TAKEDOWN IN TWO WORLDS: AN EMPIRICAL ANALYSIS

by JENNIFER M. URBAN, BRIANNA L. SCHOFIELD, *and* JOE KARAGANIS*

ABSTRACT

In its nearly two decades of existence, the Digital Millennium Copyright Act's "notice and takedown" process has become both a source of controversy and a primary means of addressing online copyright disputes. Yet there is little empirical research into the use of notice and takedown or its effectiveness. Because it depends on the private communications and actions of copyright holders, online service providers, and targets, notice and takedown operates largely within a difficult-to-study "black box." This article contributes to the literature with an empirical look at two sets of takedown notices. It digests findings from two quantitative studies of takedown

*Jennifer M. Urban is a Clinical Professor of Law at the University of California, Berkeley School of Law. Brianna L. Schofield is the executive director of Authors Alliance and a former Clinical Teaching Fellow at the Samuelson Law, Technology & Public Policy Clinic at the University of California, Berkeley School of Law. Joe Karaganis is Vice President at The American Assembly at Columbia University. This is a digested and updated version of a portion of the longer research report, *Notice and Takedown in Everyday Practice*. This research is a part of the Takedown Project, a collaborative effort led by Urban and Karaganis to develop a global network of researchers studying notice and takedown procedures in order to advance understanding of how these procedures resolve copyright and other conflicts and how they mediate freedom of expression online. More information can be found on the project's website, https://takedownproject.org.

This work would not have been possible without both data and funding resources for the coding effort. We thank Adam Holland and Wendy Seltzer of Lumen (formerly Chilling Effects) for facilitating access to the Lumen data. We are grateful for funding support from Google Inc. as a gift to The American Assembly and from the Sloan Foundation for its support through the Berkeley Law Digital Library Copyright Project. Neither funder directed our approach in any way, and neither funder reviewed any methods, data, results, or reporting before public release. We are also grateful for Nash Information Services' in-kind donation of the OpusData database and expert database customization advice.

We are indebted to the individuals who generously lent their time, skill, and expertise to the project: Kristoff Grospe for his assistance with the development and management of the studies; Martyn Joyce for building and maintaining our customized database, coding and querying interfaces, and randomization and search algorithms; Bruce Nash of Nash Information Services for help with database and algorithm design; and Nora Broege of UC Berkeley's D-Lab for data preparation and statistical analysis. We are very grateful for the detailed and helpful comments on study design, findings, and report drafts offered by Annemarie Bridy, Niva Elkin-Koren, Gwen Hinze, Martin Husovec, Deirdre Mulligan, Bill

notices in the Lumen database, reported fully in the three-study report, Notice and Takedown in Everyday Practice.[1]

The first study, of notices sent to Google Web Search, shows a heavy reliance by large rightsholders on automated infringement detection and notification, and tends to support their assertions that they focus their enforcement resources on large-scale infringement. At the same time, 31% of these notices raised questions related to their accuracy and statutory compliance. The second study provides a snapshot of takedown as used by one set of smaller notice senders — individuals and small businesses sending to Google Image Search. In this study, notices were more likely to target social media sites, blogs, and personal websites, and 70% of the notices raised questions. While both studies revealed surprisingly high percentages of notices of questionable validity, the questions raised tended to differ. The automated notices in the first study predominantly exhibited issues identifying the works in question, while the notices sent by smaller rightsholders in the second study predominantly exhibited issues with the underlying claim. Accordingly, while notice sending appears to need improvement, there is no "one-size-fits-all" approach. Both "worlds" of notice sending need to be accommodated. We therefore suggest a set of voluntary best practices, educational efforts, and modest legal reforms directed to the different issues we observed.

Rosenblatt, and participants in multiple Takedown Project workshops, the 3rd Global Congress on Intellectual Property Rights and the Public Interest, the 4th Global Congress on Intellectual Property Rights and the Public Interest, the 18th Annual Berkeley Center for Law and Technology and Berkeley Technology Law Journal Symposium "The Next Great Copyright Act," the Department of Commerce Multistakeholder Forum on Improving the Operation of the DMCA Notice and Takedown System, the Chicago-Kent College of Law's Conference on Empirical Research on Copyright Issues, and the Copyright Society of the USA's Copyright and Technology Conference.

We also thank the exceptional team of graduate students from the University of California, Berkeley School of Law who manually reviewed takedown requests: Sérgio Alves, Jr.; Christina Farmer; Shaina Hyder; Alicia Intriago; Tigist Kassahun; Shweta Kumar; Leighanna Mixter; Smita Rajmohan; and Kirsty Watkins. Lydia Anderson-Dana provided excellent research assistance in the final stages of the project. Christian Chessman of the *Berkeley Technology Law Journal* provided valuable comments and editorial guidance on this version of the work.

Mistakes are ours alone.

[1] For a full description of the three studies, see JENNIFER M. URBAN, JOE KARAGANIS & BRIANNA L. SCHOFIELD, NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, (2016) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2755628.

TABLE OF CONTENTS

I. INTRODUCTION

In 1998, Congress passed the Digital Millennium Copyright Act[2] ("DMCA"), a then-obscure law that nonetheless had broad implications for copyright holders, Internet users, and online service providers ("OSPs"). Section 512 of the DMCA gave OSPs "safe harbor" from certain secondary liability for their users' copyright infringement. In return,

[2] The On-Line Copyright Infringement Liability Limitation Act (OCILLA) — commonly known as § 512 of the DMCA — is codified as Title II of the DMCA at 17 U.S.C. § 512 (2012). Title I of the DMCA covers the separate topic of anticircumvention.

OSPs are required to implement various practices to protect copyright holders, most notably "notice and takedown" procedures. Under these procedures, copyright owners can send short "takedown notices" to USPS requesting that infringing materials be removed from the OSP's sites. Users can challenge removal through a "counter notice" process that gives the target of a notice the ability to respond and request "putback" by the OSP. The process avoids the expense of lawsuits unless a user challenges with a counter notice, in which case the rightsholder can continue the claim by filing an infringement suit.

In developing the safe harbor regime, Congress intended for it to "provide strong incentives for service providers and copyright owners to cooperate to detect and deal with copyright infringements."[3] To achieve that end, Congress divided the burdens of compliance between OSPs and copyright owners. Congress placed on OSPs the burden of responding to valid takedown notices by "expeditiously" removing or disabling access to the identified allegedly infringing content. Congress placed on copyright holders the burden of identifying infringing material because it considered that they know what material they own, and "are thus better able to efficiently identify infringing copies than service providers . . . who cannot readily ascertain what material is copyrighted and what is not."[4]

Legal and policy arguments over § 512's effectiveness and fairness have been ongoing since its inception and center on a few key issues that allocate responsibilities and costs between the parties. First, arguments over effectiveness have centered on whether its notice and takedown procedures can scale up to address determined, large-scale piracy. Second, various stakeholders have advanced arguments over fairness. Rightsholders have challenged the requirement that they identify individual infringements and have pushed in court for more expansive readings of when an OSP's level of knowledge about infringement or its activities take it out of the safe harbor. OSPs have, largely successfully, defended themselves in court by arguing that rightsholders must identify specific instances of infringement when sending notices. Finally, arguments on behalf of target interests have centered on whether the statute sufficiently manages possible overbreadth in takedown.

These disputes arise in part because § 512's drafters, acting in 1998, were unable to anticipate the Internet's massive growth or the tremendous

---

[3] S. REP. No. 105-190, at 20 (1998); H.R. REP. No. 105-551, pt. 2, at 49 (1998).

[4] S. REP. No. 105-190, at 48 (1998); H.R. REP. No. 105-551, pt. 2, at 57-58 (1998). Courts have since affirmed that the DMCA notice and takedown provisions follow longstanding copyright law by "plac[ing] the burden of policing ongoing copyright infringement — identifying potential infringing material and adequately documenting infringement — squarely on the owners of the copyright." Perfect 10, Inc. v. CCBill LLC, 488 F.3d 1102, 1113 (9th Cir. 2007).

technological and social changes that took hold in the ensuing years. At the time, search engines were relatively new, and Yahoo! had the greatest market share; Google was launched the year § 512 was passed. Peer-to-peer networks were not anticipated — Napster launched in 1999 — yet quickly became preferred technologies for copyright infringement. Facebook and other social networks followed several years later. These social networks, along with YouTube and other platforms, have emerged as both dominant content purveyors and major speech platforms.

The nature of § 512's notice and takedown procedures has also changed for some users. Faced with large-scale infringement, often via peer-to-peer networks, large copyright holders now use automated "bots" to search for copyright violations and generate thousands or millions of automated "takedown" notices to OSPs. For these rightsholders, specialized "content protection" teams or third-party rights enforcement organizations ("REOs") commonly facilitate copyright enforcement efforts. In response, OSPs that receive automated notices automate their responses out of necessity.[5]

While automating large-scale enforcement allows some copyright owners to better police their copyrights on today's Internet, relying on machines to make decisions about sometimes-nuanced copyright law also raises questions about costs to OSPs, accuracy, and ultimately, effects on expression. At the same time, individual rightsholders without access to automated systems or third-party services may struggle to enforce their copyrights online.[6] Finally, some notice senders may misunderstand — or deliberately abuse — the takedown remedy.

Given the enormous technological changes since the law was passed, the ongoing disputes about the efficacy of notice and takedown for policing online infringement, and the criticisms of improper takedowns targeting online expression, there is a growing desire to understand how, and how well, notice and takedown operates today. A 2013 U.S. Department of Commerce task force reviewed stakeholders' experiences addressing online copyright issues, including § 512.[7] And in 2016, the Copyright Of-

[5] While this is an important shift, it does not mean that all OSPs receive automated notices or engage in automated review. To the contrary, our broader study found the opposite: despite these technological shifts, the majority of OSPs still process a manageable volume of notices and engage in substantial human review. NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 28-30.

[6] *See, e.g.*, The Arts and Entertainment Advocacy Clinic at George Mason University School of Law, Response to "Section 512 Study: Notice and Request for Public Comment" at 11 (quoting Ellen Seidler).

[7] DEP'T OF COMMERCE INTERNET POLICY TASKFORCE, COPYRIGHT POLICY, CREATIVITY, AND INNOVATION IN THE DIGITAL ECONOMY (2013).

fice began a formal study, not yet complete, to evaluate § 512.[8] Responses to the Copyright Office's Notice of Inquiry[9] provide some information on stakeholders' experiences with notice and takedown.[10]

Yet there is little available empirical research. Notice and takedown largely occurs in a "black box" — via private notices and actions by private parties — making it difficult to study from the outside. Furthermore, a heightened sense of political sensitivity and fear of liability on all sides creates reluctance to share information. Accordingly, only a few empirical reviews — Urban and Quilter (2006), Quilter and Heins (2007), and Seng's 2014 statistical inquiry[11] — exist in the research literature.[12]

To help fill this gap, we undertook three empirical studies, detailed in the report, *Notice and Takedown in Everyday Practice*:[13]

- A qualitative study that documents the ways in which the notice and takedown process has been perceived and operationalized by U.S. OSPs and large rightsholders, based on confidential surveys and in-depth interviews with nearly three dozen OSPs and notice senders ("Study 1").
- A quantitative examination of a random sample of takedown notices, taken from a set of over 108 million requests submitted to

---

8 *Section 512 Study*, COPYRIGHT.GOV, https://www.copyright.gov/policy/section512 (last visited Jan. 10, 2017).

