoriginal, not fixed, or created by a government body.

128. 17 U.S.C. § 102(b) (setting rule that copyright may be claimed only in "expression" of a work of authorship, and not in its "idea").

129. Id. § 107.

130. Regarding the idea/expression dichotomy, see, for example, Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930) ("Nobody has ever been able to fix that boundary, and nobody ever can."); Benjamin A. Goldberger, How the "Summer of the Spinoff" Came to Be: The Branding of Characters in American Mass Media, 23 Loy. L.A. Ent. L. Rev. 301, 362 (2003) (dubbing idea/expression dichotomy a "deceptively simple concept"). Regarding fair use, see, for example, Dellar v. Samuel Goldwyn, Inc., 104 F.2d 661, 662 (2d Cir. 1939) (per curiam) (describing fair use as the "most troublesome [doctrine] in the whole law of copyright"); Gideon Parchomovsky & Philip J. Weiser, Beyond Fair Use, 96 Cornell L. Rev. 91, 93 (2010) ("[C]ourts generally keep the [fair use] doctrine as vague as possible and decline to provide a formula for what constitutes fair use.").

131. Legal norms are generally expressed as either standards or rules. Standards provide judges with considerable discretion, but result in deficient ex ante certainty, as well as inconsistency among judges. Rules typically provide little flexibility to judges, and might

may be, would fall short of the gold mark of perfection. This, however, does not undermine our proposal in the least, since human-based solutions are both prohibitively costly and impractical and would yield even poorer outcomes.[132]

The standard we propose is of necessity more modest than a perfect screening technology. Compared to a perfect screening technology, the best available technology standard should be thought of as a second-best solution. The best available technology is the technology that offers the best effectiveness/cost ratio.

Effectiveness is measured by the number of infringing works and noninfringing works allowed. The task of the technology is to identify and block the *greatest* number of infringing works and the *smallest* number of noninfringing works. Accordingly, effectiveness can be denoted as a fraction with number of infringing works removed as the numerator and the number of noninfringing works removed as the denominator. The larger the number (or ratio), the more effective the technology. Hence, it should be possible to array all existing technologies based on their effectiveness. To illustrate, assume that technology A removed 100 infringing works and 2 noninfringing ones. It would be assigned the number 50. Technology B on the other hand, removed 80 infringing works and 4 noninfringing ones and would therefore be assigned the number 20. Based on these numbers, technology A would be considered more effective than B. It is possible to similarly rank other technologies.

For the purpose of determining effectiveness, we give the same weight to "false negatives" (infringing posts not blocked) and "false positives" (noninfringing posts blocked). We chose to do so for two reasons. First, especially in the absence of empirical data on the relative social cost of each error, it is impossible to decide whether to err on the side of "false negatives" or "false positives" even on a particular case, let alone categorically; permitting infringements lowers the level of copyright protection, while unnecessary blocking of posts threatens to turn filters into vehicles that stifle fair use and jeopardize free speech. What is more, in order to promote optimal copyright protection in the long term, develop-

---

appear arbitrary or unfair in particular cases. The strength of rules is in providing ex ante clarity and predictability. On the distinction between rules and standards, see generally Frederick Schauer, Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life 15 (1991) (examining use of rules as "decision-making characterized by its reliance on entrenched but potentially under- and over- inclusive generalizations"); Louis Kaplow, Rules Versus Standards: An Economic Analysis, 42 Duke L.J. 557, 560 (1992) (noting "the only distinction between rules and standards is the extent to which efforts to give content to the law are undertaken before or after individuals act" (emphasis omitted)); Pierre Schlag, Rules and Standards, 33 UCLA L. Rev. 379, 381–83 (1985) (describing distinction between rules and standards); Cass R. Sunstein, Problems with Rules, 83 Calif. L. Rev. 953, 957 (1995) ("In every area of regulation . . . it is necessary to choose between general rules and case-by-case decisions.").

132. See infra Part III.A.

ers of filtering mechanisms should have an equal incentive to invest in minimizing both types of mistakes.

Effectiveness, however, is merely the first measure we propose. Cost must also be a consideration. Given marginal diminishing returns to enforcement, it is undesirable to force webhosts to adopt the most effective technology regardless of its cost. Hence, the price of various technologies must be incorporated into the calculus and technologies must be compared based on an effectiveness-per-price basis. In principle, if the marginal cost of preventing an infringement is higher than the harm inflicted by it, then liability should not attach to a website that fails to stop it.[133] Clearly, there is a point on the price continuum at which the cost of enforcement exceeds the benefit. Hence, only technologies that fall *under* that point should be compared, and then among those the technology that offers the best effectiveness-to-cost combination should be selected.

Clearly, there could be more than one technology that meets the definition of the best available technology at any given time—if more than one technology arrives at the same ratio between effectiveness and cost. Thus, the best available technology standard does not denote a single technology, but rather a range of technologies, each of which would satisfy the standard. As elaborated below,[134] webhosts that use any of the technologies that satisfy the standard will be protected from liability.

C. *From Theory to Practice*

In the legal reality we envision, content owners will be entitled to pursue users for posting infringing content. However, they will not be able to seek monetary relief against webhosts that met the best available technology standard. Hence, our proposal will reduce the number of suits against webhosts to the bare minimum, leaving only those webhosts who failed to employ the best available technology exposed to the risk of litigation.[135] In the proceeding discussion we discuss the steps necessary for implementation of our proposal and how best to perform each step.

1. *Creating a Database of Protected Content.* — The first challenge for effective filtering is informational. Currently, it is often impossible to know which expressive content is copyrighted by whom. Copyright protection springs into existence the moment original expression is fixed in

133. See, e.g., J. Randy Beck, The False Claims Act and the English Eradication of Qui Tam Legislation, 78 N.C. L. Rev. 539, 609–10 (2000) ("A rational regulatory system seeks an optimal level of enforcement—one that adequately fulfills the statutory purposes while minimizing social costs."); Michael K. Block & Joseph Gregory Sidak, The Cost of Antitrust Deterrence: Why Not Hang a Price Fixer Now and Then?, 68 Geo. L.J. 1131, 1131 (1980) (noting "an efficient enforcement policy will not deter all antitrust violations because the cost of deterring some of the violations will be greater than the harm averted").

134. See infra Part II.C.2.

135. Under the new regime, webhosts would not be obliged to meet the threshold requirements of section 512(i), to accommodate standard technological measures and to terminate repeat infringers. 17 U.S.C. § 512(i) (2006).

a tangible medium.[136] Registration of copyright claims, though encouraged,[137] is *not* a prerequisite for obtaining copyright protection.[138] Nor are deposit[139] or notice.[140] In reality, the vast majority of copyrighted content is not registered anywhere.[141] True, registration is a precondition for bringing an infringement suit,[142] but the Act allows copyright owners to register their works any time prior to the commencement of litigation.[143] Many copyright owners do not ever attach a copyright notice to their works.[144] And even in cases in which notice was affixed by the owner, it can often be easily removed from the digital copies that are uploaded onto websites.[145] This means that webhosts cannot rely on any external sources to ascertain the legal status of content.[146] There is no a priori way to create a comprehensive database of copyrighted works. All works in the United States are "born" copyrighted,[147] and as long as peo-

---

136. Id. § 102(a) ("Copyright protection subsists . . . in original works of authorship fixed in any tangible medium of expression . . . .").

137. Registration is a precondition for bringing an infringement suit, as well as for statutory damages and attorney's fees. Id. §§ 411(a), 412.

138. Id. § 408 ("[R]egistration is not a condition of copyright protection.").

139. Id. § 407(a) ("Neither the deposit requirements of this subsection nor the acquisition provisions of subsection (e) are conditions of copyright protection.").

140. Id. § 401(a) ("[A] notice of copyright as provided by this section *may* be placed on publicly distributed copies . . . ." (emphasis added)); see also Richard A. Epstein, The Dubious Constitutionality of the Copyright Term Extension Act, 36 Loy. L.A. L. Rev. 123, 124 (2002) ("[C]opyright law . . . flipped over from a system that protected only rights that were claimed to one that vests all rights, whether claimed or not.").

141. Samuelson et al., supra note 116, at 1186 ("Despite the existence of some incentives to register copyright claims with the Copyright Office, relatively few authors actually do so, which means that the public does not have access to useful information about who the owners are and how to track them down to seek permission."); see also Christopher Sprigman, Reform(aliz)ing Copyright, 57 Stan. L. Rev. 485, 513 (2004) (noting "Copyright Office data on the annual number of copyright registrations . . . suggest[s] that the rate of registration is responsive to relatively small changes in registration fees").

142. 17 U.S.C. § 411.

143. See id. § 408(a) (stating registration is permissive and can be done anytime during copyright term).

144. This claim is evident, considering that all original works that are fixed qualify for copyright protection, including all blog posts, user comments on news sites, uploaded homemade clips, and even email messages.

145. Copyright notice comes in the form of the word "copyright" or the symbol "©" (or the abbreviation "Copr.") followed by the year of first publication and the name of the copyright owner. See 17 U.S.C. § 401 (listing forms of notice); see also 37 C.F.R. § 201.20 (2010) (describing methods of affixation and positions that satisfy notice requirement). Removal of notice might therefore require simply erasing or editing out these symbols or the snippet where they appear from the work that is being uploaded.

