

VIRGINIA JOURNAL OF LAW & TECHNOLOGY

SUMMER 2014 UNIVERSITY OF VIRGINIA VOL. 18, No. 03

The State of the Discordant Union: An Empirical Analysis of DMCA Takedown Notices

DANIEL SENG[†]

© 2014 Virginia Journal of Law & Technology Association, at <http://www.vjolt.net>.

[†] Daniel Seng, J.S.D. Candidate, Stanford Law School, Stanford University and Associate Professor, Faculty of Law, National University of Singapore. The material for this article was presented at the 15th Year Retrospective of the Digital Millennium Copyright Act, Santa Clara University on March 15, 2013. I would like to thank Mark Lemley, Deborah Hensler, Eric Goldman, Jennifer Urban, Roland Vogl, Corynne McSherry, and Maria Jose Cordero for their useful feedback and suggestions, and all the participants at the 15th Year Retrospective for their many valuable and encouraging comments.

ABSTRACT

By conducting a census on half a million takedown notices and more than fifty million takedown requests in its datasets, this Article takes a detailed and systematic look at the state of the takedown process from an empirical perspective. It examines the use and issuance of takedown notices by copyright owners and reporters and the response of service providers to them. This Article further studies the relationship between the notices and requests and safe harbor provisions of the Digital Millennium Copyright Act and identifies ways in which the takedown process can be further improved to preserve the diversity and freedom of the Internet.

TABLE OF CONTENTS

I. Introduction	373
II. Study Methodology	378
A. Chilling Effects and Google Transparency Report	378
B. Ethical Considerations and Limitations	383
III. Takedown Notices in General	388
A. Content Providers	392
B. Reporters and Reporting Agents	395
C. Takedown Notice Formats	398
D. Takedown Requests in General	400
E. Trusted Reporters and their Takedown Requests ..	414
IV. Google Services' Takedown Notices and Requests by DMCA Section	418
A. Notices and Requests for Google Services	418
B. Counter Notices	426
V. Proposals for Reform	431
A. Electronic Notices	432
B. The Growth of Notices with Mega Claims and Mega Requests	433

C. URIs as Takedown Requests..... 436

D. Greater Transparency about the Trusted Copyright Removal Program 437

E. Section 512(d) Notices and Section 512(g) Counter-Notifications..... 439

VI. Conclusions 440

VII. List of Extracted Fields in Takedown Notices..... 443

VIII. Tables 444

IX. Figures 463



I. Introduction

In October 2013, the Digital Millennium Copyright Act (DMCA) celebrated its fifteenth anniversary. Enacted, among others, to limit the liability of Internet intermediaries as service providers,¹ the DMCA safe harbor provisions have been credited with ensuring the very existence of the Internet and preserving the variety and quality of its content.² They have further served as the basis for similar provisions in the intellectual property chapter in the free trade agreements entered into between the United States and Australia,³ Bahrain,⁴ Central America–Dominican Republic states,⁵ Chile,⁶ Columbia,⁷ Korea,⁸ Morocco,⁹ Oman,¹⁰ Panama,¹¹

¹ The Online Copyright Infringement Liability Limitation Act (OCILLA), 17 U.S.C. § 512 (2012); *see also* S. REP. NO. 105-190, at 8 (1998).

² *See, e.g.*, Samantha Rose Hunt, *DMCA 10 Years Old, Keeping Internet Alive*, TG DAILY (Oct. 28, 2008, 6:01 PM), <http://www.tgdaily.com/business-and-law-features/39944-dmca-10-years-old-keeping-internet-alive>.

³ Free Trade Agreement, U.S.-Austl., art. 17.11, ¶ 29, May 18, 2004, *available at* http://www.ustr.gov/sites/default/files/uploads/agreements/fta/australia/asset_upload_file469_5141.pdf; *see also* Copyright Act 1968, Part V, Division 2AA, § 116AA (Austl.).

⁴ Free Trade Agreement, U.S.-Bahr., art. 14.10, ¶ 29, Sept. 14, 2004, *available at* http://www.ustr.gov/sites/default/files/uploads/agreements/fta/bahrain/asset_upload_file211_6293.pdf.