9 Section 512 Study: Notice and Request for Public Comment, 80 Fed. Reg. 81862 (Dec. 31, 2015), https://www.copyright.gov/fedreg/2015/80fr81862.pdf.

10 *Requests for Public Comments: Digital Millennium Copyright Act Safe Harbor Provisions*, REGULATIONS.GOV, https://www.regulations.gov/docketBrowser?rpp=25&so=DESC&sb=commentDueDate&po=0&D=COLC-2015-0013 (last visited Jan. 10, 2017).

11 *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 COMP. HIGH TECH. L.J. 621 (2006); LAURA QUILTER & MARJORIE HEINS, BRENNAN CTR. FOR JUSTICE, INTELLECTUAL PROPERTY AND FREE SPEECH IN THE ONLINE WORLD: HOW EDUCATIONAL INSTITUTIONS AND OTHER ONLINE SERVICE PROVIDERS ARE COPING WITH CEASE AND DESIST LETTERS AND TAKEDOWN NOTICES (2007); Daniel Seng, *The State of the Discordant Union: An Empirical Analysis of DMCA Takedown Notices*, 18 VA. J.L. & TECH. 369 (2014).

12 There are also a few recent, but generally narrower empirical reviews of copyright takedown activity. *See, e.g.*, BRUCE BOYDEN, THE FAILURE OF THE DMCA NOTICE AND TAKEDOWN SYSTEM: A TWENTIETH CENTURY SOLUTION TO A TWENTY-FIRST CENTURY PROBLEM (2013), http://cpip.gmu.edu/wp-content/uploads/2013/08/Bruce-Boyden-The-Failure-of-the-DMCA-Notice-and-Takedown-System1.pdf; KRIS ERICKSON ET AL., COPYRIGHT AND THE ECONOMIC EFFECTS OF PARODY: AN EMPIRICAL STUDY OF MUSIC VIDEOS ON THE YOUTUBE PLATFORM AND AN ASSESSMENT OF THE REGULATORY OPTIONS (2013), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/309903/ipresearch-parody-report3-150313.pdf.

13 NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1.

the Lumen[14] archive over a six-month period (most of which relate to Google Web Search) ("Study 2").

- A further detailed quantitative examination of a random sample of notices that were sent to Google in relation to its Google Image Search service, isolated from the same six-month set of takedown requests taken from the Lumen archive ("Study 3").

In this shorter paper, we digest the main findings and analysis from the two quantitative studies of Lumen notices — Study 2 and Study 3. We do not cover the qualitative study here, but do refer to it from time to time as "Study 1" and occasionally refer to its taxonomy of "DMCA Classic," "DMCA Auto" and "DMCA Plus" OSPs.[15] A digested and updated version of Study 1, *Notice and Takedown: OSP and Rightsholder Accounts of Everyday Practice*, can be found in Volume 64, number 3, of this *Journal*.[16]

II. SUMMARY OF METHODS[17]

We analyzed a set of all takedown requests sent to the non-profit repository Lumen over a six-month period. The set comprised 288,675 notices containing well over 100 million (108,331,663) individual takedown requests — i.e., claims of infringement — typically in the form of URLs linking to allegedly infringing material. We adopted the individual takedown request — rather than the notice body, which might have many requests included within it — as the primary unit of analysis. If a Web form allows it, some senders pack many takedown requests, each of which may enforce different copyrights and target different alleged infringements, into one notice body. From the rightsholder's perspective, each request is

[14] LUMEN, https://lumendatabase.org (last visited Jan. 10, 2017). The Lumen database collects and analyizes legal complaints and takedown requests.

[15] NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 28-30. For "DMCA Classic" OSPs — the majority — the volume of notices has remained relatively infrequent despite technological changes, and substantial human review of notices is still the norm. "DMCA Auto" OSPs receive very large numbers of notices generated by automated systems, have experienced a steep increase in volume in each of the past several years, and have shifted to more automated notice-processing practices. Finally, "DMCA Plus" OSPs have adopted procedures that go beyond measures required by § 512, including filtering systems, direct takedown procedures for trusted rightsholders, hash-matching based "staydown" systems, and contractual agreements with certain rightsholders that set forth additional protections and obligations for both parties.

[16] Jennifer M. Urban, Joe Karaganis & Brianna L. Schofield, *Notice and Takedown: Online Service Provider and Rightsholder Accounts of Everyday Practice*, 64 J. COPYRIGHT SOC'Y 371 (2017).

[17] For a complete description of our methodology and study limitations, including descriptions of all variables and our material methodological decisions, see NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 78-82, 98, Appendix C, and discussion throughout Section IV.

a claim to remove one allegedly infringing item. From an OSP's perspective, each is a claim that must be acted upon. An accurate analysis thus requires reviewing individual takedown requests.

The studies have some limitations, which are more fully discussed in *Notice and Takedown in Everyday Practice*.[18] Most importantly, Lumen depends on voluntary submissions to its repository and thus cannot provide a representative sample of all the notice activity on the Internet. No such data source is available. Further, Lumen receives notices from a relative few Internet providers.[19] Moreover, Google Inc. provides the overwhelming proportion of data: it received 99.4% of the notices in our set and nearly 100% of the individual requests. Google is both a dominant company in the search market and highly unusual among OSPs in the extraordinary number of notices it receives. This cuts two ways. It suggests that notices sent to Google are an important feature of the notice and takedown ecosystem, and thus valuable to study. At the same time, it suggests that notices to Google are not necessarily representative of notices sent to other OSPs. Indeed, Study 1 suggested that the large majority of OSPs are "DMCA Classic" providers that do not engage in the "DMCA Auto" and "DMCA Plus" practices Google uses. Further, both Study 1 respondents and publicly available data suggested that Google is a clear outlier in the number of notices it receives.[20] Accordingly, the Lumen studies should be read in conjunction with Study 1, which provides a more broadly representative picture. We also made a number of choices when classifying notices and deciding how to manage missing information. Most often, we chose to err on the side of the rightsholder where judgment was required.[21]

The quantitative analysis is based on manual review and coding of two randomized samples of these requests by the Takedown Project[22] lead

---

[18] *Id.*

[19] Other organizational submitters to Lumen include or have included: Twitter, Kickstarter, Medium, Proxy.sh, Stack Exchange, Stripe, Tucows, Wikimedia Foundation, and WordPress. E-mail from Adam Holland, Project Coordinator, Lumen (July 29, 2014, 10:10 PST) (on file with authors).

[20] *See, e.g.*, NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 32, 69-73.

[21] This came up most notably whenever missing information precluded observation. Because webpage content changes over time, this decision affected a substantial number of notices, meaning that we likely undercounted potential flaws. Most importantly, in 26% of cases in Study 2, reviewers were unable to input information about the allegedly infringing material because it had been taken down or the location where it resided was no longer available. We chose to assume that takedown was appropriate in these cases unless there was contrary information, making it likely that we undercounted potential flaws in these notices.

[22] TAKEDOWN PROJECT, http://takedownproject.org (last visited Jan. 20, 2017).

researchers and a team of graduate legal researchers at the University of California, Berkeley.

Study 2 is based on a randomized sample of the entire six-month cohort of notices, the vast majority of which were automated notices sent to Google Web Search. Automated notices to Google's Web Search service were such an overwhelming feature of the overall cohort that further sampling was required to isolate other services that may exhibit different features. Accordingly, Study 3 focused on a separate randomized sample[23] of only the notices sent to Google's Image Search service.

III. MAPPING THE USE OF TAKEDOWN NOTICES AND EXPLORING OSPS' AND RIGHTSHOLDERS' ASSERTIONS ABOUT NOTICE AND TAKEDOWN

We had a number of goals for our review of takedown notices. First, we hoped to produce a general descriptive account of how notices were being used in this cohort: who was sending them, in response to what types of disputes, who was targeted, and the like. Where relevant, we hoped to compare our findings to Urban and Quilter's 2006 study — with the expectation that notice-sending practices have likely changed substantially — and with Seng's more recent study. Second, in response to the longstanding concerns about the integrity of the process expressed in the literature,[24] we were interested in whether the notices exhibited characteristics that would raise questions about their validity. Third, we wondered whether the notices would corroborate the accounts of notice and takedown we heard from OSPs and rightsholders in Study 1.

These accounts offered a number of testable assertions. First, our rightsholder respondents (all large rightsholders with valuable properties) described focusing their automated notice and takedown practice on sites that are "dedicated to infringement," and giving less human scrutiny to notices sent to such sites than they would to more general-purpose sites.[25]

---

[23] Study 3's findings are based on a random sample of 1,732 requests, drawn from the subset of 33,409 requests in the six-month dataset that were sent to Google Image Search (i.e., requests where the "notice_re line" is "Infringement Notification via Imagesearch Complaint"). This gives us an approximate margin of error of +/-2.29 at a 95% confidence interval.

[24] *See, e.g.*, Alfred C. Yen, *Internet Service Provider Liability for Subscriber Copyright Infringement, Enterprise Liability, and the First Amendment*, 88 GEO. L.J. 1833, 1888 (2000); Emily Zarins, *Notice Versus Knowledge Under the DMCA's Safe Harbors*, 92 CAL. L. REV. 257, 291-95 (2004); Wendy Seltzer, *Free Speech Unmoored in Copyright's Safe Harbor: Chilling Effects of the DMCA on the First Amendment*, HARV. J.L. & TECH. (2010).

[25] NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 34-35. Large rightsholders also described sending notices to search services, and especially to Google Web Search, as one important tactic within their overall enforce-

Second, OSPs consistently reported that the most challenging and resource-intensive aspect of takedown lay in attempting to identify allegedly infringing material on their systems in response to inaccurate or imprecise location pointers in notices.[26] Third, OSPs also expressed concerns about problems with the underlying legal claims in some notices, reporting that non-infringing material was regularly targeted.[27] OSPs noted that in their experience, one-off, first-time, low-volume, or similar "small senders" were most likely to apparently misunderstand the notice and takedown process, mistake the statutory requirements, or use it for clearly improper purposes. Given this, OSPs typically give notices from small or one-off senders more scrutiny prior to takedown, and several always subject these notices to human review.[28] In reviewing the notices, we checked the accuracy of these assertions (with the understanding that the non-representative nature of the sample meant that we could not prove or disprove these assertions for all cases).

A. *Study 2: In Six Months of Lumen Notices, an Onslaught of Automated Notices to Google Web Search*

Study 2 considered takedown requests across the entire six-month Lumen dataset. Using the methods described above, we examined the dataset in two ways. Where the source data were machine-readable, we were able to run queries across the entire six-month dataset of 108,331,663 requests.[29] For a deeper look, we hand-coded a randomized sample from the six-month dataset, containing 1,827 takedown requests,[30] from 1,766 separate notice bodies.[31]

---

ment strategies. *Id.* at 69-72. This behavior was easy to observe. *See infra* Section III.A.