146. Even the Copyright Office's registry is neither obligatory nor exhaustive. See 17 U.S.C. § 408(a) (noting registration is permissive and can be done anytime during copyright term); see also Jane C. Ginsburg, The U.S. Experience with Mandatory Copyright Formalities: A Love/Hate Relationship, 33 Colum. J.L. & Arts 311, 338–41 (2010) (discussing effects of registration).

147. 17 U.S.C. § 102(a) ("Copyright protection subsists . . . in original works of authorship fixed in any tangible medium of expression . . . .").

ple continue to create, the contours of the copyrighted world are reshaping.[148] As a consequence, webhosts must rely on copyright owners to provide information about the works they want protected—against which webhosts could check users' uploads to seek a match.

However, copyright owners cannot realistically be expected to provide information about their works to each and every website. Imposing such a requirement on them is asking too much. Nor should webhosts be required to contact copyright owners for data. Given the ever growing number of copyright owners and the fact that copyrights are transferable,[149] no webhost can be expected to succeed in identifying all rights holders, let alone contacting them. The cost of compliance with such a requirement would clearly be prohibitive.

What should the solution be then? The solution we envision is the creation of a single database that will contain all relevant information about copyrighted content.[150] The establishment of a central information

148. Some commentators suggest shifting to a formalities-based system, in which copyright protection would be *conditioned* upon a work's registration in a database. Such a change would not only solve the problem but avoid it in the first place. See Lawrence Lessig, For the Love of Culture: Google, Copyright, and Our Future, New Republic, Feb. 4, 2010, at 24, 29 (arguing for establishing an "absolute obligation [for domestic authors] to register their work . . . [when] [f]ailure to register would mean that the work would pass into the public domain"); see also David Fagundes, Crystals in the Public Domain, 50 B.C. L. Rev. 139, 182 (2009) ("[A]dvocates of increased formalities have suggested that the Copyright Office create a public database of copyrighted works in order to centralize and publicize their ownership status as a means of buttressing the public advantages brought by more formality."). Such approaches probably fall outside of the Berne Convention as it currently stands. See generally Berne Convention for the Protection of Literary and Artistic Works art. 5(1), opened for signature July 24, 1971, 1161 U.N.T.S. 31 (amended Sept. 28, 1979) [hereinafter Berne Convention]. More nuanced approaches call for creating stronger incentives for copyright owners to comply with formalities. See, e.g., Samuelson et al., supra note 116, at 1198–1202 (recommending creation of system that would encourage registration); Sprigman, supra note 141, at 555 (proposing system of "new-style" formalities that, although nominally voluntary, are de facto mandatory for any rightsholder whose work may have commercial value).

149. See 17 U.S.C. § 204(a) (providing requirements for transfers of copyright ownership).

150. In practice, modern filters' databases do not contain the works themselves; rather, these databases contain fingerprints (i.e., identifiable components of files (also known as robust hashes or visual signatures)) or watermarks (i.e., embedded visible or invisible marks that remain in the file even when the file is copied or altered). See June M. Besek, Anti-Circumvention Laws and Copyright: A Report from the Kernochan Center for Law, Media and the Arts, 27 Colum. J.L. & Arts 385, 447–48 (2004) (describing watermarks); see also Brad Stone & Miguel Helft, New Weapon in Web War over Piracy, N.Y. Times, Feb. 19, 2007, at C1; Advisory Comm. to the Cong. Internet Caucus, 111th Annual Technology Policy Exhibition Demonstrators, http://www.netcaucus.org/events/2008/kickoff/demonstrators.shtml (on file with the *Columbia Law Review*) (describing various filtering technologies); Michael Liedtke, Audible Magic Emerging as Top Copyright Cop in Digital Revolution, USA Today, Mar. 23, 2007, http://www.usatoday.com/tech/news/techinnovations/2007-03-23-magic-police_N.htm (on file with the *Columbia Law Review*) (describing rise of particular filtering technology company); David Kravets, Analysis: FCC Comcast Order Is Open Invitation to Internet Filtering, Wired

repository will provide content owners with a convenient way to convey information to webhosts and will also allow webhosts access to the information they need to engage in the filtering process.[151]

The central database we envision may either be created under the auspices of the Copyright Office or under the supervision of a private body. Entrusting the creation and management of the database to the Copyright Office will provide three advantages. First, the Copyright Office is the most natural reference point for most copyright owners, many of whom have had past dealings with the Copyright Office for other purposes.[152] Second, and relatedly, the Copyright Office already has information about all the copyrighted works that have been registered. Furthermore, since registration with the Copyright Office is a precondition for filing an infringement suit, it makes little sense to force content owners to register their works twice with two separate entities. Third, and most importantly, it is critical to ensure access to the proposed database to all website operators and technology developers on a nondiscriminatory basis. After all, the information in the database will provide the foundation for the filtering process and all systems must be compatible with it and have uninterrupted access to it.

Alternatively, the task of managing the database may be entrusted to a private (or semiprivate) body similar to the Internet Corporation for Assigned Names and Numbers (ICANN), which handles registration of domain names.[153] This alternative is attractive since it can result in a competitive registration market. ICANN's registry, for example, is fed by various secondary registrars, and not directly by content owners. The various registrants compete among themselves to attract registrations. This competition, in turn, lowers fees and encourages participation by more owners.[154] A similar dynamic may emerge in the context of registration of

---

Threat Level Blog (Aug. 20, 2008, 12:53 PM), http://blog.wired.com/27bstroke6/2008/08/analysis-fcc-co.html (on file with the *Columbia Law Review*) (arguing FCC supports internet service provider use of filtering to block illegal internet activity). On video filtering, see Delaney et al., supra note 121.

151. Our central database is likely to have an additional incidental benefit, in providing the market with updated information regarding copyright ownership, thus facilitating licensing markets for copyright works in general. See Samuelson et al., supra note 116, at 1186 (noting lack of information in market regarding copyright ownership harms licensing markets).

152. Copyright owners deal with the Copyright Office when they register and deposit their works. See 17 U.S.C. § 411 (registering work); id. § 407 (depositing work). Furthermore, because the Copyright Office holds a database of copyrighted works, it is the natural resource for potential licensees who wish to learn the identities of the copyright owners of works they wish to license.

153. Internet Corp. for Assigned Names and Numbers, ICANN Information, http://www.icann.org/general (on file with the *Columbia Law Review*) (last visited Aug. 2, 2011).

154. Professor Lawrence Lessig offered to create a mechanism to turn to a formalities-based copyright system that would be based on the ICANN model. See Lawrence Lessig, Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity 284 (2004). Clearly, competition that would drive down the costs of

copyrighted content. In this case, competition may emerge not only over price, but also over the quality of content marking technologies that may facilitate the filtering process.

The main argument against a central registration system in copyright law is that formal requirements from copyright owners might deprive authors of value that stems from their works.[155] Specifically, the concern is that owners would lose protection because they failed to comply with some fine detail of the registration requirements.[156] These concerns do not apply to our model. Under our model, owners who would choose not to register their works will be able to continue to use their own enforcement measures, as well as to follow the notice and takedown procedure the current system provides, and which our model retains.[157] Therefore, our proposal does not worsen the situation for content owners relative to the current regime.[158]

2. *Determining the Best Available Technology*. — The next step in implementing our proposal is determining which technologies come under the best available technology standard. One option is to defer to the courts that will make these determinations on a case-by-case basis.

The advantages of a post hoc judicial process are that both parties will be forced to participate in discovery, revealing detailed and accurate facts regarding their practices. Courts are also neutral and relatively unsusceptible to capture[159] and other public choice problems.[160] Most im-

---

registration would be more important in Lessig's proposal, where lack of registration implies denial of copyright.

155. The abandoning of formalities under U.S. law, including a registration requirement, resulted mainly from the need to comply with the Berne Convention. See Sprigman, supra note 141, at 543–45 (analyzing reasons why Berne Convention itself adopted negative approach towards formalities).

156. See Molly Shaffer Van Houweling, Author Autonomy and Atomism in Copyright Law, 96 Va. L. Rev. 549, 605–06 (2010) (noting concern that "strict formality requirements were traps for unwary authors").

157. Owners would not lose copyright protection if they do not register, and therefore our proposal does not run into a problem with the Berne Convention that forbids signatories from implementing formalities. The United States became obligated by the Berne Convention following the Berne Convention Implementation Act, Pub. L. No. 100-568, 102 Stat. 2853 (1988) (codified as amended in scattered sections of 17 U.S.C.).

158. As we explain, our proposal retains the notice and takedown mechanism that the current regime provides. See supra note 11 and accompanying text; infra note 206 and accompanying text.

159. See Thomas W. Merrill, Capture Theory and the Courts: 1967–1983, 72 Chi.-Kent L. Rev. 1039, 1050 (1997) (describing belief that "agencies were likely to become 'captured' by the business organizations that they are charged with regulating"); Richard Pierce, Institutional Aspects of Tort Reform, 73 Calif. L. Rev. 917, 935 n.104 (1985) ("'Capture' refers to the tendency of some agencies to favor the industry they are required to regulate by protecting the industry from outside competition and stifling innovation that threatens the status quo in the industry.").