⁵ Free Trade Agreement, U.S.-Dom. Rep.-Cent. Am., art. 15.11, ¶ 27, Aug. 5, 2004, *available at* http://www.ustr.gov/sites/default/files/uploads/agreements/cafta/asset_upload_file934_3935.pdf.

⁶ Free Trade Agreement, U.S.-Chile, art. 17.11, ¶ 23, June 6, 2003, *available at* http://www.ustr.gov/sites/default/files/uploads/agreements/fta/chile/asset_upload_file912_4011.pdf.

Peru,¹² and Singapore.¹³ They also have served as the template for the enactment of similar defenses in the European Union,¹⁴ including the United Kingdom,¹⁵ and the People's Republic of China.¹⁶ Considering that the safe harbor provisions form no

⁷ Free Trade Agreement, U.S.-Colom., art. 16.11, ¶ 29, Nov. 22, 2006, available at <http://www.ustr.gov/sites/default/files/col-ipr.pdf>.

⁸ Free Trade Agreement, U.S.-S. Kor., art. 18.10, ¶ 30, June 30, 2007, available at http://www.ustr.gov/sites/default/files/uploads/agreements/fta/korus/asset_upload_file273_12717.pdf; see also Copyright Act, No. 9625, Apr. 22, 2009, arts. 102-104 (S. Kor.).

⁹ Free Trade Agreement, U.S.-Morocco, art. 15.11, ¶ 28, June 15, 2004, available at http://www.ustr.gov/sites/default/files/uploads/agreements/fta/morocco/asset_upload_file797_3849.pdf.

¹⁰ Free Trade Agreement, U.S.-Oman, art. 15.10, ¶ 29, Jan. 19, 2006, available at http://www.ustr.gov/sites/default/files/uploads/agreements/fta/oman/asset_upload_file715_8809.pdf.

¹¹ Trade Promotion Agreement, U.S.-Pan., art. 15.11, ¶ 27, June 28, 2007, available at http://www.ustr.gov/sites/default/files/uploads/agreements/fta/panama/asset_upload_file131_10350.pdf.

¹² Trade Promotion Agreement, U.S.-Peru, art. 16.11, ¶ 29, Apr. 12, 2006, available at http://www.ustr.gov/webfm_send/1031.

¹³ Free Trade Agreement, U.S.-Sing., art. 16.9, ¶ 22, May 6, 2003, available at http://www.ustr.gov/sites/default/files/uploads/agreements/fta/singapore/asset_upload_file708_4036.pdf; see also Copyright Act, Ch. 63, Part IXA (2006) (Sing.).

¹⁴ Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on Certain Legal Aspects of Information Society Services, in Particular Electronic Commerce, in the Internal Market, 2000 O.J. (L 178) 1.

¹⁵ The Electronic Commerce (E.C. Directive) Regulations 2002, SI 2002/2013, regs. 17, 18, 19 (U.K.).

¹⁶ Regulations on the Protection of the Right to Network Dissemination of Information Networks (promulgated by the State Council, May 18, 2006,

part of any international intellectual property and copyright treaties, the DMCA safe harbors have indeed gone global. And the world has embraced the DMCA.

The ensuing litigation and cases that interpret the safe harbor provisions therefore come as no surprise.¹⁷ What is surprising however is the relative paucity of research about takedown notices¹⁸—the procedural mechanism introduced in

effective July 1, 2006) 19 STATE COUNCIL GAZ. 468, at arts. 20–23, (China); *see also* Interpretations of the Supreme People’s Court on Several Issues Concerning the Application of Law in the Trial of Cases Involving Copyright Disputes over Computer Network (promulgated by the Sup. People’s Ct., Nov. 22, 2006, effective Dec. 8, 2006) (China).