[26] NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 38-40.

[27] *Id.* at 40.

[28] *Id.*

[29] We used machine-readable data across the entire six-month set for the following high-level findings: the total percentage of takedown requests submitted using a Google notice submission form; the total number of requests that came from members of Google's Trusted Copyright Removal Program; the total percentage of requests sent to Google Web Search; Table 2, Google Services Represented in Lumen; and Figure 6, Top Senders Targeting File Sharing Sites Dead More than 8 Months. All other findings are based on samples.

[30] This number includes only the completed codings and excludes notices that are not DMCA notices and those that contained some other database error (such as where the allegedly infringing material selected for coding is actually a link that the sender provided to the allegedly infringed work).

[31] Study 2 findings are based on a random sample of 1,827 requests drawn from the six-month dataset. As the full dataset contains 108,331,663 takedown requests, this sample provided a margin of error of +/- 2.29 at a 95% confidence interval, and +/-3.02 at a 99% confidence interval.

At the highest level, the Lumen data display two major features: large rightsholders' heavy use of notices to Google Web Search, and the overwhelming predominance of automated sending by these rightsholders and REOs acting for them.

- 99.8% of the takedown requests in the entire six-month dataset were requests to Google Web Search.
- 95.4% of the requests came from members of Google's Trusted Copyright Removal Program ("TCRP") — a program that allows members, usually large copyright holders or their agents, to submit large volumes of requests.[32]
- 98.9% of the takedown requests (or 86.7% of the notices) were submitted using an automated Google notice submission form.

The sheer numbers of notices to Google Web Search and the apparent high level of automation[33] confirm rightsholders' Study 1 descriptions of their practices and contrast sharply with OSPs' Study 1 descriptions of DMCA Classic processes.

The striking dominance of Google Web Search also confirms what major rightsholders told us in Study 1 about the particularized role of search (and especially Google Web Search) in their current use of notice and takedown.[34] Google submits notices for a long list of other services, including social media and cloud hosting services that might be expected to attract notices. Yet compared to Google Web Search, these other services barely register:

[32] Google's TCRP is set up for copyright owners or their enforcement agents ("trusted users") who have a "consistent need to submit thousands of URLs each day" and who have a "proven track record of submitting accurate notices." GOOGLE, HOW GOOGLE FIGHTS PIRACY 14 (2014), https://drive.google.com/file/ d/ 0BwxyRPFduTN2NmdYdGdJQnFTeTA/view.

[33] It is theoretically possible that these notices were not generated and sent automatically, but this is highly unlikely given what we learned in Study 1 about the standard industry practices of the types of senders most dominant in the sample.

[34] NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, Section III.F.

TOP 10 GOOGLE SERVICES REPRESENTED IN LUMEN (MAY–OCT 2013)	PERCENTAGE OF REQUESTS IN LUMEN (MAY–OCT 2013)
1. Google—Web Search	99.810%
2. Google—Blogger	.090%
3. Google—Image Search	.030%
4. Google—Docs	.004%
5. Google—Sites	.004%
6. Google—Picasa	.002%
7. Google—Plus (Photos)	.001%
8. Google—Plus	.001%
9. Google—Android Market	.001%
10. Google—Orkut	.001%

Table 2: Top 10 Google Services Represented in Lumen (May–October 2013)

1. *The Rise of Third-Party Senders and a Shift to Movies, Music, and Adult Content*

At the next highest level of observation, notices sent to Google Web Search appear to have undergone two major shifts since Urban and Quilter's 2006 study. First, the shift towards automation goes hand-in-hand with a shift from major rightsholders sending their own notices to hiring third-party agents to detect infringement and send notices. Second, the major entertainment industries have moved from playing a minor role to become the dominant group requesting Google Web Search takedowns.

Section 512 notices must be sent by either the copyright owner or a person authorized to act on behalf of the copyright owner.[35] Urban and Quilter's 2006 study found that nearly all (98.5%) notices sent to Google services between March 2002 and August 2005 were sent by rightsholders themselves.[36] However, when they reviewed notices sent to a Texas webhost and connectivity provider between late 2004 and mid-2007, they saw an emerging role for trade associations and third-party REOs using automated methods to send notices to ISPs.[37] REOs, trade associations,

[35] See, e.g., 17 U.S.C. § 512(c)(3)(A)(vi) (2012) (stating that the notice must include a statement made under penalty of perjury that the complaining party is authorized to act on behalf of the owner of an exclusive right that is allegedly infringed).

[36] Urban & Quilter, *Efficient Process*, *supra* note 11, at 654.

[37] Jennifer Urban & Laura Quilter, Undue Process: Challenges for Rightsholders and Service Providers in Implementing Section 512's Notice and Takedown Provisions (Jan. 2009) (unpublished conference manuscript) (on file with authors). Quilter and Heins also identified the emergence of REOs in their 2007 interview study of OSPs. *See* QUILTER & HEINS, *supra* note 11, at 14-17.

and large owners also began using automated methods to target search engines. By 2012, Seng found that at least nine out of ten notices sent in 2012 that were archived in the Lumen database were sent by agents rather than by copyright holders themselves.[38] Nearly 46% came from just four large trade associations, led by the British recording industry association BPI.[39]

Consistent with Seng's findings, the trend towards outsourcing of takedown notices to agents is evident in our sample. Agents of copyright owners sent 91.8% of the requests in our sample; copyright owners themselves sent only 7.5%. *See* fig. 1, below.

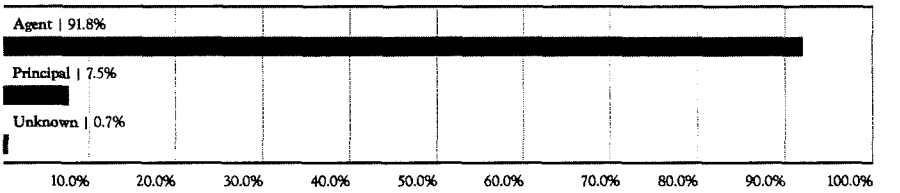

Figure 1: Sender Role for Entire Six-Month Set

Both third-party REOs and trade associations are highly active senders: REOs sent nearly half (49.3%) of takedown requests and trade associations sent 38%. Law firms trailed far behind at only 0.3%.

This shift to third-party services and trade associations appears in tandem with a major increase in the entertainment industries' sending of takedown requests, at least to Google Web Search. In their 2006 study, Urban and Quilter found little use of takedown notices to Google Web Search by the entertainment industries; the bulk of notices came from small Internet businesses, computer software companies, and game companies.[40] In that study, the movie and music industries combined were responsible for only 3% of § 512(d) notices.[41]

Today, the story is very different. Corroborating rightsholders' descriptions in Study 1, major rightsholders appear to use notices to Google Web Search as an important tactic within their enforcement strategy. The music, adult entertainment, and movie/television industries, taken together, sent by far the most takedown requests to Google Web Search in our sample. The largest percentage of requests, 44%, was issued by or on behalf of copyright owners in the music industry, followed by 28.1% requests issued by or on behalf of copyright owners in the adult entertainment industry, and 17% on behalf of copyright owners in the movie/

[38] Seng, *supra* note 11, at 448.
[39] *Id.* at 396.
[40] Urban & Quilter, *Efficient Process*, *supra* note 11, at 651.
[41] *Id.*

television industry. Other industries made up a significantly smaller proportion of requests: software (7.5%); games (5.4%); books (4.2%); web design (only 0.4%); and photography (only 0.2%).[42]

When did this shift occur? Seng's work identified a significant shift from the music and movie industry sending practices identified in Urban and Quilter's 2006 study.[43] Seng found that 32.1% of takedown requests sent by the top-fifty notice-senders sent between 2008–12 were sent by the music industry, 30.1% by the adult entertainment industry, and 20.8% by the movie/television industry.[44] Our data show a similar distribution, though we observed a further uptick in requests sent by the music industry. *See* fig. 4, below.

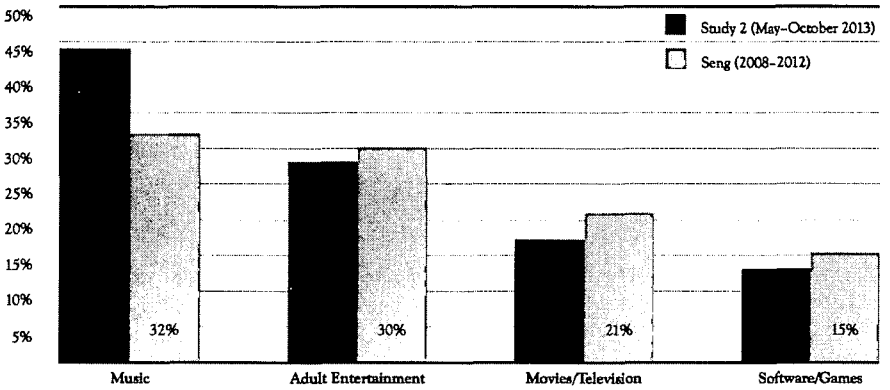

Figure 4: Percentage of Takedown Requests by Industry of Principal

As would be expected from the large representation of entertainment companies in the sample, the types of works for which removal was requested were heavily weighted toward audiovisual works and sound recordings, with pictorial and graphic works making up most of the remainder.[45]

---

[42] These numbers add up to more than 100% because copyright owners may be associated with more than one industry.

[43] Seng, *supra* note 11. Urban and Quilter's paper was published in 2006, but their data collection ended in 2005. Urban & Quilter, *Efficient Process*, *supra* note 11. However, in a later conference paper, looking at late 2004 to mid-2007 notices to The Planet, they also identified the emergence of REOs, supporting Seng's finding. Urban & Quilter, Undue Process, *supra* note 37.

[44] Note that Seng's data included notices sent to providers eligible for safe harbors under all of § 512's safe harbors, though, like our Study 2 data, the majority of notices were sent to Google Web Search (and were therefore § 512(d) notices). *See* Seng, *supra* note 11, at 419-20.

[45] For further descriptive statistics, see NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, Appendix A.

2. *Targeting Torrents: Over Two-Thirds of Requests Targeted Torrent or File Search Sites*

We also independently classified the type of site ultimately targeted by each takedown request in our sample — that is, the site whose link rightsholders were asking to be removed from Google's Web Search index.[46] In line with what we heard from major rightsholders in Study 1, the large majority — over two-thirds — of the requests in our main sample targeted file sharing sites — predominantly torrent and file search sites.[47]

Torrent sites were targeted by 35.9% of requests; and file search sites by 33.9%. A substantial chunk (12.7%) could not be classified. Of the rest, cyberlocker sites[48] (9.9%) made up the most significant group at nearly one in ten, followed by aggregator sites[49] (5.9%), forums/fan sites (4.4%), and video streaming sites (3.7%). E-commerce sites (0.9%), social media sites (0.5%), and personal websites/blogs (0.3%) made up only a handful. *See* fig. 5, below.