160. See, e.g., Craig Allen Nard, Legal Forms and the Common Law of Patents, 90 B.U. L. Rev. 51, 56–57 (2010) (contending "judicial primacy acts as a bulwark against the more politicized legislative process or capture-prone administrative rulemaking"); Francesco Parisi, Public Choice Theory from the Perspective of Law, *in* The Encyclopedia

portantly, courts are best suited to make findings regarding the infringing status of works and fair use, which are required in order to rank the filters according to their cost-effectiveness.

Nonetheless, we believe that this option should be rejected. Assigning the task of determining the best available technology to the courts will come at a significant cost—both in terms of time and in terms of money. Courts are ill equipped to make technological judgments. Worse yet, since new technologies appear on the scene all the time—a challenge discussed in detail in Part III.C—courts will not be able to rely on precedents and will have to consider every new technology de novo. This means that litigation may result in conflicting decisions, an outcome that will substantially undermine the certainty our proposal is designed to create.[161] Judicial processes might also result in strategic behavior by repeat players[162] and strike suits by content owners who have the financial resources to withstand lengthy and expensive litigation.[163] Finally, courts typically confine themselves to the dispute before them. They tend to consider only the claims and interests of the specific parties to the dispute, and are unable to gather information about third parties who are likely to be affected by their decisions.[164]

Hence, we propose that the list of best available technologies be determined ex ante by an agency at the Copyright Office, after consulting members of the copyright and technology industries, webhosts, and the general public. This agency would have the requisite expertise to determine which technologies should qualify at any given time. Of course, it would reconvene periodically to revisit the list and update it as new and

---

of Public Choice 214, 222 (Charles K. Rowley & Friedrick Scheider eds., 2004) ("To the extent to which judicial bodies are independent from political forces and shielded from interest group pressure, the process of judicial lawmaking can be considered immune from the collective decision making failures . . . ."). Public choice problems arise when regulatory bodies are tilted in favor of the narrow interests of strong, concentrated, and homogeneous groups at the expense of the social interest. See generally Daniel A. Farber & Philip P. Frickey, The Jurisprudence of Public Choice, 65 Tex. L. Rev. 873, 883–906 (1987) (discussing public choice theory in legislation); William M. Landes & Richard A. Posner, The Independent Judiciary in an Interest-Group Perspective, 18 J.L. & Econ. 875, 877–87 (1975) (proposing economic theory of independent judiciary); Sidney A. Shapiro, Keeping the Baby and Throwing Out the Bathwater: Justice Breyer's Critique of Regulation, 8 Admin. L.J. Am. U. 721, 722 (1995) (describing public choice theory).

161. This adds to the general uncertainty that is inherent to the ex post nature of judicial rulemaking. See, e.g., Shyamkrishna Balganesh, The Pragmatic Incrementalism of Common Law Intellectual Property, 63 Vand. L. Rev. 1543, 1593 (2010) ("Since the process of [judicial] rulemaking is invariably ex post, to the individual actor the law remains uncertain until a decision is actually rendered.").

162. See generally Marc Galanter, Why the "Haves" Come Out Ahead: Speculations on the Limits of Legal Change, 9 Law & Soc'y Rev. 95, 98 (1974) (defining "repeat player" as "a unit which has had and anticipates repeated litigation which has low stakes in the outcome of any one case, and which has the resources to pursue its long-run interests").

163. See supra note 84 and accompanying text.

164. See Balganesh, supra note 161, at 1593 (discussing consequences of fact that "common law courts create rules from the context of a specific dispute").

improved technologies emerge. An alternative approach would be to assign this task to a representative body that would conduct ex ante review of filtering technologies. This model would closely resemble standard-setting organizations (SSOs) that agree on technological standards in other areas.[165] Hence, policymakers would be able to draw on the experience we have had with standard-setting organizations in determining the composition and procedures of our technology setting body.

Clearly, ex ante mechanisms are also not free from concerns. They impose participation costs on the relevant parties and may incur considerable public choice problems.[166] There is also a risk that the decisions of the technology setting body might create externalities for underrepresented groups.[167] Imbalanced decisions, if they occur, might have adverse implications across the board, because all webhosts would need to use a filter that operates under a suboptimal balance between legitimate and infringing content.

For the reasons discussed above,[168] we believe that despite these concerns, the proposed ex ante mechanism will provide the superior and most efficient mechanism, and is the only way to ensure stability and certainty. Moreover, judicial review on the decisions of the technology setting body will likely mitigate these risks and inform the process of distinguishing infringements from fair use.

3. *Creation of Filtering Clearinghouses*. — Once the list of the best available technologies has been compiled, the next challenge is to decide how filtering should be carried out in practice. For the reasons explained above, individual screening is not cost-effective.[169] Demanding each webhost to incur the cost of installing, maintaining, and updating its own filtering software will lead to wasteful duplicative expenditures without any meaningful advantages to content owners or to society. In addition, the cost of compliance may be prohibitive for many small nonprofit web-

---

165. Standard-setting organizations are industry groups that set common standards in a variety of areas, and are especially prevalent in the areas of technology and telecommunications. See, e.g., Mark Lemley, Intellectual Property Rights and Standard-Setting Organizations, 90 Calif. L. Rev. 1889, 1892–93 (2002) (describing role of standard-setting organizations).

166. See supra notes 159–160. Such concerns are often raised in the context of copyright legislative processes. See, e.g., Robert P. Merges, One Hundred Years of Solicitude: Intellectual Property Law, 1900–2000, 88 Calif. L. Rev. 2187, 2234–35 (2000) (reviewing rent seeking by special interest groups in realm of intellectual property legislation); William F. Patry, Copyright and the Legislative Process: A Personal Perspective, 14 Cardozo Arts & Ent. L.J. 139, 141 (1996) (describing lobbying efforts of interest groups in copyright lawmaking process).

167. For example, mash-up artists or other creators that rely on fair use, as well as the public's diffuse interests in preserving fair uses, might continue to be underrepresented compared to the narrowly focused interests of content industries, causing the latter to prevail despite representing the less socially optimal outcome.

168. See supra note 161 and accompanying text.

169. See supra Part II.A.2.

sites that may lack the financial resources to install and operate such a system.

A better solution would be to refer all user generated content to several central clearinghouses that would conduct the filtering for the webhosts. Such clearinghouses would provide a significantly more cost-effective method for filtering. First, they would prevent redundant expenditures of webhosts. Second, they would obviate the need for filtering vendors to tailor filters for each webhost individually, thus allowing for a more efficient and less costly production process of filters. Third, clearinghouses are likely to be able to negotiate a better price with the producers of filtering technologies than with individual webhosts.

In theory, one clearinghouse could suffice for conducting filtering under our model. But, per our above discussion, it is best not to give monopoly power to one center, in order to ameliorate the risk of supra-competitive pricing and imbalanced blocking of materials.[170] In addition, the presence of several clearinghouses is also likely to support competition in the market for filtering products and thereby sustain a steady improvement of such technologies. The existence of a number of clearinghouses would also divide the task and reduce the amount of time it takes to review content and clear it for posting.

A variant of this general framework would be to allow large websites to filter in-house. Indeed, it may be more efficient for some webhosts to integrate the filtering vertically into their operations, as opposed to purchasing the service from an outside provider. Ronald Coase suggested that transactions will be organized in the firm when the cost of doing so is lower than the cost of using the market.[171] While the clearinghouse structure for filtering is generally more efficient, there could be exceptions. For example, companies whose core business is creating technology and are *also* operating webhosts (i.e., Google) might incur considerably lower costs in creating a filter than other webhosts, or perhaps even in purchasing it from an external company.[172]

In order not to intervene in what might sometimes be a more efficient state of affairs, we propose that webhosts would be able to request the inclusion of their own, homemade filtering systems into the list of best available technologies and, if the filter is at least as effective as the best available technologies, the use of that filter would entitle the company to the safe harbor.[173]

---

170. See supra note 123 and accompanying text.

171. Ronald H. Coase, The Nature of the Firm, 4 Economica 386, 390–92 (1937). For other factors affecting the boundaries of the firm, see Sanford J. Grossman & Oliver D. Hart, The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, 94 J. Pol. Econ. 691, 693–95 (1986).

172. This analysis may explain why YouTube only uses its homemade filter as opposed to third-party software. See supra note 123.

173. Since the cost of the filter in this case is going to be borne exclusively by the webhost itself, there is no need for external checks on the cost parameter. Thus, the filter only needs to be effective, under the definition in Part II.B.

D. *The Benefits of the Technological Safe Harbor Model*

The technological safe harbor model is designed to create appropriate incentives for content owners, webhosts, and technology providers to join forces and improve copyright enforcement.

Consider, first, the effect of our model on content owners. The proposed regime provides content owners with a strong motivation to share information about their content with webhosts. Specifically, while registration is not mandatory,[174] our model correlates the level of protection a copyright owner would receive with the level of data she would provide. The more accurate the information a content owner provides, the easier it will be to detect infringements of her content. As a result, content owners would have a strong incentive to update the central database by providing precise information about works they wish to protect.