¹⁷ *See* DANIEL SENG & IGNACIO GARROTE FERNÁNDEZ-DÍEZ, WORLD INTELL. PROP. ORG., COMPARATIVE ANALYSIS OF NATIONAL APPROACHES TO THE LIABILITY OF INTERNET INTERMEDIARIES FOR INFRINGEMENT OF COPYRIGHT AND RELATED RIGHTS (forthcoming 2014) (on file with the World Intell. Prop. Org.) (reviewing the legislation and caselaw that deal with Internet intermediary liability from Argentina, Australia, Belgium, Brazil, Canada, Chile, People’s Republic of China, Colombia, Egypt, Finland, France, Germany, India, Italy, Jamaica, Japan, Korea, Malaysia, Mexico, New Zealand, Nigeria, Russia, Rwanda, Senegal, Singapore, South Africa, Spain, United Kingdom and the United States); *see also* Béatrice Martinet Farano, *Internet Intermediaries’ Liability for Copyright and Trademark Infringement: Reconciling the EU and U.S. Approaches* (Transatlantic Technology Law Forum Working Papers No. 14, 2012), available at

http://www.law.stanford.edu/sites/default/files/publication/300252/doc/slsp_public/farano_wp14-4.pdf; M. Peguera, *The DMCA Safe Harbors and Their European Counterparts: A Comparative Analysis of Some Common Problems*, 32 COLUM. J.L. & ARTS 481 (2009).

¹⁸ In this paper, the term “notice” is used to refer to the legal document addressed by the complainant to the receiving organization such as an online service provider or an individual. The term used in 17 U.S.C. § 512(c) is “notification.” Google’s Copyright Transparency Report refers to these notices as “requests,” whereas Twitter’s Transparency Report refers to them as “notices.” *See FAQ: What is a Copyright Removal Request*,

the DMCA which empowers copyright owners to make takedown “requests”¹⁹ for each notice—to have service providers expeditiously remove or disable access to allegedly infringing content. This is a novel mechanism because it bypasses judicial oversight over copyright disputes. Despite academic claims that the takedown process is having little or no impact on piracy,²⁰ it is the mainstay of content providers for managing online infringement because it is fast, cheap, and efficient.

But little is actually known about the state of takedown notices today. Who are the copyright owners that use takedown notices? Who are the reporters serving the takedown notices on service providers? Are there different types of reporters? Who are the intermediaries who, as service providers, are served with the notices? Are there different types of notices? And which safe harbor provisions in the DMCA are addressed in these takedown notices and requests?

GOOGLE,
http://www.google.com/transparencyreport/removals/copyright/faq/#what_is_a_copyright_removal_request_is_a_copyright_removal_request (last visited Sept. 8, 2014).

¹⁹ This paper uses the term “requests” to refer to the “[i]dentification of the material that is claimed to be infringing or to be the subject of infringing activity and that is to be removed or access to which is to be disabled, and information reasonably sufficient to permit the service provider to locate the material” in 17 U.S.C. § 512(c)(3)(A)(iii). See *infra* text accompanying note 114.

²⁰ Another study concluded that anti-piracy efforts, including those through takedown notices, appear to have a limited impact on piracy among one-click hosts. See *generally* TOBIAS LAUINGER ET AL., INTERNATIONAL SECURE SYSTEMS LAB, CLICKONOMICS: DETERMINING THE EFFECT OF ANTI-PIRACY MEASURES FOR ONE-CLICK HOSTING (2012), available at <http://www.iseclab.org/papers/clickonomics.pdf>.

To date, the only study that has undertaken a comprehensive empirical study of takedown notices is the seminal 2006 study by Jennifer Urban and Laura Quilter.²¹ In reviewing 876 notices submitted to the Chilling Effects project until August 2005,²² the authors provided some answers to the above questions. In their dataset, which had 59 percent of 35 U.S.C. § 512(d) search notices and 36 percent of § 512(c) hosting notices,²³ they concluded that corporations and business entities were the primary users of takedown notices, but individuals constituted a significant minority.²⁴ They also found, somewhat surprisingly, that there was no significant use of § 512(c) hosting and §512(d) search notices by the movie and music industries.²⁵ The vast majority of these notices were sent by the rightholders themselves or their attorneys, with agents, enforcement agencies, and trade associations accounting for only 4.9 percent of all § 512(c) notices and only 1.3 percent of § 512(d) notices.²⁶

Since then, there has been an explosion in the use of takedown notices. It is no longer uncommon for Google and the so-called cyberlocker sites to receive and process thousands of takedown notices each month. And the sheer volume of these notices has also made it very difficult to undertake a rigorous empirical study of them. Part II of this Article briefly explains the methodology that is used to overcome this problem, describes the dataset and analysis used, and addresses

²¹ Jennifer M. Urban & Laura Quilter, *Efficient Process or “Chilling Effects”?* *Takedown Notices under Section 512 of the Digital Millennium Copyright Act*, 22 SANTA CLARA COMPUTER & HIGH TECH L.J. 621 (2006).