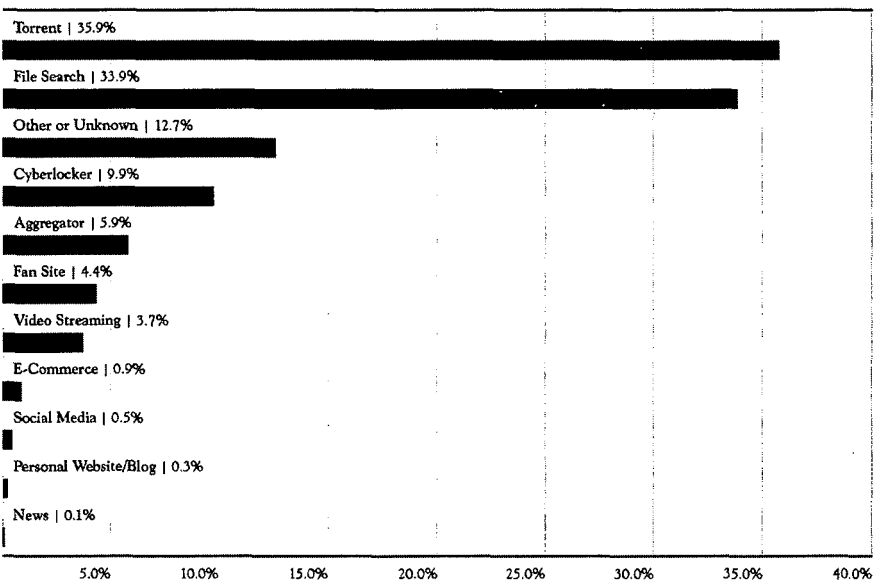

Figure 5: Types of Target Sites (Study 2)

46 For Google Web Search, we consider the site that the sender requests be removed from Google's search index to be the "targeted site."

47 File search sites are dedicated search sites that link to downloadable content across the Internet.

48 A cyberlocker site is a site that hosts user files.

49 Aggregator sites curate and present multiple options for accessing content.

Even accounting for the 12.7% we could not classify, a large majority of targets were types of sites likely to serve infringers, supporting rightsholders' assertions in Study 1 that they focus their takedown activity on likely large-scale infringement.[50]

The predominance of likely infringement sites as targets also appears to represent a shift compared to Urban and Quilter's observations in 2006. In their study, a significant portion of notices to Google Web Search targeted bloggers, fan sites, and other "hobbyists," as well as educational sites and critics.[51] More than half of targeted sites appeared to be competitors of the sender.[52] As large content purveyors have come to dominate sending to Google Web Search, the nature of the most commonly targeted sites has changed.

3. *Questions of Accuracy and Substantive Judgment: Misidentification Issues Predominated in Automated Notices*

The professionalized notices observed in Study 2 overwhelmingly appeared to be automated notices sent by "bots," reflecting the fact that most were sent by professional senders targeting large-scale infringement.[53] Unsurprisingly, the questions most commonly raised by Study 2 notices appeared to reflect issues created by automated decision-making and the vast scale on which automated requests are sent and processed. Our review of these notices corroborated OSPs' complaints that notices regularly left them struggling to identify and remove the complained-of material: the most common issues raised by Study 2 notice were problems

---

[50] There are alternative explanations. For example, the sheer number of file sharing, torrent, and similar sites may simply outweigh social media, fan sites, and similar sites. Without a census of the Internet, we cannot be sure. But this seems unlikely. Further, rightsholder and OSP descriptions generally agreed on the type of sites large rightsholders target, and anecdotal reports from OSPs participating in the Copyright Office's § 512 roundtables also tend to support this assertion. *See generally* NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, Section III ("Study 1); Auttomatic, Inc., Response to "Section 512 Study: Notice and Request for Public Comment" at 2 (stating that WordPress.com received 541 notices of claimed infringement in one month); Wikimedia Foundation, Response to "Section 512 Study: Notice and Request for Public Comment" at 1 (stating that Wikimedia Foundation received 41 takedown notices in 2015); Internet Archive, Response to "Section 512 Study: Notice and Request for Public Comment" at 2 (stating that Internet Archive receives a "relatively small number of DMCA notices").

[51] Urban & Quilter, *Efficient Process*, *supra* note 11, at 655-56. Urban and Quilter used somewhat different categories than we did, so there is not a one-to-one comparison, but the basic difference is clear. *Id.*

[52] *Id.*

[53] *Supra* Section III.A.1-2.

identifying the allegedly infringing material ("AIM") or the allegedly infringed work ("AIW").

In total, nearly a third (31%) of Study 2 takedown requests were potentially problematic, including both "fundamentally flawed" requests and "questionable" requests. (Where a single request presented multiple potential issues, we counted it only once.) "Fundamentally flawed" requests, which represented complete mismatches between the AIW and the alleged infringement, made up 4.2% of these. Of the remaining takedown requests, 28.4% had other characteristics that raised questions about their validity. These broke down into several categories of potential issues, centered on challenges identifying the works in question.

a) Fundamental Mismatches

One in twenty-five of the takedown requests (4.2%) in Study 2 were fundamentally flawed because they targeted content that clearly did not match the allegedly infringed work. This translates to approximately 4.5 million requests[54] across the entire six-month set that could be expected to suffer from this problem. These mismatches were of several types. In some cases, the sender was targeting material owned by the stated principal, but not the material identified in the notice. For example, in one request[55] sent by REO Vobile on behalf of Paramount, the AIW was the Paramount movie *An Officer and a Gentleman* and the AIM was the Paramount movie *Anchorman: The Legend of Ron Burgundy*. In other cases, the mismatch was apparently the result of automated systems making over-broad use of keywords. For example, one request[56] identified the copyrighted work as works by the artist "Usher," but the notice targeted a file for the movie *The House of Usher*. Still other mismatches occurred when the provided URL led to a search results page that did not actually list content related to the AIW.

b) Statutory Requirements — Identification Issues Predominated

Section 512 imposes a number of requirements on takedown notices aimed at allocating costs fairly, ensuring basic due process, and preventing abuse. These include statements that the owner or an authorized agent is the one requesting removal; that the claim is accurate and made in good

---

54 The margin of error for our sample is +/-2.29 with 95% confidence, so we can expect a range from 2 million to 7 million in the entire 108.3 million.

55 *DMCA (Copyright) Complaint to Google*, CHILLING EFFECTS (Aug. 7, 2013), https://chillingeffects.org/dmca512c/noticecgi?NoticeID=1123159 (Copyright Claim #3, Allegedly Infringing URL #40).

56 *BPI DMCA (Copyright) Complaint to Google*, CHILLING EFFECTS (Aug. 1, 2013), https://chillingeffects.org/dmca512c/noticecgi?NoticeID=1112904 (Copyright Claim #12, Allegedly Infringing URL #440).

faith; that there is information sufficient for a target to respond to the claim; and more substantively, that there is sufficient information about the AIW and AIM to allow the OSP to locate the AIM and assess the request.[57]

A surprising percentage of takedown requests — 19.5%, a little less than one in five — raised questions about whether the request was in compliance with these basic elements of notification.

The vast majority of questionable requests reflected the issue raised by OSPs in Study 1: failing to sufficiently identify the allegedly infringed work[58] (6%) or, most commonly, the material alleged to infringe[59] (13.3%). Nearly all of the notices that raised statutory compliance issues presented one of these two problems. Conversely, only a handful of requests in our sample raised questions about the more technical statutory requirements:[60] a signature of an authorized party;[61] information reasonably sufficient to contact the complaining party;[62] a statement of good faith;[63] a statement that the notice is accurate;[64] and a statement under penalty of perjury[65] that the complainant is authorized to act on behalf of the copyright owner.[66]

---

57 17 U.S.C. § 512(c)(3)(A)(i)-(vi) (2012).

58 *Id.* § 512(c)(3)(A)(ii).

59 *Id.* § 512(c)(3)(A)(iii). This element also requires that senders provide information reasonably sufficient to permit the service provider to locate the material. *Id.*

60 We expect this is, in large part, due to the following factors: the professionalized nature of the senders in our sample; some rightsholders' use of the standardized Automated Copyright Notice System; and Google's shift to a webform that demands these pieces of information. As noted above, users of Google's TCRP sent the vast majority of these notices — all pre-vetted, knowledgeable senders that make regular use of the takedown system. As regards the webform, all of the requests in the coded set were submitted using Google's online form. In Study 1 interviews, OSPs suggested that using webforms increases the likelihood that the required technical statements are included. For example, Google's webform system will not accept a takedown request unless senders provide an email address where they can be contacted.

61 17 U.S.C. § 512(c)(3)(A)(i) (2012).

62 *Id.* § 512(c)(3)(A)(iv).

63 *Id.* § 512(c)(3)(A)(v).

64 *Id.* § 512(c)(3)(A)(vi).

65 *Id.*

66 The dominance of identification issues may also be, in part, a product of our conservative methodology: we did not count as questionable the 36.8% of requests that failed to specify the copyright owner of the allegedly infringed work. Typically, these requests were from trade associations that listed "member companies" as the copyright owner rather than identifying the specific copyright owner of the identified work.

In the largest questionable group, information in the notice made it difficult to locate the allegedly infringing material targeted for removal. Most often, the notice provided a URL that led to a search results page or aggregator page that included multiple works, making identifying the AIM problematic. We note that our 13.3% finding probably represents an undercount of the requests that exhibit this issue. This is because we chose to count the 26% of notices targeting material that could no longer be located as valid absent additional information.[67] Nonetheless, issues with AIM identification were even more frequent than they were in Urban and Quilter's 2006 study.[68]

Though less common, problems identifying the allegedly infringed work also arose in a significant number —at least 6%, or about one in sixteen — of takedown requests. This is one of the most contested issues in debates over § 512, especially the question of when a sender's "representative list" of allegedly infringed works is sufficient.[69] Our 6% count,

---

[67] NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 79-80.