As far as webhosts are concerned, our mechanism will eliminate the twin risks of expensive litigation and monetary liability for copyright infringement by users. Indeed, it will also eliminate the legal uncertainty webhosts currently face. The use of the clearinghouses by webhosts will automatically qualify them for safe harbor status. If a copyright infringement suit is brought against a webhost, all it will have to do to defeat the lawsuit will be to show that it referred users' content to a clearinghouse for filtering at the time of the infringement. Since the webhosts' practice of referring content to a clearinghouse will probably become public information—as webhosts who follow such practices have an interest in making this fact known—very few cases, if any, will even go to court. Indeed, copyright owners will have no incentive to bring a lawsuit against a webhost that refers content to one of the model's clearinghouses; the webhost will be protected under our technological safe harbor, and the lawsuit will surely be rejected. What is more, our model reduces the cost of webhosts' compliance with the safe harbor provision. As analyzed above, the cost of referring materials to a filtering process saves webhosts the much higher cost of implementing and running the filters themselves.[175]

In combination, these two effects—copyright owners' incentives to provide data and webhosts' incentives to filter—would lead to a dramatic improvement in the efficiency of online copyright enforcement. Enhanced enforcement would augment the economic incentives of copyright industries—as well as amateurs—to produce creative content.[176] Fi-

174. See 17 U.S.C. § 408 (2006) (defining statutory registration requirements).

175. See supra Part II.A.2.

176. Enforcement of copyright law would enable the system to accomplish its goal, which is to enhance the incentive to create. See, e.g., Mazer v. Stein, 347 U.S. 201, 219 (1954) ("The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in 'Science and useful Arts.'" (quoting U.S. Const. art. I, § 8, cl. 8)); see also Julie E. Cohen, *Lochner* in Cyberspace: The New Economic Orthodoxy of "Rights Management,"

nally, it would provide webhosts with legal certainty, which is likely to increase the investment in webhosting platforms.[177]

Importantly, our proposal has another salutary dynamic effect. Its implementation would lead to competition in the market of filtering mechanisms, initiating a virtuous cycle of increasing productivity and efficiency.[178] Under our model, filtering clearinghouses, competing to attract webhosts to use their services, would create constant demand for better filters. This demand would induce technology providers to engage in accelerated development and improvement of filtering systems.

In other words, our proposed safe harbor is "technology-endorsing" in the sense that it ensures that superior new technologies will be constantly adopted. This, in turn, would generate competition in the market for filtering technologies as technology providers would constantly race for the next best available technology.

Innovation in the filtering arena may generate spillovers beyond the realm of copyright enforcement and creative industries—although this is an effect we cannot predict or quantify at this point in time.[179] Filtering systems require optimization of search, identification, and comparison technologies, as well as automation of otherwise human calculations.[180] Each of these technologies could be used separately or in combination with other technologies to enhance production in other industries.[181] As Mark Lemley and Anthony Reese have noted, "Economic evidence strongly suggests that those unanticipated future benefits, or 'spillover' effects, often exceed the immediate value of most new technologies."[182]

---

97 Mich. L. Rev. 462, 471 (1998) ("By guaranteeing authors certain exclusive rights in their creative products, copyright seeks to furnish authors and publishers, respectively, with incentives to invest the effort necessary to create works and distribute them to the public.").

177. See, e.g., Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., 535 U.S. 722, 730–31 (2002) (recognizing, in patent context, "clarity is essential to promote progress, because it enables efficient investment in innovation"); cf. Kaplow, supra note 131, at 613 (arguing absence of clear legal precedent defining rules raises costs to businesses contemplating future acts); Edward Lee, Technological Fair Use, 83 S. Cal. L. Rev. 797, 822–23 (2010) (discussing importance of certainty to innovation and investment in new technologies).

178. Cf. Michael Abramowicz & John F. Duffy, Ending the Patent Monopoly, 157 U. Pa. L. Rev. 1541, 1555–58 (2009) (discussing virtues of competition in framework of proposal to replace U.S. Patent and Trademark Office's current monopoly on patent examination with competitive marketplace).

179. Cf. Clayton M. Christensen, The Innovator's Dilemma 150 (rev. ed. 2003) ("[N]either manufacturers nor customers know how or why the products will be used . . . [or] what specific features of the product will and will not ultimately be valued.").

180. See infra Part III.A. (discussing functionality of filtering systems in context of fair use).

181. Clearly, the new tools that would develop can also be put to bad use, such as undue censorship, and various uses that can create privacy issues. Like most technologies, filtering systems are nothing but tools that can be utilized differently by different parties.

182. Lemley & Reese, supra note 66, at 1387.

III. Challenges and Objections

This Part addresses four practical and conceptual obstacles to our proposal. First, we tackle the argument that an automated model cannot, in principle, deal with fair use cases. Second, we discuss the model's ability to handle tolerated use. Third, and finally, we address the challenge of dynamism—namely, the ability of our model to cope with ongoing improvements in filtering technologies.

A. *The Challenge of Fair Use*

An obvious challenge for our model is how to deal with fair use. Determining whether a certain use is "fair" is an intricate task.[183] There is no shared understanding of the doctrine even among judges. Fair use cases are often marked by frequent reversals, split courts, and inconsistency even on the Supreme Court level.[184] As Judge Posner put it, the fair use test does not "constitute an algorithm that enables decisions to be ground out mechanically."[185]

We agree that technology cannot provide a perfect solution to the challenge of fair use. We contend, however, that our model can handle fair use cases at least as well as the extant regime. To demonstrate that, we would like to take a step back and describe the filtering process in its entirety.

In our vision, filtering systems would classify content into one of three categories. The first category (Category A) would include instances of blatant copyright infringements, namely posts that constitute verbatim copies of copyrighted works. Identifying Category A posts should be fairly straightforward. Such posts will not only match a database item, but would also comprise that single item in its entirety, or nearly in its en-

---

183. The test, which applies to six favored uses—criticism, comment, news reporting, teaching, scholarship, and research—requires a careful analysis of four broadly worded nonconclusive factors, enshrined in section 107 of the Copyright Act. 17 U.S.C. § 107 (2006) ("(1) the purpose and character of the use . . . ; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used . . . ; and (4) the effect of the use upon the potential market for . . . the copyrighted work."). These four factors are illustrative but not exclusive, and courts can consider other parameters as well. See id. But see infra note 196 and accompanying text (noting courts treat factors as limitative, not illustrative).

184. See Pierre Leval, Toward a Fair Use Standard, 103 Harv. L. Rev. 1105, 1106–07 (1990) ("Reversals and divided courts are commonplace."). But see Barton Beebe, An Empirical Study of U.S. Copyright Fair Use Opinions, 1978–2005, 156 U. Pa. L. Rev. 549, 574–75 (2008) ("[O]utside of the cases that reached the Supreme Court . . . , [our fair use case law] has not been marked by especially high reversal, dissent or appeal rates.").

185. Chi. Bd. of Educ. v. Substance, Inc., 354 F.3d 624, 629 (7th Cir. 2003); see also Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 588 (1985) (Brennan, J., dissenting) (asserting that section 107 factors "do not mechanistically resolve fair use issues").

tirety.[186] Category A cases represent the most significant—and legitimate—concern of copyright owners, because these posts form clear substitutes for their products and their source of livelihood.[187] Accordingly, Category A posts should automatically be blocked ex ante from appearing on the site.

Our second proposed category (Category B) would consist of noninfringing posts. This category would comprise posts that contain no traces of an identifiable work, as well as posts that are licensed or otherwise authorized.[188] All these cases raise no issue of infringement.

The third category (Category C), would contain the "hard cases"—posts that may qualify as fair uses. We concede that technology cannot provide an error-proof solution in this case. It should be borne in mind, though, that human review is not a silver bullet either—as our fair use jurisprudence may demonstrate.[189]

We believe then that the optimal solution is a combination of technological and human review. Initially, the degree of reliance on human agents might be quite substantial. Over time, however, as filtering technologies improve, the role of humans is likely to diminish.[190]

Although filters cannot be expected to conduct a full analytical fair use examination,[191] they could use proxies and quantitative measures as indications of whether the fair use factors point towards a positive or negative fair use finding, and increasingly, better align their process with courts' fair use decisions.[192] For example, technology can test whether a work is published—an element of the second fair use factor.[193] It can also easily find the amount taken under the third factor. Filters might also be able to develop proxies for commercialism[194] (e.g., the post's designated

186. The filter would only need to verify that the post has not been authorized or uploaded by the copyright owner of the underlying work herself. Such information can be incorporated into the original database.

187. See Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 591 (1994) ("[W]hen a commercial use amounts to mere duplication of the entirety of an original, it clearly 'supersede[s] the objects' of the original and serves as a market replacement for it, making it likely that cognizable market harm to the original will occur." (citation omitted)).