²² *Id.* at 641.

²³ *Id.* at 644.

²⁴ *Id.* at 652.

²⁵ *Id.* at 651.

²⁶ *Id.* at 654.

the possible limitations and ethical considerations arising therefrom. Part III summarizes the broad general findings from the analysis, including the nature and types of takedown notices and requests and the overall state of takedown notices in the relevant industries. Using notices addressed to Google for its case study, Part IV presents the findings of the notices and counter notices by the DMCA safe harbors addressed, and Part V concludes with some suggestions for changes to the notice and takedown mechanisms set up under § 512.

II. STUDY METHODOLOGY

A. Chilling Effects and Google Transparency Report

Although takedown notices form the lynchpin of the DMCA safe harbor scheme, the practical operation of it is not well researched. One reason why much discussion so far has been based on anecdotal evidence²⁷ is that the DMCA itself does not mandate the publication of the takedown notices or indemnify the parties who do so. Currently, service providers receiving takedown notices may have legitimate reasons not to publish them. For instance, notice publication by a hosting platform could be a precursor to a public admission of actual or constructive knowledge of the scale of infringing activity on its

²⁷ See, e.g., INTERNET POLICY TASK FORCE, U.S. DEP'T OF COMMERCE, COPYRIGHT POLICY, CREATIVITY, AND INNOVATION IN THE DIGITAL ECONOMY 57 (2013), available at <http://www.uspto.gov/news/publications/copyrightgreenpaper.pdf> (citing anecdotal reports which refer to the problem of inaccurate notices and surmising that the problem is one that is “subject to dispute”).

platform leading to possible loss of its safe harbor protection.²⁸ It may also attract lawsuits from copyright owners not originally parties to the notices. Likewise, a content provider may also be reluctant to publicly document its enforcement efforts because its disclosure amounts to the very publication of an infringing link or resource which the content provider is seeking to disable or remove.²⁹ Finally, the sheer volume of takedown notices sent has complicated efforts to document the notices and requests, let alone study them. While a court has ruled that such re-publications of takedown notices constitute fair use,³⁰ the ruling made by the court turned on a highly factual analysis which may yield different results in different circumstances.

Thankfully, notwithstanding these legal challenges, some companies, notably Google³¹ and Twitter,³² have made public commitments to publish their notices in the Chilling Effects project repository. The Chilling Effects notices

²⁸ 17 U.S.C. § 512(c)(1)(A)(i)–(ii) (2012). For instance, communications among the service providers' employees discussing uploaded content may constitute evidence of knowledge or awareness of specific instances of infringement. *Viacom Int'l, Inc. v. YouTube, Inc.*, 676 F.3d 19, 34 (2d Cir. 2012).

²⁹ See Jonathan Bailey, *Why I Don't Fear Chilling Effects . . . and You Shouldn't Either*, PLAGIARISM TODAY (Jan. 17, 2011), <http://www.plagiarismtoday.com/2011/01/17/why-i-dont-fear-chilling-effects-and-you-shouldnt-either>.

³⁰ *Perfect 10, Inc. v. Google, Inc.*, No. CV04-9484 AHM (SHx), 2010 WL 9479060, at *13 (C.D. Cal. July 30, 2010).

³¹ See, e.g., *Removing Content from Google*, GOOGLE, <https://support.google.com/legal/troubleshooter/1114905?hl=en> (last visited Sept. 8, 2014).

³² See *Copyright and DMCA Policy*, TWITTER, <http://support.twitter.com/articles/15795-copyright-and-dmca-policy> (last visited Sept. 8, 2014).

repository is a joint project between the Electronic Frontier Foundation and several U.S. law schools to document, index, tag, and make publicly available takedown notices and their detailed contents.³³ To conduct this study, the Chilling Effects repository was “crawled”³⁴ for individual takedown notices, using a reasonable interval of five to ten seconds between each notice to avoid overloading the Chilling Effects repository. The contents of each notice and its meta-information are saved by the crawlers onto a MySQL database. As Chilling Effects enables the submission of cease and desist notices from diverse categories, including domain name, cybersquatting, trademark, defamation, and patent disputes,³⁵ using the Chilling Effects–supplied meta-information, notices which are not related to copyright disputes are first filtered out. The residual data encompasses notices submitted for “copyright,” “DMCA safe harbor,” “anti-circumvention,” “derivative works,” and “fair use” disputes.³⁶

³³ CHILLING EFFECTS, <https://www.chillingeffects.org/> (last visited Sept. 8, 2014).