[68] Urban & Quilter, *Efficient Process*, *supra* note 11, at 674 (finding that one in eleven notices displayed significant statutory flaws). Urban and Quilter counted these notices slightly differently — they included notices that did not specify the complainant's contact information — but this difference only increases the relative proportion of notices with identification problems in our sample. *See id.*

[69] Where multiple copyrighted works at a single site are covered by the same notification, § 512 permits a sender to identify the allegedly infringed copyrighted work through a "representative list" of such works. 17 U.S.C. § 512(c)(3)(A)(ii) (2012). In *Perfect 10 v. Google*, the court held that a reference to the totality of the sender's image collection does not identify what may have been infringed and is not a representative list under the statute. Perfect 10, Inc. v. Google, Inc., No. CV 04-9484 AHM SHX, 2010 WL 9479060 (C.D. Cal. July 30, 2010), *aff'd*, 653 F.3d 976 (9th Cir. 2011). Therefore, notices that identified the copyrighted works by referencing an electronic folder of 15,000 images and offering the service provider a username and password to access the sender's website were not sufficient to identify the copyrighted work as required by the DMCA. *Id.* at *9. A list of artists' names without specifying any particular songs or allegedly infringing links is also insufficient to constitute a "representative list." Arista Records, Inc. v. Mp3Board, Inc., No. 00 CIV. 4660 (SHS), 2002 WL 1997918, at *15 (S.D.N.Y. Aug. 29, 2002) ("Although the DMCA permits a copyright owner to identify a 'representative' list of works . . . in this case, a bare list of musical artists whose songs were allegedly linked to did not constitute a representative list of works, or notice equivalent to a list of representative works that can be easily identified by the service provider."). On the other hand, providing a web address to where copyrighted works are located in conjunction with other identifying information may satisfy the requirements of providing a "representative list" as well as the requirement to provide information reasonably sufficient for the service provider to locate the infringing material. ALS Scan v. RemarQ Cmtys., Inc., 239 F.3d 619, 625 (4th Cir. 2001) (finding that the sender provided a representative list of infringing material and was sufficient to enable the defendant to locate the infringing material where the sender 1) identified two of the defendant's newsgroups that were cre-

again, is probably an undercount of the notices that would raise this question under current case law because of the conservative approach we took when categorizing notices.[70]

c) Potential Fair Uses

Though identification questions predominated the issues presented by Study 2 takedown requests, a significant number — about one in fifteen (6.6%) — exhibited at least one characteristic that likely would weigh favorably toward fair use, suggesting that more in-depth, human review was needed to identify any noninfringing materials. These requests predominantly targeted such potential fair uses as mashups or remixes, or links to search results pages including mashups or remixes. An additional notable group of requests targeted ringtones. The remainder varied widely — from cases where the AIM copied only a small portion of the AIW to cases where, based on surrounding information, the AIM was apparently being used for educational or instructional purposes.

We note that we could not do a full fair use analysis and, to foster intercoder reliability, focused on characteristics that reviewers could observe and record relatively easily.[71] We likely missed some potential fair

ated for the sole purpose of publishing ALS Scan's copyrighted works, 2) asserted that virtually all of the images on the two sites were plaintiff's copyrighted material, 3) referred the defendant to two Web addresses where it could find photographs of its models and obtain its copyright information, and 4) noted that its material could be identified because it included plaintiff's name and/or copyright symbol.). In *Perfect 10, Inc. v. Giganews, Inc.*, the district court found that Perfect 10 had neglected the obvious means of identifying material (Message-ID) in favor of screenshots of search results, which it viewed as not "reasonably sufficient to permit the service provider to locate the material." Perfect 10, Inc. v. Giganews, Inc., No. CV 11-07098- AB SHX, 2014 WL 8628031, at *8 (C.D. Cal. Nov. 14, 2014) (quoting 17 U.S.C. § 512(c)(3)(A)(iii)), *aff'd*, 847 F.3d 657 (9th Cir. 2017).

[70] Requests were counted as questionable only when the sender failed to identify any specific work at all: when the only information was a link to a website homepage, or when it was a link to another webpage where an allegedly infringed work was not apparent. If the sender provided *any* additional information, such as a title for the work or a general description of it, then we assumed that the information was sufficient to identify the AIW and did not count the notice as questionable. This likely excluded a large number of notices that could be considered questionable using a less conservative metric. For example, many notices from the adult entertainment industry identify the AIW by including a very vague description such as "video and image series by [principal name]" and a link to a website with a list of hundreds or even thousands of titles or images. If any such description was included, no matter how vague, we did not count that request as problematic.

[71] We selected characteristics to review based on: 1) characteristics that are generally considered to weigh in favor of fair use — for example, a transformative purpose for the use, or a relatively small amount taken for the use — and that 2)

uses; at the same time, the final merit of any potential fair use claims within this set will vary. Our goal was simply to observe whether automated systems appeared to generate any significant number of notices for which more contextualized human review is needed to check for fair use. It appears so: around 7 million notices out of the full 108.3 million can be expected to present these issues.[72]

d) Other Issues

A smaller percentage of takedown requests (2.3%) appeared to use the DMCA's takedown measures — which apply only to copyright infringement — when other concerns were involved. Trademark concerns were explicitly or implicitly raised in 1.3% of requests. Anti-circumvention issues — which typically arose where the AIM is a product key, rather than the copyrighted work it protects — were explicitly or implicitly raised in 1.1%. (We note that this overall category is very near the margin of error. Sub-categories are within the margin of error, and thus should not be used to draw conclusions about the overall proportions of these items in the full dataset.)

4. Study 2: Discussion

To a large degree, Study 2 corroborated the main themes we heard in Study 1 from DMCA Auto and Plus OSPs and from large rightsholders: that large rightsholders employ a high degree of automation; that they direct notices to a subset of OSPs that includes search providers; that they

reviewers could reasonably ascertain objectively from the information available in the notice (including information reasonably available at the location given by the links provided for the AIW and AIM). To fulfill the second critereon, we did not ask reviewers to decide if a use was, for example, "transformative," but instead asked them simply to identify characteristics of the AIM or its use that strongly suggest transformativeness (e.g., the AIM is a remix). Similarly, we did not ask them to judge whether the AIW was likely to have, at best, a "thin" copyright (making it more likely that a follow-on use falls within fair use), but instead asked them simply to identify AIMs that were rulesets, recipes, and the like. Because our methodology relied on using these more objective characteristics, we likely undercounted the number of works that exhibited characteristics that potentially favored fair use. At the same time, no one characteristic is dispositive of fair use. With a remix, for example, the obvious transformativeness of the use may ultimately be offset by the amount used and/or the effect on the market for the original work. Because the fair use factors must all be considered together for a full analysis, any actual determination would need further review. For more information on the characteristics that we examined, please see NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, Appendix C.

72 *See infra* Figure 7. The margin of error for our sample was+/-2.29 with 95% confidence, so the expected range of requests in this category is between 4.7 million and 9.6 million.

focus their efforts on protecting major entertainment industry copyrights; and that they focus on targeting the most obvious infringement sites.

It also echoed some of the potential problems described by OSPs in Study 1, particularly identification issues. Despite the efforts of some major senders to limit algorithmic mistakes, close to a third of requests presented questions about their validity. Further, though some categories of potential issues are proportionally small, the sheer number of notices means that even single-digit percentages each translate into millions of potential problems across the entire six-month set. *See* fig. 7.

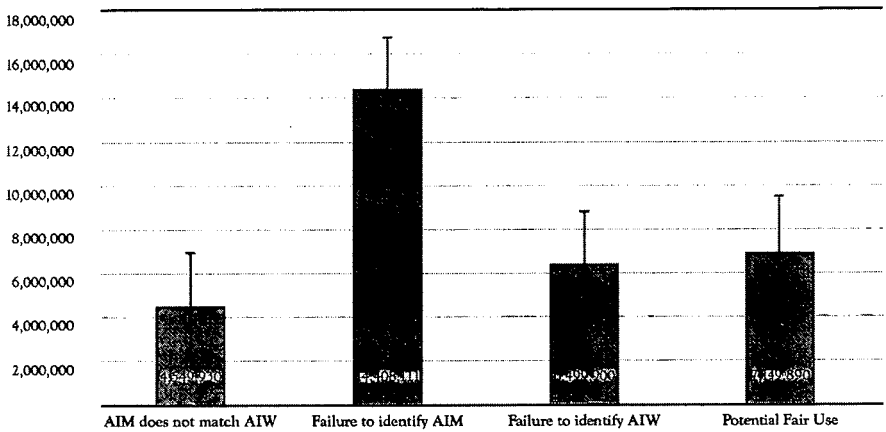

Figure 7: Problematic Takedown Requests in Six Months of Notices Sent to Lumen (Showing Error Bars Based on Margin of Error of +/-2.29 at 95% Confidence Interval)

Leaving potential fair uses aside, we suspect that the imprecision we saw in Study 2 requests reflects the challenges large rightsholders experience when using automated methods to manage large-scale infringement. As described in Study 1, both infringers and legitimate sites increasingly are dynamic, complex assemblages of content drawn from multiple sources, making precise identification of the source more difficult for rightsholders and OSPs alike. Because the ultimate targets of these particular notices were overwhelmingly sites that are likely to provide infringing materials, we expect that any overall negative effect on expression from "matching" mistakes is limited.

Still, there is reason to look for improvement. As discussed in Section IV below, § 512's identification requirements are not merely technical: they both preserve the enforcement cost allocations Congress chose in passing § 512 and help ensure fairness in the notice and takedown process. We suggest some improvements targeted to automated mistakes there.

B. *Study 3: In Notices Sent to Google Image Search, Individuals, Small Businesses, and Substantive Issues Predominated*

Study 3 considers only the notices in our six-month cohort that were sent to Google's Image Search service. Because notices to Google Web Search dominated the overall six-month set so completely, we isolated this subset with the expectation that notices to Google Image Search might present a different profile.

Once defined, the Google Image Search subset contained 33,409 requests,[73] housed in 2,777 notices. We reviewed and coded a randomized sample of 1,732 of these Google Image Search takedown requests[74] from 607 unique notices.

The Google Image Search requests indeed presented a strikingly different profile from the notices to Google Web Search we observed in Study 2. In fact, they differed in almost every characteristic. Where the Study 2 requests were automated notices, and sent by agents for large rightsholders targeting determined infringement, Study 3's Image Search requests did not appear to be automated, and tended to be sent by less professionalized claimants — individuals and small businesses — targeting social media, personal websites, and blogs. And instead of identification issues, Study 3 requests more often raised questions about the underlying claim.

1. *Senders: Small Copyright Owners Acting for Themselves, One Highly Prolific Sender, and Overseas Disputes*

As with Google Web Search, Google Image Search requests were also dominated by a few prolific senders: over three-quarters (78.3%) were sent by only seven senders (out of a total of 197 senders in the coded sample). Their profile, however, was very different from the professionalized Google Web Search senders. Most were individuals or small businesses. *See Table 3*, below.

---

[73] This number is an approximation because of the manner in which structured data from Google's Image Search complaint web form interacted with our coding engine algorithm. Some links were irrelevant repeats of the same information in different form.

Consequently, the margin of error for this sample — +/-2.29 at a 95% confidence interval — is artificially inflated because it is based on the full count of 33,409 requests. This is a conservatively wide margin of error. For more information, see NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1 n.285.

[74] As with the sample from the main population, this number includes only DMCA requests. In addition to the exclusions discussed above, *supra* note 30, approximately forty requests were excluded from the set because they possibly referred to child pornography. We instructed reviewers to immediately cease coding any such requests, did not review them, and did not include them in this analysis.