188. A license can be compulsory, express, or implied. See, e.g., 17 U.S.C. §§ 114–115 (making licenses compulsory); Effects Assocs., Inc. v. Cohen, 908 F.2d 555, 555–56 (9th Cir. 1990) (describing express or implied licenses). By authorization we also include posts that are being uploaded by the copyright owner of the underlying work.

189. See supra note 184 and accompanying text.

190. See infra note 205 and accompanying text.

191. See supra note 183.

192. This is what software does in other contexts, such as Wall Street trading that is now dominated by computer programs that utilize algorithm-based statistical analysis to decide which stocks to invest in.

193. In some instances, such as in the case of photography, however, the question of whether a work was published can be more complicated, although it is doubtful that the complexity would carry into the fair use realm. See generally Thomas F. Cotter, Toward a Functional Definition of Publication in Copyright Law, 92 Minn. L. Rev. 1724 (2008).

194. Commercialism serves as an indication against fair use. See Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 584–85 (1994) ("*Sony* stands for the proposition that the

audience) and transformativeness[195] (e.g., the proportion of original to unoriginal material in the post)—the two main components of the first fair use factor.

Making fair use determinations technologically might be further facilitated by the fact that despite Congress's invitation to judicial flexibility, courts without exception apply the four statutory factors exclusively and attach a binary value to each factor—whether it does or does not point towards a fair use conclusion.[196] Likewise, the filter can take into account the de facto hierarchy between the factors,[197] where "the first and fourth factors dominate the analysis, with the third and second factors trailing in significance."[198] Commentators have identified other patterns in the fair use jurisprudence.[199]

Since filtering technology is not advanced enough to make qualitative determinations, the switch to technology-based enforcement may be aided by the adoption of quantitative benchmarks that will substitute to some degree for qualitative criteria. Specifically, lawmakers (or industry representatives during the process of creating the list of "best available technologies") can agree that certain insignificant appropriations of

---

'fact that a publication was commercial as opposed to nonprofit is a separate factor that tends to weigh against a finding of fair use.'" (quoting Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 562 (1985))); *Harper & Row*, 471 U.S. at 562 ("The fact that a publication was commercial . . . tends to weigh against a finding of fair use."); Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 449 (1984) (noting noncommercial use is presumed to be fair use).

195. Transformativeness serves as an indication for fair use. See, e.g., *Campbell*, 510 U.S. at 578–79 (opining transformative uses further the goal of copyright law).

196. See Beebe, supra note 184, at 563–64 (noting that empirically, courts treat 17 U.S.C. § 107 (2006) as limitative and not only illustrative).

197. Courts have generally regarded the second and the third factors as relatively insignificant. See Compaq Computer Corp. v. Procom Tech., Inc., 908 F. Supp. 1409, 1421 (S.D. Tex. 1995) ("The third factor . . . is generally considered the least important factor of the fair use analysis." (citing *Sony*, 464 U.S. at 449–50)); Dow Jones & Co. v. Bd. of Trade of Chi., 546 F. Supp. 113, 120 (S.D.N.Y. 1982) ("The nature of the copyrighted work seems to be the least important and most unclear of the four factors . . . ."). But see *Campbell*, 510 U.S. at 578 ("All [factors] are to be explored, and the results weighed together, in light of the purposes of copyright."); *Harper & Row*, 471 U.S. at 553 (referring to second factor as "highly relevant to whether a given use is fair").

198. Lydia Pallas Loren, The Pope's Copyright? Aligning Incentives with Reality by Using Creative Motivation to Shape Copyright Protection, 69 La. L. Rev. 1, 31 (2008); see also Michael W. Carroll, Fixing Fair Use, 85 N.C. L. Rev. 1087, 1103 (2007) (arguing that second factor "tends to do little work in swaying the outcome" of the test); Matthew Sag, God in the Machine: A New Structural Analysis of Copyright's Fair Use Doctrine, 11 Mich. Telecomm. & Tech. L. Rev. 381, 434 (2005) (discussing lesser importance of second factor).

199. See, e.g., Michael J. Madison, A Pattern-Oriented Approach to Fair Use, 45 Wm. & Mary L. Rev. 1525, 1645–65 (2004) (concluding fair use doctrine is "explained best as an analytical tool that focuses on social and cultural patterns"); Pamela Samuelson, Unbundling Fair Uses, 77 Fordham L. Rev. 2537, 2537 (2009) (identifying several "policy-relevant clusters" that can predict result of fair use test).

copyrighted content will be per se fair use and, thus, will not constitute copyright infringements.

A blueprint for such quantitative caps can be found in the academic literature. Kevin Goldman, together with one of us, has advanced specific ceilings for uses of copyright content without liability.[200] These or other criteria could be employed by policymakers or by industry and public representatives in the process of creating the list of best available technologies to expand the range of permissible uses of content. Minimal appropriations neither present a risk of sales displacement nor otherwise usurp the market of the original work.[201] Thus, they should not be blocked by the filtering process.

Another possibility is that each content owner would set its level of permissible use individually. We elaborate on this option in our discussion on tolerated use in the next section.[202]

All this, however, will not obviate the need for some level of human oversight. Since fair use determinations include qualitative variables—such as whether the "heart of the [copyrighted work]" was taken as part of the third factor,[203] the benefit that accrues to the public from the use as part of the first factor, and the market effect of the use as part of the fourth factor[204]—some degree of human oversight is inevitable with respect to cases falling into Category C, at least in this stage.[205]

---

200. Parchomovsky & Goldman, supra note 80, at 1511–17. For example, the contemplated ceiling for *literary works* would be the lesser of fifteen percent or three hundred words total (of both the original and the copy). For *sound recordings and musical compositions*, the lesser of ten percent or ten seconds cumulatively. For *audiovisual works,* the lesser of thirty seconds or ten percent of the original work, as well as display of an architectural, choreographic, or pictorial work for thirty seconds or less, provided that those thirty seconds comprise no more than ten percent of the new work. Id.

201. See id. at 1520 ("[T]he implementation of safe harbors is unlikely to significantly chill the incentive to create new books, songs, and film. The safe harbors we propose should have a rather minimal effect on the revenues of most copyright owners.").

202. See infra note 212 and accompanying text.

203. Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 564–65 (1985).

204. Another important example is identifying parodies. A parody is the "use of some elements of a prior author's composition to create a new one that, at least in part, comments on that author's works." Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 580 (1994); see also Dr. Seuss Enters., L.P. v. Penguin Books USA, Inc., 109 F.3d 1394, 1401 (9th Cir. 1997) (rejecting defendants' argument that book was both satire and parody); Lucasfilm Ltd. v. Media Mkt. Grp., Ltd., 182 F. Supp. 2d 897, 901 (N.D. Cal. 2002) (denying preliminary injunction against pornographic animated version of Star Wars-based parody).

205. This does not mean that technology could not surprise us in the future. Garry Kasparov famously asserted that a computer would never be able to defeat a Grandmaster at chess because it lacks intuition and a strategic line of thinking. Yet Kasparov himself lost to "Deep Blue," when the more advanced computer effectively traded intuition and strategy for millions of calculations. Feng-Hsiung Hsu, Behind Deep Blue: Building the Computer That Defeated the World Chess Champion 92 (2002).

Finally, our proposal leaves intact the notice and takedown procedure and a reverse notice and takedown mechanism for users.[206] Content owners who identify putatively infringing content will be able to notify webhosts about the presence of the content and require that it be removed—precisely as they can under section 512. Hence, under our proposal, content owners are not worse off than they are now. Likewise, users will be expeditiously notified about blocking and removal actions and will be able to demand that a post be reinstated, in return for waiving their anonymity.[207]

Adopting our proposal will not leave fair use worse off than it is today. Under our model, filtering developers would internalize the decision of whether to refer materials to human review or not. Because technology providers will not want their effectiveness ranking to be lowered, they will be careful not to filter out content that has a legitimate claim for fair use.[208] On the other hand, cost and efficacy considerations would push developers to automate more decisions over time as long as accuracy would not be reduced. On the whole, filtering system manufacturers have a strong incentive to respect fair use and improve the ability of their technologies to handle fair use cases.

B. *Efficient Tolerated Uses*

A second challenge for our proposal involves tolerated uses. The term was coined by Tim Wu to refer to uses of copyrighted works that are technically infringing, yet are consciously not enforced by their copyright owners.[209] Wu further explained that there are various reasons why tolerated uses are not being enforced, ranging from simple laziness to the recognition by the owner that a particular use might be harmless or even beneficial to her.[210]

The subgroup of tolerated uses that inflicts *no harm* on content owners creates more benefit than harm (because these uses benefit the users who perform and access them, and create no harm to the copyright owner) and is therefore desirable from a social standpoint. Accordingly, it is important to allow such uses to exist. Efficient tolerated uses are actually dealt with quite effectively under the current regime, because when owners avoid sending takedown notices regarding such uses, the content

206. 17 U.S.C. § 512(g) (2006).

207. Id. § 512(g)(3); see also Perfect 10, Inc. v. CCBill LLC, 488 F.3d 1102, 1112 (9th Cir. 2007) ("Accusations of alleged infringement have drastic consequences: A user could have content removed, or may have his access terminated entirely. If the content infringes, justice has been done. But if it does not, speech protected under the First Amendment could be removed.").