³⁴ “Crawling” refers to a process in which a computer program automatically collects web pages to create a local index or local collection of web pages. See JUNGHOO CHO & HECTOR GARCIA-MOLINA, THE EVOLUTION OF THE WEB AND IMPLICATIONS FOR AN INCREMENTAL CRAWLER, PROCEEDINGS OF 26TH INTERNATIONAL CONFERENCE ON VERY LARGE DATABASES (VLDB) (1999), available at <http://ilpubs.stanford.edu:8090/376/1/1999-22.pdf>.

³⁵ *If You Have Received A Cease and Desist Notice*, CHILLING EFFECTS, <https://www.chillingeffects.org/domain/input.cgi?print=yes> (last visited Sept. 8, 2014).

³⁶ Reliance is not placed solely on the Chilling Effects supplied meta-tags. By not limiting the dataset to just notices tagged as those relating to the “DMCA safe harbor,” this analysis can include those notices submitted by complainants which raise copyright, anti-circumvention, derivative works,

This data is further supplemented by the downloadable version (“raw data”) of Google’s takedown notices for web searches since 2011, published as part of Google’s Copyright Transparency Report (GTR).³⁷ It should be noted that the raw data does not contain all the entries in a takedown notice—it only contains some information such as the notice dates, names of the copyright owner and reporter, and the requests for which Google took no action. It should also be noted that the Google raw data is a subset of the Chilling Effects repository—the repository contains not just Google’s web search notices but also notices directed to other Google services, including Blogger, Picasa, and Shopping. In addition, while the raw data from the Transparency Report collects data starting from 2011, the Chilling Effects repository data dates back to 1997. Finally, there is a time lag between notices available via Google’s raw data and the Chilling Effects repository,³⁸ since notices submitted to the repository have to be first processed before they can be published.³⁹

and fair use issues which are not also tagged as “DMCA safe harbor” notices.

³⁷ The downloadable version of Google’s web search takedown notices, made available almost on a daily basis since July 2012, does not contain the entire content in takedown notices which are otherwise available on the Chilling Effects repository. However, it does provide a link to the full notice on the Chilling Effects repository.

³⁸ *See, e.g., Perfect 10, Inc. v. Google, Inc.*, No. CV04-9484 AHM (SHx), 2010 WL 9479060, at *2 (C.D. Cal. July 30, 2010) (noting Google’s assertion that all notices it receives are forwarded to the Chilling Effects website and that the missing notices are due to a backlog in Chilling Effects’ processing system).

³⁹ *See Legal Removal Requests*, GOOGLE, <https://support.google.com/legal/answer/3110420?hl=en> (last visited Sept. 24, 2014).

Overall, more than half a million notices ($N_0=545,545$) are captured in the base dataset, which is around 3.7 GB. For this study, the cut-off date for the notices in the base dataset is December 31, 2012.⁴⁰ The notices were additionally post-processed to correct for errors. This yielded a slightly reduced dataset of 539,000 notices ($N_1=539,558$) between January 2001 and December 2012 (“first dataset”).⁴¹

After a six-month long exploratory study, a decision was made to also conduct a census of all form-based⁴² notices submitted to Google within the first dataset.⁴³ Multiple parsers

⁴⁰ To deal with the time lag in processing notices as noted above, some notices from 2013 were also captured but were excluded in the final analysis. The actual data capture process ended on February 16, 2013 for a few reasons. The shortest time lag between the date of the notice and its publication by the Chilling Effects repository is about four business days. Some notices appear to take as long as one month to be published on Chilling Effects. In order to accommodate these “late” notices, the data capture process was extended to February 16, 2013.