Sender	Sender Size	Percentage of Requests in Google Image Search Coded Sample
Ella Miller	Individual	52.9%
Lil Duck Duck	Small Business	8.5%
Purzel-Video GmbH	Small Business	6.1%
Anatoli Ivanov	Individual	3.5%
Alpha Sky Productions	Unknown	3.1%
Ragalahari	Small Business	2.9%
Oasis Costumes	Unknown	1.2%

Table 3: Percentage of Requests in Google Image Search Coded Sample

Strikingly, nearly 53% of the Google Image Search takedown requests were sent by one individual sender, whom we will call Ella Miller.[75] As discussed further below, all of Miller's notices were targeted toward improper subject matter. Because of Miller's outsized role and her general failure to send valid notices, we have frequently separated out the non-Miller notices in figures and for analysis. Miller's notices are as much a feature of the set as any other sender's notices, but it can also be helpful to look separately at her notices because some of their features would otherwise obscure observations that were more typical across all senders in the set.

a) Individual and Small-Business Senders

Across the set, including Miller, copyright claimants themselves sent the vast majority of takedown requests. Principals sent 94.4% of the requests to Google Image Search — a major difference from the agent-heavy Google Web Search requests.[76] Agents sent only a handful (.5%). The remaining 5.1% could not be classified. Setting aside Miller's individual requests did not change things much: principals sent nearly nine out of ten (88.1%) of the non-Miller requests, agents sent 1.1%, and 10.8% could not be classified. *See* fig. 8, below.

[75] "Ella Miller" and "Anatoli Ivanov" (found in Table 3) are both pseudonyms.

[76] Principals were categorized as such where the sender's name was the same as the principal's name (in the case of an individual copyright owner), or the sender's organization was the same as the principal organization (in the case of an organizational copyright owner).

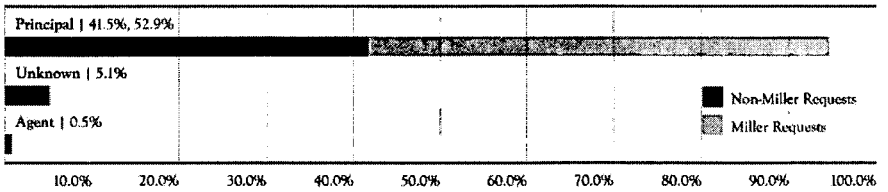

Figure 8: Sender Role (Google Image Search Sample)

Also unlike Study 2, the large majority of Google Image Search senders were individuals or small businesses.[77] Small businesses with fewer than 100 employees — and often many fewer — sent 41.1% of the non-Miller requests. Only a handful of senders were "small-medium" businesses of 100 to 999 employees (.7% of the non-Miller requests) or "medium" businesses with 1,000 to 9,999 employees (.2% of the non-Miller requests). In nearly a quarter (24%) of the non-Miller requests, the business size could not be categorized — most likely, these are also very small businesses, but we cannot be sure.[78] Overall, individuals and small businesses sent three-quarters (75%) of the non-Miller Google Image Search requests (increasing to 88.2% when the Miller notices were included).[79] *See* fig. 9. It appears that large, well-resourced senders do not focus on Google Image Search — in our dataset at least, takedown requests came from much smaller players.

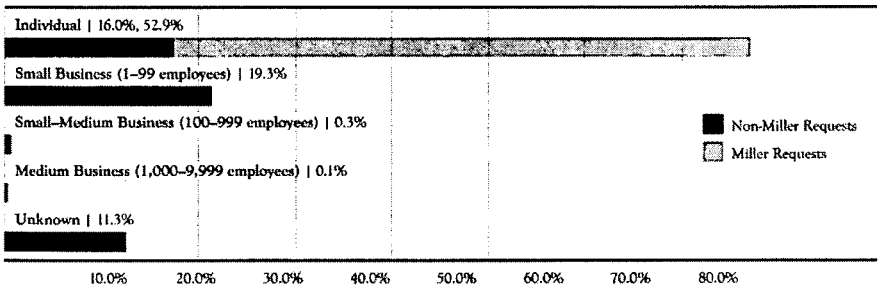

Figure 9: Sender Size (Google Image Search Sample)

[77] Where a sender provided a business name, we categorized the sender as a business. Individuals that provided a business name were counted as "businesses" even if they were self-named sole proprietorships. We used the LexisNexis Academic database to categorize business senders by size. *See Get Company Info*, LEXISNEXIS ACADEMIC, https://www.lexisnexis.com/hottopics/lnacademic (last visited Feb. 5, 2016).

[78] When Miller's notices are included, as in Figure 9, uncategorized notices make up 11.3%. These businesses could not be categorized because they did not appear in the LexisNexis Academic database. *See id.*

[79] Excluding Miller's requests, businesses sent nearly two-thirds (66.1%) of requests and individuals sent just over one-third (33.9%) of the requests. Including Miller's requests, individual requests rose to 68.9%.

Unsurprisingly, Google Image Search senders tended to represent visual creators. Though large corporations were absent, smaller image-focused businesses were well represented. The most prominent industries were the art/design[80] and adult entertainment industries, with each representing over one-fifth (23.8% and 20.3% respectively) of the non-Miller requests. The photography industry made a notable showing, with 10.5% of the non-Miller requests. E-commerce sites represented 5.1% of the non-Miller requests, and the industry was unknown or could not be classified in 6.3% of the non-Miller requests.

b) Overseas Disputes

Disputes originating outside of the United States were also a notable feature: overseas senders sent the majority of requests in the Google Image Search notices. All of Ella Miller's requests originated from Sweden, and appeared to relate to "flame-wars" between individuals located in Sweden, though some of these arguments may have taken place on servers in the United States.[81] Leaving Miller aside, senders based abroad sent over half (56.5%) of the notices to Google Image Search. This included senders based in Germany (16.4%), India (7.5%), Israel (7.5%), Great Britain (3.3%), and China (2.6%). Other foreign countries each had a smaller share but combined made up 19.2%. Senders based in the United States sent 43.5% of the requests (dropping to 20.5% if the Miller notices are included).

2. Targeting Social Media, Personal Websites, and Blogs

The Google Image Search targets also look different from the torrent and file search sites that made up two-thirds of target sites in the overall six-month dataset. In the Google Image Search set, the "worst of the worst" infringement sites gave way to social media postings, personal websites,[82] and blogs.

A quarter (24.6%) of the non-Miller requests targeted links to material on social media sites, and 15.6% targeted links to material on personal websites or blogs. News sites (7.4%), e-commerce sites (6.9%), forum/fan

[80] A significant number (81.8%) of the requests originating from the art/design industry were for greeting-card designs, with one sender (who was the second-most prolific sender after Ella Miller) sending the majority of these requests.

[81] Much of the targeted content appeared on a Flashback discussion forum, which is apparently based in the United States. *See* FLASHBACK, https://www.flashback.org (listing the website's corporate owner and address as "Flashback International Inc 1461 First Avenue, New York, NY 10075-2201, USA").

[82] Targeted sites tagged as "personal websites or blogs" may include small businesses, particularly where the business model is one based on in-site advertising revenue. Sites selling goods, however, were instead tagged as "e-commerce" sites.

sites (6.3%),[83] video streaming sites (6.1%), torrent sites (5.9%), aggregator sites (3.8%), cyberlockers (1.7%), corporate sites (1.6%), file search sites (1.0%), and educational sites made up the remainder of the targeted sites in the non-Miller requests. A large number (47.7%) of targeted sites could not be categorized, typically because the targeted site was no longer live. *See* fig. 13, below.

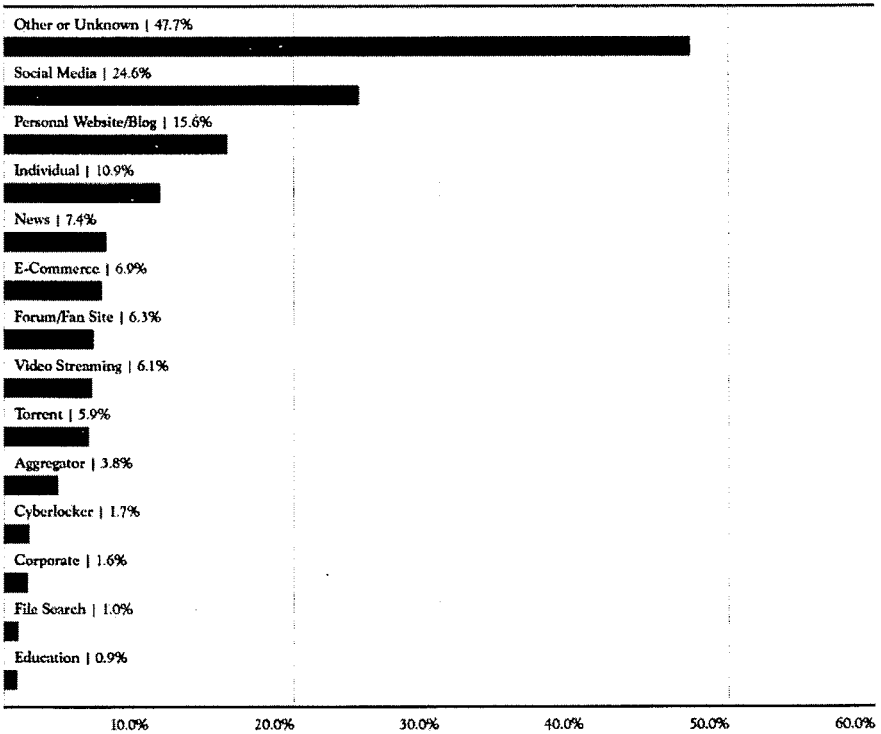

Figure 13: Types of Targets (Google Image Search Sample; Non-Miller Requests)

We also noted where it was apparent that an individual user's content was targeted, typically because the request was a direct link to a single post on a message board thread or where the targeted site was obviously a blog that belonged to an individual user. More than one in ten (10.9%) of the non-Miller requests clearly targeted an individual user.

[83] In an oversight, we neglected to create separate variables for forums and fan sites. These types of target sites were captured together in one variable, which was used to indicate fan sites or forums (used by fans or not). For more information, see NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, Appendix C.

3. *Questions of Accuracy and Substantive Judgment: Substantive Issues with the Underlying Claim Predominated*

The Google Image Search requests presented significantly more potential substantive problems, and a greater variety of potential problems, than the Study 2 notices. Overall, including the Miller requests — all of which appeared to be improper subject matter — seven out of ten (70%)[84] of the Google Image Search requests presented serious questions about their validity. Without the Miller requests, 36.3% of the remaining Google Image Search takedown requests were questionable.

a) Ella Miller's Requests

Ella Miller is a European individual who was embroiled in an online dispute about modeling photographs taken of her. Apparently in response to discussion of these photographs on various web forums, Miller sent thousands of the notices in our Google Image Search set.

Miller appeared to provide a supportive example for the "rule of thumb" to more closely scrutinize notices from individual senders that OSPs described using in Study 1. All of her requests appeared to be improper subject matter for DMCA takedown. Though she identified them as DMCA notices[85] — indicating a copyright claim — none were copyright complaints. Though we could not discern details, the notices typically targeted message board threads and blog posts that appeared to be critical of Miller. Most notices pointed to written material that Miller alleged to be defamatory, harassing, slanderous, or threatening.[86] An additional handful of requests identified the AIM as a photograph of Miller. However, these requests were still grounded in defamation and similar torts rather than copyright.[87]

---

[84] For purposes of calculating the total number of questionable requests, a request that has multiple questionable characteristics was only counted once.

[85] Senders who use Google's Web form for Image Search answer a series of questions that are intended to filter requests into the appropriate complaint track. The notices are then labeled with an "issue type" that identifies the type of complaint. *See Removing Content from Google*, GOOGLE, https://support.google.com/legal/troubleshooter/1114905?hl=en#ts=1115648 (last visited Feb. 5, 2016).