208. See supra Part II.B. (explaining how effectiveness is measured).

209. Wu, Tolerated Use, supra note 53, at 617.

210. Id. at 619 ("Reasons can include simple laziness or enforcement costs, a desire to create goodwill, or a calculation that the infringement creates an economic complement to the copyrighted work—it actually benefits the owner.").

typically remains on the website.[211] Our model, in contrast, would treat all technically infringing contents indiscriminately—and may ultimately block even efficient tolerated uses.

Admittedly, we only have partial solutions to the efficient tolerated use problem. One solution is built into the system: Content owners could decide to tolerate infringements by simply refraining from including information about their works in the central database that informs the screening process. Yet, the ex ante nature of such decisions would still not enable owners to tolerate ex post some uses of works they generally want protected.

Furthermore, content owners may adopt intermediate positions that would permit certain types of uses, or permit users to appropriate certain percentages of their works.[212] For example, a content owner may decide that copying of up to twenty percent of her work would be exempt from liability and set her screening preferences accordingly. The owner can further decide to set differential tolerance levels for different works. For example, for newly released works, set the filter automatically to block any posting that exceeds seventy percent correspondence of audio and video, while for older material the tolerance setting might rise to ninety percent. This can salvage most tolerated uses. Nevertheless a problem will remain with respect to "mash-ups." Mash-ups are video clips that contain original visual content with copyrighted music in the background. And such content clearly exists.

A possible way to ameliorate this problem could be to inform content owners of all blocked uses—immediately or periodically—and allow them to override the filter's decision and "unblock" the material at any time.[213] Admittedly, this is still a partial solution, because copyright owners may opt in to the default "blocked" status merely because this is the default or because they do not internalize the benefit from this use.

We concede that our proposal does not effect a Pareto optimal solution—one that makes at least one person better off while leaving no one worse off. Yet, there is nothing surprising about that. As Guido Calabresi powerfully demonstrated, very few legal regimes, if any, satisfy the Pareto principle.[214] Virtually all legal changes create winners and losers. Hence, we are not daunted by the fact that our proposal will not make everyone

---

211. As discussed at length in Part I, under the current notice and takedown regime, webhosts have no incentive to remove infringing content without the request of the copyright owner.

212. See Robert Merges, To Waive and Waive Not: Property and Flexibility in the Digital Era, 34 Colum. J.L. & Arts 113, 126–27 (2011) (suggesting "selective waiver" mechanism, in which copyright owners would waive their rights in certain categories).

213. Such a mechanism would be similar to YouTube's current mechanism. This mechanism flags to content owners when their material appears on the site and allows them to choose whether to remove it or share with Google the ad revenues resulting from it.

214. Guido Calabresi, The Pointlessness of Pareto: Carrying Coase Further, 100 Yale L.J. 1211, 1212 (1991).

better off. This cannot be the test for making legal changes. A much more useful test is to compare the benefits of our proposal with the cost thereof. Our proposal fares well on this criterion for three reasons.

First, it is important to remember that not all tolerated uses are efficient; many tolerated uses are actually inefficient, namely, overall they create more harm than benefit.[215] The level of tolerated uses is to a large extent a function of the cost of enforcement—a costly, ineffective enforcement process, like the one that exists today, may prevent enforcement against harmful infringers.[216] In the absence of empirical data regarding the nature of tolerated uses, particularly the reasons why they are tolerated, it is impossible to know whether our model comprises an overall superior or inferior policy to handle tolerated uses.

Second, the attraction of tolerated uses is to a large extent a false allure. Tolerated uses are an elusive concept. A copyright owner may tolerate a use for a certain period of time, but the owner may decide to start vigorously enforcing its rights at any time.[217] Indeed, as the number of unauthorized uses made of a certain copyrighted work goes up, it becomes more likely that the owner of the copyright will commence legal action against the infringers. This stems from the simple fact that the owner's decision whether to "tolerate" a use is not only a function of the cost of enforcement but also of the expected value she can derive from the lawsuit.

Finally, and most importantly, the societal interest in protecting webhosts is far more important than the interest in maintaining all currently available tolerated uses. The degree to which webhosts are exposed to liability affects all users and content owners. It is a key determinant of the future of the Internet. Tolerated uses, while important, present a relatively minor issue in comparison to webhosts' liability and the need for a balanced system of copyright enforcement.

C. *The Challenge of Improvements*

Naturally, the competition we envision in the market for filtering technologies is expected to constantly yield new and better filtering applications. In light of this expectation and the adaptive nature of the best available technology standard, one may wonder whether our technological safe harbor will ever shelter webhosts. After all, it is possible that newer and slightly better filtering systems will appear on the market on a monthly basis and their appearance may result in a constant reranking of technologies, with new filtering technologies pushing older ones off the best available technology list.

---

215. See Wu, Tolerated Use, supra note 53, at 619 (noting possible reasons for tolerated uses).

216. See supra Part I.A (discussing inefficiencies in current enforcement regime).

217. See Edward Lee, Warming up to User-Generated Content, 2008 U. Ill. L. Rev. 1459, 1486–88 (2008) (noting owners sometimes use a "hedging" strategy—waiting to see whether use is harmful or beneficial before taking any enforcement measures).

This challenge would arise with respect to any adaptive legal standard, especially those involving technology, because technology improves at an especially fast pace.[218] Although the problem of improvements may appear daunting at first blush, it is much less formidable upon closer inspection. In fact, a simple way to deal with this problem is to base the list of the "best available technologies" on periodic reviews of the technology, which will be revised and updated.[219] The intervals between updates will be determined based on the rate at which new filtering systems are produced and the length of the examination period necessary to establish the effectiveness of new technologies. Given that there will only be several clearinghouses and that improvements to filtering technologies are not easy to produce, the cost of updating the list of best available technologies should not be too high. It should be borne in mind that other software applications, games, electronics, and other consumer goods are routinely reviewed, evaluated, and compared to rival products.[220] Such reviews often involve multidimensional analysis. In our case, by contrast, the comparison focuses on two dimensions alone, effectiveness and cost, which should expedite the review process.

Once a periodic review is completed, the filtering centers that need to replace their technology will be given a certain grace period to do so. During that period, all webhosts that referred content to the filtering centers will be fully protected from liability.

IV. Possible Extensions of the Best Available Technology Standard

Webhosts are not by any means the only entities that are exposed to copyright liability by dint of the actions of their users. Copyright liability poses a critical threat to other internet businesses and entities, many of which may also benefit from our proposal. This Part explores the possibility of extending our model to operators of other digital platforms and internet actors. In particular, we examine the applicability of our model to operators of peer-to-peer file sharing platforms—who enjoy no safe harbor under current law—and access providers—for whom our model could potentially replace the safe harbor they currently enjoy under section 512(a).

---

218. Adaptive standards are open ended and allow the flexibility to leave certain areas open for future revisions. Adaptive standards are not foreign to copyright law. The "tangible medium of expression," for example, was designed to "avoid the artificial and largely unjustifiable distinctions . . . under which statutory copyrightability in certain cases has been made to depend upon the form or medium in which the work is fixed." H.R. Rep. No. 94-1476, at 52 (1976), reprinted in 1976 U.S.C.C.A.N. 5659, 5665.

219. Updates to the software might occur between the periods.

220. Technology journals and even mainstream newspapers typically contain sections that compare and review new products on the market.

A. *Applying the Technological Safe Harbor to Peer-to-Peer Services*

Peer-to-peer systems are distinct from webhosts in that they do not store users' materials on their websites, but rather enable users to access materials that reside on other users' hard drives.[221]

Peer-to-peer file sharing services are not explicitly covered under section 512.[222] This should be no surprise, given that the first widespread file sharing service, Napster, was only created after the enactment of the DMCA.[223] Inimical to peer-to-peer services, courts have thus far declined to read section 512 to protect these services for various reasons.[224]

Without the protective wings of the DMCA that exempts webhosts from using filters (in the absence of a "red flag"),[225] peer-to-peer services have been repeatedly criticized when they refrain from deploying filtering mechanisms. Most notably, in *Grokster*, the Supreme Court viewed file sharing services' failure to install filtering devices as evidence of their unlawful intent, which in turn tipped the scales towards holding them secondarily liable for inducing a copyright infringement.[226] At the same time, the *Grokster* Court itself clarified that lack of filtering cannot by itself suffice to establish secondary liability.[227] Accordingly, the precise effect of

---

221. See A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004, 1010–11 (9th Cir. 2001) (explaining how peer-to-peer systems operate).

222. See 17 U.S.C. § 512 (2006) (failing to specify peer-to-peer services).

223. Helman, Secondary Liability, supra note 21, at 134 (noting that excluding newer technologies from section 512 "does not reflect a deliberate decision, but rather the obvious lack of predictive powers").