⁴¹ Only notices with the “Topic Code” metafield marked in Chilling Effects as “anticircumvention,” “copyright,” “dmca512,” “derivative,” “fairuse,” or “response” were counted.

⁴² The notices are submitted in a variety of formats, ranging from fax to post, from email to web form. As will be illustrated later, since the vast majority of contemporary notices are in web form, the second dataset examines these notices exclusively.

⁴³ Twitter notices were excluded by design from the second dataset and not included in the automated parsing process because of a significant discrepancy in numbers. Only 2426 Twitter notices for 2012 were detected in the Chilling Effects’ repository. However, Twitter reports that 6646 notices have been served on them in 2012. See *Copyright Notices: January 1 – June 30, 2012*, TWITTER, <https://transparency.twitter.com/copyright-notices/2012/jan-jun> (last visited Sept. 8, 2014) [hereinafter *Twitter Transparency Report 1*]; *Copyright Notices: July 1 – December 31, 2012*, TWITTER, <https://transparency.twitter.com/copyright-notices/2012/jul-dec> [hereinafter *Twitter Transparency Report 2*] (last visited Sept. 8, 2014). This discrepancy could stem from a time or processing lag between

were developed by the author to read the contents of each notice (with about fifteen format variants) to extract all relevant fields, such as the names of the copyright owner, the reporting agent, and the recipient service provider, the service complained of, the nature of the copyright work allegedly infringed, the nature of the infringement, and the takedown requests.⁴⁴ In all, thirty-six fields are extracted from each of the 501,286 notices, or 91.89% percent of the base dataset.⁴⁵ Comprising 56,991,045 takedown requests, this dataset (“second dataset”) (N₂=501,286) with its indexes and relational data weighs in at around 6.8 GB (uncompressed size). It is this dataset that is used to analyze the notices and requests in Part IV of this paper.

B. Ethical Considerations and Limitations

Any empirical study is subject to various ethical considerations and limitations. This study is no exception.

One possible ethical issue that may arise stems from the fact that human subject data appear incidentally as the names of copyright owners or reporters of copyright infringement. The Chilling Effects repository addressed this by attempting to redact the names and identities of individual parties in the

Twitter’s own reporting of its takedown notices and its subsequent submission to the Chilling Effects’ repository.

⁴⁴ The parsing takes into account, as far as possible, the linguistic and formatting variations within each notice. For a complete list of the fields captured, see *infra* Part VII, List of Extracted Fields in Takedown Notices, available at http://vjolt.net/vol18/issue3/seng/Seng_Appendix01.pdf.

⁴⁵ To verify that the parsers worked as required, a visualization tool was developed to look at each parsed notice individually. Sampling was also done on the parsed notices to check for errors and statistical techniques were applied to detect outliers and adjust the parsing algorithms iteratively.

notices. Despite this, the parsers developed for this study managed to extract some individuals' names and identities from the notice contents. To preemptively address any privacy and ethical concerns, data about individuals was aggregated and any personally identifiable information was removed from the published observations. Names of individuals are not referred to in this paper, and Stanford's Institutional Review Board has reviewed and approved these methodologies.

Objections may also be raised about the lack of representativeness of the datasets. This study makes primary use of the Chilling Effects repository, where notice submissions are voluntary and only a small group of service providers like Google and Twitter have made public commitments to submit their notices on a regular basis. Thus, this study inherits the same limitations as the Chilling Effects repository. In particular, because Google notices predominate, the observations made here may not relate easily to other online service providers and to the online service provider industry in general.⁴⁶ The Chilling Effects repository also does not contain any notices received by hosting sites like Megaupload (before it was shut down) and Rapidshare, even though these sites—one study pins them as the top two file hosting sites, accounting for 25 percent of all file sharing traffic in from 2011 to 2012⁴⁷—acknowledge that they have received and processed millions of takedown notices from

⁴⁶ Thus, the first dataset contains only six Facebook takedown notices and no notices directed to Microsoft/Bing. The six Facebook takedown notices appear to have been email notices submitted to Chilling Effects by their reporter.