[86] Such concerns are certainly sympathetic, though not proper subject matter for a § 512 notice. In a previous study, Urban and Quilter similarly observed the use of § 512 for defamation, privacy, and related claims. Urban & Quilter, Undue Process, *supra* note 37. The lack of a similar notice and takedown regime for such claims may lead complainants to shoehorn non-copyright complaints into the § 512 structure; individual senders may or may not realize that this is an inappropriate use of the copyright takedown notice.

[87] The few requests that identified a photograph as the allegedly infringing material requested takedown of professional portraits of Miller herself. If these notices had raised copyright claims, they still would raise questions about whether Miller

b) Requests from Other Senders

Leaving Miller aside, more than a third (36.3%) of the remaining Google Image Search requests also presented substantive problems. These broke down into several categories: questionable subject matter for a takedown request; potential fair use defenses; questions surrounding ownership of the allegedly infringed work; and a handful of less-frequent issues. *See* fig. 15, below.

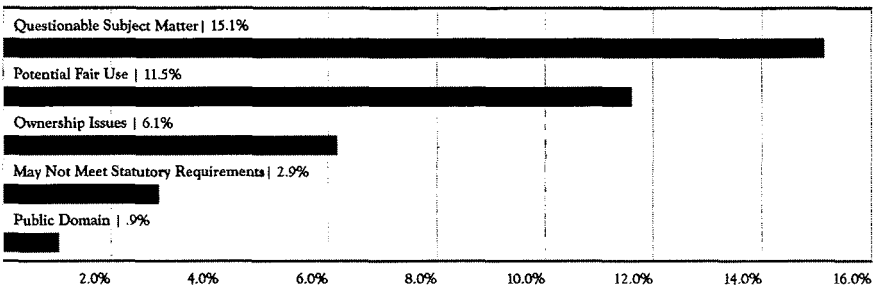

Figure 15: Problematic Takedown Requests (Google Image Search Sample; Non-Miller Requests)

i) Subject Matter Issues

Like Miller's requests, a significant number — 15.1%, close to one in six — of other senders' requests raised issues outside of copyright, raising questions about the use of the DMCA takedown process. The specific issues varied. Most prominent were privacy concerns — typically arising when senders appeared in a photograph that they wanted removed from the Google Image Search index. Another notable group involved product photographs.[88] A small percentage of requests raised trademark concerns, including a few where the work alleged to be infringed was a short trademarked (but uncopyrightable) phrase, such as "Point Master" or "Mos-

in fact owned the copyright to the images — which would more commonly accrue to the photographer — or was authorized to send notices on behalf of the copyright owner.

[88] While these can be couched as copyright claims, product photograph disputes are often closely related to brand control, distribution channels, and competition issues. Any copyright interest in a product photo is likely to be thin: because they may lack the requisite originality to merit copyright protection, product photos raise questions about the copyrightability of the original work. As most photos involve at least some minimal creativity (for example, courts have held that sufficiently creativity for at least thin copyright protection may come from creative use of lighting, angles, or composition), the more direct issue may be fair use. This is especially salient as requests to remove product photos typically involved a sender and a target competing in the same line of business. Cases like these present an OSP with a difficult judgment call.

quito Genie." Others identified the allegedly infringed work as visual elements of the trademark or logo but expressed concerns unrelated to copyright. For example, in one request, a French theme park company wanted the identified material removed because it had recently updated its branding and the AIM was "not consistent with our current brand."[89] Finally, a small percentage of requests raised defamation claims or concerns about "online impersonation" and "harassment."

ii) Potential Fair Uses

One in nine of the non-Miller requests (11.5%) were flagged with characteristics that weigh favorably toward fair use, suggesting that further review could reveal a fair use defense.[90] Over half of these were requests to take down allegedly infringing material on news sites. Others included requests where the allegedly infringing material was apparently being used for educational purposes, such as a scientific photograph of bacteria under a microscope or photographs of architectural and historical sites in Israel. In some cases, the allegedly infringing material was a part of a curated collection of images. Examples included greeting card designs that were used as examples of "cute invitations" or "wording for a 60th anniversary card." Other requests in this category exhibited a variety of other characteristics that suggested considering fair use would be warranted; for example, cases where only a small amount of the original work was copied, or where the allegedly infringing material adapted, commented on, or criticized the original work.

iii) Ownership Issues

In nearly one out of sixteen of the non-Miller requests (6.1%), the allegedly infringed work was a photograph in which the sender was the subject of the photograph. Since copyright ownership typically vests with the photographer, the subject of the photograph does not usually own the copyright to the photograph in which they appear. It is thus unclear whether the sender has the authority to send a takedown request in such cases.[91] Finally, a small number of the non-Miller requests (.9%) targeted material that appeared to be in the public domain and therefore not pro-

[89] Translated to English from the original French text.

[90] As with the Study 2 sample, we could not conduct a full fair use analysis, and followed the methodology described *supra*, note 71, and NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, Appendix C. As such, the strength of potential fair use claims will vary.

[91] It is possible that, in some cases, the photographer transferred copyright ownership to the subject of the photograph. And in the age of the "selfie" stick and front-facing camera phones, it is possible that the subject of the photograph may also be the photographer and, as such, the copyright owner. Excluding instances

tected by copyright. (This is within in the margin of error and should be considered only anecdotal.)

iv) Identification Issues Were Not a Predominant Feature

Compared to the requests in Study 2, a relatively small percentage of the non-Miller requests (2.9%) presented problems with identifying the works in question. A relatively small number of requests linked to search aggregator pages, leading to problems with identifying the allegedly infringing material on the linked page. A negligible number failed to identify the allegedly infringed work.

4. Study 3: Discussion

Overall, strikingly different questions predominated from those posed by the automated Study 2 notices. Study 3's individual and small-business senders were relatively unlikely to misidentify the allegedly infringed works or the allegedly infringing works — it was, generally speaking, straightforward to identify the works in question and compare them. Study 3 senders were, however, notably more likely to send notices that raised substantive questions about the underlying claim (for example, whether the claim was proper subject matter for a § 512 takedown request).

The high number of substantive questions raises concerns about negative effects on expression. And where Study 2 senders' focus on likely infringement sites reduced concern that their mistakes affected expression interests, Study 3 senders' targeting of social media, personal websites, blogs, and individual user content heightens this concern.

As to why substantive mistakes were a stronger feature of the Study 3 notices, this may be related to the less-professionalized nature of typical senders to Google Image Search, who, as individuals and small business owners, may be less likely to be knowledgeable about the complexities of copyright law and when takedown notices are appropriate than the sophisticated copyright-industry senders predominant in Study 2. This would comport with what OSPs in Study 1 told us about their triage systems: OSPs commonly flag notices from individual or one-off senders for additional review because their experience tells them that these notices are more likely to be mistaken or abusive than notices from more sophisticated senders.[92]

The high proportion of substantive mistakes may also stem in part from the large percentage of senders (56.5%) who appeared to be outside

where the complained-of image appeared to be a self-portrait drops the number of notices with ownership issues down to about one in eighteen (5.6%).

[92] NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 40.

of the United States. These senders may be even less likely to understand the nuances of U.S. copyright law — for example, its subject-matter limitations and its fair use provision — than smaller domestic senders.[93]

The different issues presented by, on the one hand, automated notices, and, on the other hand, notices sent by smaller senders, suggest different interventions, which we discuss below.

IV. ANALYSIS AND RECOMMENDATIONS

Our six-month snapshot of takedown requests revealed two "worlds" of notice and takedown practice. In one, represented by Study 2 notices to Google Web Search, large rightsholders with valuable copyrights fight determined infringers with onslaughts of automated notices. In the other, represented by Study 3 notices to Google Image Search, individuals and small businesses seek to resolve a variety of disputes (including copyright, but also privacy, defamation, and other disputes) using takedown notices.

Though not necessarily representative of notice and takedown overall, these studies both corroborated the picture painted for us by OSPs and rightsholders in the broader Study 1 and echoed some of the major themes of discontent surrounding § 512. Study 2 showed large rightsholders making heavy use of automated takedown and, as they described in Study 1, focusing it toward likely infringers. At the same time, Study 2 also supported OSPs' complaints that misidentified or imprecise notices could create a substantial burden. Study 3 gave credence to OSPs' "rule of thumb" to subject notices from "small" or "one-off" senders to more scrutiny to catch substantive problems. Study 3 also buttressed OSPs' concerns about problems with the underlying legal claims in some notices and their reports that non-infringing material was regularly targeted. In microcosm, these sets of notices depicted some of the central controversies around notice and takedown: the frustrations of using it to address large-scale piracy; the debates over the proper allocation of enforcement costs; and the concerns about its potential effect on competition policy and expression interests.

The two "worlds" of notices present quite different questions and suggest different methods for improving notice and takedown. The issues raised by Study 2 notices almost certainly stemmed from the vast scale of notice sending observed and the automated nature of the notices. Misidentification of the works in question predominated over substantive problems (though potential fair uses should not be forgotten). Because

[93] Though the OSP in question is the United States company Google Inc., the predominance of overseas senders also raised deeper questions about jurisdiction and the appropriate use of § 512 takedown. *See* NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 102-03.

these notices overwhelmingly targeted likely infringing sites, we are less concerned that questionable notices in this set would negatively affect expression. However, improvements are still needed. A "mere" misidentification raises both process and cost-allocation issues, and in some cases can directly affect expression.

First, § 512's identification requirements are fundamental as a matter of process: Without the ability to compare the allegedly infringed work with the material for which takedown is requested, an OSP (or the responsible user, or another reviewer) cannot assess whether a takedown request is proper. Accordingly, notices that do not substantially comply with these requirements are not sufficient to confer knowledge of infringement — and thus potential secondary liability if the OSP does not remove the material — on the service provider.[94] (In contrast, notices that exhibit the more technical statutory deficiencies — such as a failure to substantially comply with the requirements for a signature, good faith statement, and statement that the notice is accurate and that the complaining party is authorized — may still be considered in determining whether an OSP has actual or red flag knowledge of infringement if the OSP does not attempt to follow up with the sender to cure the problem.)[95]

Second, identification issues potentially undo the cost-benefit balance Congress struck in § 512 by reallocating search costs from rightsholders to OSPs. Indeed, these issues illustrate the longstanding tussle between rightsholders and OSPs over the appropriate allocation of enforcement costs. By requiring notices to sufficiently identify infringements before they could trigger actionable knowledge under § 512, Congress chose to allocate infringement search costs, in the main, to notice senders. This both mirrored copyright holders' traditional duty to identify infringements and reflected Congress' judgment that copyright holders were in the best position to identify infringements online.[96] Though rightsholders have long chafed at this choice and have heavily litigated the identification requirements, courts have affirmed them even as the scale of potential infringements has risen.[97]

A shift in this cost allocation becomes a policy concern if it begins to undermine § 512's competition policy by restricting less-well-resourced

94 17 U.S.C. § 512(c)(3)(B)(ii) (2012).