224. See, e.g., In re Aimster Copyright Litig., 334 F.3d 643, 655 (7th Cir. 2003) ("[Section 512] provides a series of safe harbors . . . but none in which Aimster can moor."); Columbia Pictures Indus., Inc. v. Fung, No. CV 06-5578 SVW(JCx), 2009 WL 6355911, at *15–*18 (C.D. Cal. Dec. 21, 2009) (debarring section 512 from file-sharing service defendant for number of reasons). Peer-to-peer service providers thus incurred secondary liability in a number of high profile cases. See *Columbia Pictures*, 2009 WL 6355911 at *19 (granting plaintiffs' summary judgment motion under inducement of infringement theory of secondary liability); see also Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 545 U.S. 913, 930 (2005) (noting while Copyright Act does not expressly render anyone liable for infringement committed by another, "doctrines of secondary liability emerged from common law principles and are well established"); *Napster,* 239 F.3d at 1025 ("At this stage of the litigation, plaintiffs raise serious questions regarding Napster's ability to obtain shelter under § 512 . . . ."); Arista Records LLC v. Lime Grp. LLC, 715 F. Supp. 2d 481, 508–09, 513 (S.D.N.Y. 2010) (applying *Grokster* to find defendant liable for inducement of copyright infringement and noting "[Lime Wire]'s failure to mitigate infringing activities"); Lemley & Reese, supra note 66, at 1369–72 (noting peer-to-peer systems often fall outside of section 512, either because they do not qualify as ISPs or because they do not fall under any of four categories to which safe harbors apply).

225. 17 U.S.C. § 512(m).

226. *Grokster*, 545 U.S. at 939. Justice Breyer was reluctant to draw conclusions from the absence of filtering. Id. at 958 (Breyer, J., concurring).

227. Id. at 939 n.12 (majority opinion) ("[I]n the absence of other evidence of intent, a court would be unable to find contributory infringement liability merely based on a failure to take affirmative steps to prevent infringement, if the device otherwise was capable of substantial noninfringing uses."). The Court's caution might have stemmed from its attempt to remain consistent with Sony Corp. of America v. Universal City Studios,

lack of filtering on the imposition of liability on grounds of inducement remains unclear. Tim Wu has interpreted *Grokster* to suggest that while failure to filter may not prove bad intent in itself, filtering may negate a finding of bad intent, effectuating a "safe harbor" from inducement claims.[228] A more attenuated approach is that filtering may serve as an inconclusive indication of good faith, because filtering does not go along with an infringement-oriented business plan.[229]

It is equally unclear under present law what would be considered adequate filtering. In *A&M Records, Inc. v. Napster, Inc.*, the Ninth Circuit held that the goal "is to get [infringements] down to zero."[230] In contrast, on remand, the district court in *Grokster* ordered the defendant to use a filter that meets "the most effective means [to prevent infringements]" standard.[231]

We believe that our technological safe harbor regime represents a better way to go forward. As with webhosts, the model would protect peer-to-peer services from all types of copyright liability so long as they employ the best available technology for filtering in their file sharing software.

Inc., 464 U.S. 417, 442 (1984), where an analogous device to a filter, the "broadcast flag," was not required. See Meng Ding, Note, *Perfect 10 v. Amazon.com*: A Step Toward Copyright's Tort Law Roots, 23 Berkeley Tech. L.J. 373, 384–85 (2008) ("The broadcast flag was, in large part, analogous to the 'filtering mechanism' . . . in *Grokster*. Since the *Sony* Court did not require it, it is hard for the *Grokster* Court to require it without disturbing *Sony*."). As a comparison, an Australian decision required installation of filtering mechanisms in peer-to-peer services. See Universal Music Austl. Pty Ltd. v Sharman License Holdings Ltd. (2005) 220 ALR 1, 5–7, 74.

228. See Tim Wu, The Copyright Paradox, 2005 Sup. Ct. Rev. 229, 247 ("*Grokster* creates a kind of safe harbor . . . . It may be read to suggest that a product that *does* filter is presumptively not a product that is intended to promote infringement, even if it does, in practice, facilitate infringement.").

229. See Monotype Imaging v. Bitstream, 376 F. Supp. 2d 877, 888–89 (N.D. Ill. 2005) (finding no inducement because, inter alia, defendant had taken measures to avoid infringing uses).

230. 284 F.3d 1091, 1097 (9th Cir. 2002) (quoting district court opinion). The district court ordered Napster to switch its current filter—which was based on file names, and was thus susceptible to circumvention of file names by users—to a superior audio fingerprinting technology. Id. The Ninth Circuit upheld this order. Id. at 1098–99 ("Napster's original filtering mechanism was unsuccessful in blocking all of plaintiffs' noticed copyrighted works . . . . It was a proper exercise of the district court's supervisory authority to require use of the new filtering mechanism . . . ."); see also Mark A. Lemley & Philip J. Weiser, Should Property or Liability Rules Govern Information?, 85 Tex. L. Rev. 783, 802 (2007) (noting that *Napster*'s "order effectively required Napster to shut down, which indeed it did").

231. Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 518 F. Supp. 2d 1197, 1241 (C.D. Cal. 2007) (ordering defendant to apply "the most effective means available to reduce the infringing capabilities"). The court recognized that "perfect" filtering was unattainable. Id. at 1235 ("[Plaintiffs] would require [the service] to shut down until it was capable of installing a 'perfect' filter that could prevent any infringement from occurring. Yet, the undisputed evidence currently indicates that there is no filtering mechanism that can 'exhaustively' stop every single potential infringement on a peer-to-peer network . . . .").

A technological safe harbor regime would create better incentives for the operation of legal peer-to-peer services: It would incentivize copyright owners to cooperate with file sharing services instead of suing them and driving them to bankruptcy.[232] In addition, it would incentivize file sharing services to internalize the costs of infringement.[233] Finally, it would open an additional market for filtering technologies—that of peer-to-peer services. The formalization of a technological safe harbor would also clear the uncertainty that currently shrouds the status of filtering in the context of peer-to-peer services.[234]

There is yet another reason to establish a technological safe harbor for operators of peer-to-peer platforms: Affording a safe harbor to webhosts but not to file sharing services will result in a policy that favors one kind of platform over another, and consequently discourage the creation of legal peer-to-peer services.[235] This would be regrettable, because peer-to-peer services comprise a rapid, resource-efficient, and overall positive tool for information sharing.[236]

One crucial factor, however, renders our safe harbor regime considerably more appealing when applied to webhosts than to peer-to-peer services. By contrast to the webhosting context, in the peer-to-peer world users typically do not need to be in continuous contact with the service provider in order to share files after downloading the file sharing software.[237] Thus, if a peer-to-peer service shifts to an improved filter, as the technological safe harbor demands, the new filter would only apply to the service's new customers (or those who continually update their software), and would not affect the software that is already in use. Users who use the old software would still be using an outdated filter. In this sense, the dynamic nature of our model is likely to be substantially weaker in the peer-to-peer framework.

---

232. See, e.g., supra note 224 (describing several cases in which courts have declined to read 17 U.S.C. § 512 (2006) to protect peer-to-peer services).

233. The current regime creates no incentive for peer-to-peer services to avoid infringement, because no consistent, realistic standard for filtering applies to them. See supra notes 230–231 and accompanying text.

234. See supra notes 230–231 and accompanying text.

235. See also Lemley & Weiser, supra note 230, at 831 (arguing "irrational, technology-based distinctions create distortions, unfair competitive advantages, and arbitrage opportunities").

236. See Elkin-Koren, supra note 3, at 19–25 (describing how peer-to-peer networks promote efficiency and freedom); Lital Helman, When Your Recording Agency Turns into an Agency Problem: The True Nature of the Peer-to-Peer Debate, 50 IDEA 49, 87–88 (2009) (describing benefits of file sharing).

237. Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 545 U.S. 913, 921–22 (2005).

B. *Applying the Technological Safe Harbor to Access Providers*

The potential liability of Internet Service Providers (ISPs)[238] for infringing content passing through their system is governed by section 512(a). The section establishes a safe harbor for ISPs, which shelters them from monetary liability for copyright infringement.[239] Like webhosts, ISPs are also exempted under the existing regime from the need to engage in filtering.[240] To qualify under the safe harbor established by section 512(a), ISPs must be neutral and refrain from interfering in the transmission of content through their pipelines beyond what is necessary for routing it.[241] For reasons mostly unrelated to copyright enforcement, however, ISPs have begun to move away from neutrality in recent years, by deploying "intelligent" routers within their networks, which enable them to inspect and monitor the traffic they carry.[242] The more involved ISPs become in monitoring content, the less eligible they might become for the section 512(a) safe harbor.[243]

At first blush, our model seems to provide a balanced practical framework for administering ISP liability.[244] Our model would provide an

---

238. The term ISP is used at times in a wide sense, to encompass both access providers and other types of online service providers. We use the term narrowly to address only access providers, typically broadband services.

239. 17 U.S.C. § 512(a) (2006) provides:

> A service provider shall not be liable for monetary relief . . . for infringement of copyright by reason of the provider's transmitting, routing, or providing connections for, material through a system or network controlled or operated by or for the service provider, or by reason of the intermediate and transient storage of that material in the course of such transmitting, routing, or providing connections . . . .

The safe harbor is conditioned upon the ISP fulfilling the statute's conditions. See infra note 241 and accompanying text.

240. 17 U.S.C § 512(a), (m).

241. To qualify under section 512(a), an ISP must meet five conditions: (1) the transmission must not be initiated by the ISP itself; (2) the transmission must be automatic, and involve no selection of the information by the ISP; (3) the ISP must not select the recipient of the transmission; (4) the ISP must not host the information except transiently, as is necessary to transmit it; and (5) the ISP must not modify the information. Id. § 512 (a)(1)–(5).

242. See Bridy, supra note 92, at 106 (noting ISPs' transition to "intrusive traffic management or shaping"); Paul Ohm, The Rise and Fall of Invasive ISP Surveillance, 2009 U. Ill. L. Rev. 1417, 1432–37 (describing invasive new monitoring by ISPs); see also Peter K. Yu, The Graduated Response, 62 Fla. L. Rev. 1373, 1387 (2010) (exploring possibility of ISPs undertaking inspection and monitoring).

243. Rob Frieden, Internet Packet Sniffing and Its Impact on the Network Neutrality Debate and the Balance of Power Between Intellectual Property Creators and Consumers, 18 Fordham Intell. Prop. Media & Ent. L.J. 633, 645 (2008) ("The output of affordable deep packet inspection . . . raises questions whether non-neutral network operation disqualifies ISPs for a safe harbor exemption from liability for carrying copyright infringing traffic provided by § 512 of the DMCA.").

244. This discussion assumes that ISP filtering would be legal, either as "reasonable network management" or otherwise. Indeed, although filtering by ISPs may run into problems with net neutrality principles, see, e.g., Ohm, supra note 242, at 1492, it is

alternative to the diminishing section 512(a); one that would allow ISPs to conduct some controlled surveillance of copyright infringements, stopping it at the source, yet would include safeguards to assure blocking is not excessive. What is more, the issue of filtering by ISPs has been on the global public agenda for some time now.[245]

Yet, in our opinion, applying our system to ISPs will yield considerably more harm than benefit. First, surveillance of users' online behavior by ISPs is likely to have implications beyond the copyright realm. Chief among those are innovation policy,[246] privacy,[247] censorship,[248] and others.[249] Applying our system to ISPs would encourage ISPs to engage in surveillance from the narrow prism of copyright law, without attending to all the complicated issues this measure might imply beyond.

Second, despite the similar result for the user whose content gets blocked, filtering by ISPs might prove far more hazardous than filtering

---

estimated that the FCC is not likely to prevent ISPs' monitoring of networks for copyright-related reasons. See Bridy, supra note 92, at 132 ("The FCC . . . is unlikely to intervene in the name of net neutrality to prevent this private (re)ordering by, for example, prohibiting content blocking or filtering by ISPs."). Note, however, that an attempt to explicitly include copyright filtering as legitimate network management failed. Compare 155 Cong. Rec. S1738 (daily ed. Feb. 5, 2009) (showing text of proposed amendment 417, referring to deterring copyright infringement), with American Recovery and Reinvestement Act of 2009, Pub. L. No. 111-5, 123 Stat. 115 (failing to include any reference to copyright infringement).

245. See, for instance, the Anti-Counterfeiting Trade Agreement (ACTA) (Proposed Draft 2010), available at http://www.ustr.gov/webfm_send/2417 (on file with the *Columbia Law Review*), which might pave the way for ISP filtering in the United States, the European Union, Australia, Japan, Canada and a number of other nations.

246. See Christopher S. Yoo, Network Neutrality and the Economics of Congestion, 94 Geo. L.J. 1847, 1851 n.13 (2006) (noting "network neutrality proponents defend their proposals almost exclusively in terms of the economic benefits of innovation"); see generally Tim Wu, Network Neutrality, Broadband Discrimination, 2 J. on Telecomm. & High Tech. L. 141 (2003) (discussing regulation of broadband providers).

247. See Ohm, supra note 242, at 1432–37 (describing aggressive expansion of network monitoring and predicting that "will eviscerate user privacy"); see also Electronic Communications Privacy Act (ECPA), 18 U.S.C. §§ 2510–2712 (2006) (addressing interception of and access to wire and electronic communications).

248. See Frank Pasquale, Beyond Innovation and Competition: The Need for Qualified Transparency in Internet Intermediaries, 104 Nw. U. L. Rev. 105, 119–24 (2010) (discussing potential bias when carriers can influence content); Kevin Werbach, The Centripetal Network: How the Internet Holds Itself Together, and the Forces Tearing It Apart, 42 U.C. Davis L. Rev. 343, 367 (2008) (opposing filtering because of censorship's threat "to the structure and universality of the Internet itself"); Letter from Open Internet Coal. to Cong. (June 29, 2009), available at http://www.openinternetcoalition.org/files/OIC_DPI_Iran_062909.pdf (on file with the *Columbia Law Review*) (describing repressive regimes' use of deep packet inspection for spying and censorship and warning of potential misuse in United States).

249. See, e.g., Pasquale, supra note 248, at 108 (pointing to "a new host of concerns about privacy, culture, and power online"); see also Comcast Corp. v. FCC, 600 F.3d 642, 644 (D.C. Cir. 2010) (involving Comcast's monitoring of BitTorrent traffic, and concluding FCC lacks authority to enforce network neutrality rules over broadband internet providers).

by webhosts. ISP filtering would involve scrutinizing *all* internet traffic in real time, and necessarily degrade the performance of the system.[250] Similarly, contrary to the free business model of webhosts, the costs associated with filtering by ISPs (e.g., purchasing and maintaining filtering hardware and software) are likely to be passed on to subscribers, decreasing the affordability of broadband services. Worse yet, the lack of transparency that exists regarding ISPs' practices[251] and the lack of meaningful alternatives and competition among ISP services[252] render ISPs much more likely to surrender users' interests in the face of pressures from content owners concerning the operation of filters.[253]

It appears that filtering by ISPs involves substantially more complex considerations than filtering by webhosts. This, of course, does not imply that copyright law should not be a factor in the analysis. On the contrary, we believe that it is an important consideration. Yet, in this case, it should be balanced against potentially more weighty considerations. Hence, in so far as establishing a technological safe harbor for ISPs is concerned, we believe we should proceed with caution. That said, we suggest that if ISPs would ever be allowed to conduct filtering for copyright purposes, then the standard of best available technology should act as a safeguard to ensure that the incentives are in place to guarantee that users' rights are protected.

Conclusion

This Article proposes a technological safe harbor that would protect webhosts from monetary liability for copyright infringement. Per our proposal, webhosts would be immune from liability as long as they employ the best filtering technology available on the market. We demonstrate that adopting our proposal would yield several important advantages relative to the existing legal regime. First, it would give webhosts the legal certainty they need in order to grow and develop. Second, it would substantially improve effective and balanced copyright enforcement, by creating incentives for copyright owners, webhosts, and technology providers to collaborate towards an optimal scheme of copyright enforcement. Furthermore, the design of our proposal would dramatically reduce litiga-

250. Tim Wu, Has AT&T Lost its Mind?, Slate (Jan. 16, 2008, 10:15 AM), http://www.slate.com/id/2182152 (on file with the *Columbia Law Review*) (responding to AT&T's proposal to examine all traffic carried on its network).

251. Preserving the Open Internet, Broadband Industry Practices, 74 Fed. Reg. 62,638, 62,640 (Nov. 30, 2009) (to be codified at 47 C.F.R. pt. 8) (noting ISPs generally do not disclose their network management practices).

252. See Ohm, supra note 242, at 1476 (noting "[i]n most parts of the United States, the only two choices [for internet connectivity] are DSL from the telephone company and a cable modem from the cable company").

253. See Pasquale, supra note 248, at 119–23 ("Given that most consumers have only one or two options, if any, for broadband connectivity, network providers can easily use their services to subtly advance their own political or cultural agendas without much fear of losing customers.").

tion in this area of the law and completely obviate lawsuits, such as *Viacom v. YouTube*. Finally, the proposal, if implemented, would promote dynamic efficiency by spurring competition in the market for filtering systems. Competition in the market for filtering technologies will lead to more accurate detection of infringement and as a result, would enhance balanced copyright enforcement over time.

At the end of the day, legal challenges that are born out of technological advancements often call for technological solutions. The challenge of online infringement is a case in point in our opinion. Given the exponential rate at which new content—infringing and noninfringing—is posted onto websites and the fact that creativity in digital media is largely based on preexisting materials, webhosts cannot possibly detect every single case of copyright infringement. Nor is it socially desirable to require them to screen content manually. The imposition of such a requirement would dramatically increase operation costs for all webhosts, which may lead to the demise of many small websites. Accordingly, filtering technologies appear to mark the only viable way to move forward.

This realization is not ours alone; it is shared by leading industry participants who have adopted variants of our proposal through private agreements. Given the generality of the problem, transaction costs, and pressing need for a solution, we believe that legislative intervention is in order. The blueprint we provide in this Article may aid lawmakers in crafting innovative legal solutions not only to the challenge of webhosting, but also to the ongoing struggle of file sharing via the peer-to-peer platform, and may even inform some changes in the policy concerning the liability of ISPs. In sum, it provides a general analytical framework for reshaping online liability.