95 *Id.*

96 *See supra* note 4 and accompanying text.

97 *See, e.g.*, Perfect 10, Inc. v. CCBill, LLC, 488 F.3d 1102, 1114 (9th Cir. 2007); UMG Recordings, Inc. v. Veoh Networks, Inc., 665 F. Supp. 2d 1099, 1108-09 (C.D. Cal. 2009), aff'd sub nom. UMG Recordings, Inc. v. Shelter Capital Partners, LLC, 667 F.3d 1022 (9th Cir. 2011), *opinion withdrawn and superseded on reh'g*, 718 F.3d 1006 (9th Cir. 2013), *and aff'd sub nom.* UMG Recordings, Inc. v. Shelter Capital Partners, LLC, 718 F.3d 1006 (9th Cir. 2013); Hendrickson v. eBay, Inc., 165 F. Supp. 2d 1082, 1093 (C.D. Cal. 2001).

OSPs' ability to rely on the safe harbors. If high levels of identification mistakes are common to automated noticing,[98] then the cost of complying with the takedown provisions — already resource-intensive — could increase substantially for all OSPs that receive automated notices and perhaps become unfair or unsustainable for the vast majority of OSPs that lack the "Big Five's" resources.[99] It also becomes a policy concern if the cost to OSPs of managing identification mistakes — fruitlessly searching their systems for infringements that cannot be found — inordinately limits their ability to find and address notices with substantive problems or their ability to efficiently respond to non-automated notices from smaller senders.

Third, imprecise notices also have the potential to cause substantive problems. Some of the measures OSPs described in Study 1 illustrate. Common examples included requests to remove search result pages that may include both infringing and non-infringing content, requests to remove comment threads that may contain an infringing link somewhere in the thread, and requests to remove pages with dynamic content that may no longer contain the material in question.[100] The last case illustrates that, beyond raising search costs for rightsholders and OSPs alike, the challenges posed by dynamic content also increase the potential for ripple effects in the event of an imprecise takedown. At the limit, OSPs reported that takedown requests could become de facto takedowns of whole sites, either through the volume of requests or the targeting of top-level pages. Most OSPs reported acting conservatively, taking down content in order to avoid liability even if it meant also removing non-infringing content on the targeted page.[101]

Given these considerations, mistargeting by automated systems should be addressed. Substantive issues with notices — like those seen in Study 3 and to a lesser degree in the potential fair uses in Study 2 — raise

---

[98] That is, rather than identification mistakes being only a feature of the Google Web Search notices we could directly observe. This seems unlikely, especially since OSPs in Study 1 generally complained of identification problems. *See supra* note 27 and accompanying text.

[99] In Study 1, Google was a universal reference point for OSPs' cost concerns, but other large companies' resources also leave many OSPs' far behind. This is especially true of any of the "Big Five" OSPs, which by market capitalization are the largest companies, of any kind, in the world. Natalie Sherman, *Are Google, Amazon and Others Getting Too Big?*, BBC NEWS (June 8, 2017), http://www.bbc.com/news/business-39875417 (describing the five largest companies worldwide as Apple, Alphabet (Google Inc.'s parent company), Microsoft, Amazon, and Facebook).

[100] NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 39.

[101] *Id.* On the other hand, some OSPs were able to develop policies that avoid the most sweeping requests. Search service respondents, for example, generally reported that they reject takedown requests targeting home page URLs. *Id.*

immediate questions about potential effects on expression and should also be addressed. Automated systems for infringement detection, noticing, and OSP processing are certainly necessary in the face of large-scale infringement (which may be automated as well), and they should remain cost-effective. But these systems can be improved.

By and large, we think voluntary measures — perhaps encouraged by policymakers — are the best way to efficiently improve automated systems while keeping them useful to rightsholders. We recommend a number of "best of breed" policies and practices, often suggested by rightsholders or OSPs in Study 1. Good practices for rightsholders using automated methods include:

- As we saw in Study 2, adopting initial filtering practices that focus automated efforts on "rogue" target sites identified through human review to limit the fallout from mistakes;
- Devising targeting algorithms that go beyond simple filename or URL matching to more accurately identify infringing materials;
- Employing human spot-checks of algorithmic decision making (by, for example, sampling notices for human review);
- Developing technical and human methods for better identifying allegedly infringing materials, and filtering out obvious potential non-infringing uses; and
- Employing policies and guidelines for enforcement agents and enforcement staff to help them avoid misidentifications and identify fair use and relevant tolerated uses.

Similarly, OSPs that use automated measures should engage in cross-checks to reduce overbroad takedowns. Often, these mirror practices recommended for rightsholders. For example, we recommend that OSPs:

- Develop automated methods to weed out obviously flawed notices; conduct human spot-checks of automated notices and use them to refine automated methods; and escalate questionable notices for human review.[102]

These suggestions could help address both problems introduced by automated notice sending and problems introduced by non-automated, but substantively questionable notices that arrive embedded within the tide of automated notices.

Fully addressing the substantive problems we saw in the Study 3 "world," however, requires a different approach. Best practices for automated processing could help OSPs automatically flag notices that claim improper subject matter — for example, by flagging notices that mention

---

102 For more detail on our recommendations related to automation problems, as well as additional recommendations, see NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 134-36.

"private," "defamation," or other keywords — or exhibit other machine-readable characteristics, but overall, automation is unlikely to fully address the wide variety of issues we observed. Accordingly, we also recommend that OSPs use human review where possible (for example on samples of notices or on notices that are triaged as potentially problematic). OSPs should implement triage systems that will flag for additional scrutiny notices that are more likely to be problematic (such as a notice from a sender that is an apparent competitor of a target or is from a one-time or new sender). Finally, OSPs should question or reject notices in appropriate circumstances.[103]

Front-end educational efforts directed at senders may also be effective at improving accuracy. While some poor-quality notices — for example, those directed at a competitor of the sender — may be truly abusive, others may simply reflect a sender's limited knowledge of copyright law. Given that copyright is specialist subject-matter even for lawyers, this would be both unsurprising and understandable. Accordingly, non-expert senders may benefit from guidance at the point of sending a notice. Educational materials could be provided by OSPs, trade groups, non-profits, or institutions like the Copyright Office or law school centers. Ideally, the same materials could provide senders with "how-to" information and guidance about appropriate takedown requests, and targets with copyright guidance and an easy-to-use counter notice function.

Finally, some modest statutory reforms could improve noticing of all types. We stress that these should be limited and take into account the complexity and diversity of the § 512 system. As further described in Study 1, different OSPs and senders are differently situated, with different needs and challenges. This leaves statutory reform efforts that benefit some parties at risk of causing unintended harm to others. Statutory changes should not inordinately raise costs on any party, and should maintain the safe harbor's pro-competition effects.

Consequently, our recommendations for statutory reforms are limited, and largely follow the suggestions of others. They focus on evening out the liability risks for sending and responding to notices, with goals of: encouraging high-quality notices without imposing unreasonable costs on rightsholders; lowering the cost to OSPs of not taking down material when they receive a flawed notice; and encouraging the use of counter notices where appropriate. As counter notice senders already must, notice senders should equally declare under penalty of perjury that their substantive

---

[103] For more detail on our recommendations related to human error, including recommendations for improving small senders' access to more sophisticated enforcement methods, see *id.* at 137-40.

claims in a takedown notice are accurate.[104] Similarly, we suggest that the standard for recovery for damages from misrepresentations in a notice or counter notice be slightly lowered, from "knowing, material" misrepresentations, to "reckless" misrepresentations. While this is still far too high a standard to deter legitimate notice sending, it could deter the worst offenses and allow for those harmed by bad notices to have some chance of recovery. Finally, we recommend remedying the lopsided liability risk posed by the current statutory damages regime so that targets have more confidence to send counter notices[105] and OSPs have more confidence to reject problematic notices.[106]

---

104 *See, e.g.*, DENA CHEN, MUSETTA DURKEE, JARED FRIEND, & JENNIFER URBAN, UPDATING 17 U.S.C. § 512'S NOTICE AND TAKEDOWN PROCEDURES FOR INNOVATORS, CREATORS, AND CONSUMERS 14-15 (2011). Disclaimer: Public Knowledge's white paper was prepared by Samuelson Law, Technology & Public Policy students, representing Public Knowledge under Urban's direction. The opinions in the white paper belong to Public Knowledge and not necessarily any author of this report, though, based on our research, we endorse some of them here as well.

105 From all available research, including Study 1, it appears that counter notices are rarely used. *See, e.g.*, NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 44-46. These suggestions respond to reasons why OSPs thought targets did not send counter notices even when they had strong arguments; however, more direct research into targets might suggest other approaches. Further, these changes may not substantially increase counter noticing to search providers, specifically, because search providers do not have to communicate notices to targets (and often cannot, as they do not have a service relationship with targets). Further, it is currently unclear whether search providers must accept counter notices and whether responding to a counter notice by replacing the material in question ("putback") is protected for search providers as it is for other OSPs. The issue is highly technical; *see* NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 132-33 for discussion. To partially address this, we also suggest a technical fix that would make clear that search engines must accept counter notices and that they have the same protection for "putback" as other OSPs. *Id.*

106 We recommend adopting reforms proposed by the Department of Commerce's Internet Policy Task Force, along with some modest extensions, including giving courts discretion to depart from the "per infringed work" calculation in cases of non-willful secondary liability for online services, specifying factors in the Copyright Act that courts must consider when assessing statutory damages, and increasing the availability of statutory damages reductions in cases of innocent infringement. *See* DEP'T OF COMMERCE INTERNET POLICY TASKFORCE, WHITE PAPER ON REMIXES, FIRST SALE, AND STATUTORY DAMAGES (2016), at 86-99; NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1, at 129-30. For additional recommendations for what courts should and should not do when awarding statutory damages under the current Copyright Act regime, see Pamela Samuelson & Tara Wheatland, *Statutory Damages in Copyright Law: A Remedy in Need of Reform*, 51 WM. & MARY L. REV. 439, 501-09 (2009), http://scholarship.law.wm.edu/wmlr/vol51/iss2/5.

None of these changes should limit rightsholders' ability to use automated systems, nor should they substantially raise costs for any sincere notice sender. However, they should deter the worst behavior on the part of senders, and give OSPs and targets more incentives to challenge problematic takedowns.

V. CONCLUSION

After nearly two decades of technological and social change, § 512 notice and takedown remains at the center of online copyright enforcement. Though notice and takedown is difficult to study, its importance to OSPs, rightsholders, and online users makes it important to look into the "black box" as far as possible. The partial view afforded by our look at Lumen notices gives some useful insights into takedown practice and some of the challenges it poses. We observed some important features of the notice-and-takedown system and found some obvious opportunities for improvement. Our suggestions are intended to increase notice quality, while sustaining or improving § 512's usefulness for rightsholders, OSPs, and online users.[107]

[107] Still, the picture remains incomplete. Further research — especially further studies of DMCA Classic OSPs and studies of smaller rightsholders — could fill out the map of the notice and takedown ecosystem and suggest additional improvements.

PART II

RECENT DEVELOPMENTS IN COPYRIGHT