⁴⁷ Josep Sanjuàs-Cuxart et al., *Measurement Based Analysis of One-Click File Hosting Services*, 20 J. OF NETWORK AND SYS. MGMT., 276, 277 (2012), available at <http://personals.ac.upc.edu/jsanjuas/papers/httpfs.pdf>.

content providers.⁴⁸ Nor does the repository include the substantial numbers of takedown notices directed to Usenet service providers,⁴⁹ which have witnessed substantial litigation over their legal liability in Europe.⁵⁰ Nonetheless, where a takedown notice has been served against the hosting or aggregating site, the content providers will most likely serve similar takedown notices against a widely used search engine like Google⁵¹ to prevent the same content from being located.⁵²

⁴⁸ See Superseding Indictment at 11, *United States v. Kim Dotcom*, Criminal No. 1:12CR3 (E.D. Va. Feb. 16, 2012), available at [http://www.justice.gov/usao/vae/victimwitness/mega_files/Certified Mega Superseding Indictment \(2-16-2012\).pdf](http://www.justice.gov/usao/vae/victimwitness/mega_files/Certified%20Mega%20Superseding%20Indictment%20(2-16-2012).pdf).

⁴⁹ Of the four key Usenet/newsgroup service providers—EasyNews, Giganews, Astraweb, and Newzbin—there is only one notice directed to Giganews in the Chilling Effects repository. See *Giganews Utilizes Section 512*, CHILLING EFFECTS (Feb. 19, 2002), <http://www.chillingeffects.org/dmca512c/notice.cgi?NoticeID=173>. This is the case despite the substantial increase in such takedown notices that Usenet service providers are receiving. See Enigmax, *Usenet Feels the Heat as Copyright Holders Try to Strip Away Content*, TORRENTFREAK (Nov. 9, 2012), <http://torrentfreak.com/usenet-feels-the-heat-as-copyright-holders-try-to-strip-away-content-121109>. One reporting organization claims to have sent out over five million takedown notices to Usenet service providers as of May 2013. See *Usenet Stats*, MORGANELLI GROUP, <http://www.morganelligroup.com/stats.php> (last visited Sept. 13, 2014).

⁵⁰ See, e.g., *Twentieth Century Fox Film Corp. v. British Telecomm.*, [2011] EWHC 1981 (Ch) (Eng.); *Twentieth Century Fox Film Corp. v. Newzbin Ltd.*, [2010] EWHC 608 (Ch) (Eng.).

⁵¹ See Danny Sullivan, *Google Still World's Most Popular Search Engine By Far, but Share of Unique Searchers Dips Slightly*, SEARCH ENGINE LAND (Feb. 11, 2013, 9:00 AM), <http://searchengineland.com/google-worlds-most-popular-search-engine-148089>.

⁵² See Grant Crowell, *Copyright Law: What Search Marketers Should Know (Part I)*, SEARCH ENGINE WATCH (Sept. 26, 2007), <http://searchenginewatch.com/article/2064501/Copyright-Law-What-Search-Marketers-Should-Know-Part-1>; see also Chris Cameron, *The British Are Coming! (To Serve Google a DMCA Notice)*, READWRITE (June

Examining the takedown requests received by Google does arguably provide good empirical insights into the identities and nature of the sites targeted by content providers and reporters.

Another limitation is that the Chilling Effects repository does not contain takedown notices as part of Google/YouTube's Content ID⁵³ system. Content ID notices, which form the bulk of video-related notices,⁵⁴ fall outside the rubric of the DMCA,⁵⁵ although a recent change by YouTube to require a content provider to file a formal DMCA takedown in response to an uploader's challenge to a Content ID takedown⁵⁶ has helped to bridge the gap. But given the anecdotal evidence that millions of Content ID takedowns are generated yearly and that many of such takedowns involve

22, 2010),
http://readwrite.com/2010/06/22/the_british_are_coming_to_serve_google_a_dmca_notice.

⁵³ The Content ID system works because participating content providers have submitted ten million digital fingerprints of their audio and video works to YouTube, which are matched against every uploaded video. See *How ContentID Works*, YOUTUBE, <https://support.google.com/youtube/answer/2797370?hl=en> (last visited Sept. 14, 2014).

⁵⁴ Hundreds of thousands of allegedly infringing videos have been taken down by content providers through YouTube's Content ID system. *A Guide to YouTube Removals*, ELEC. FRONTIER FOUND., <https://www.eff.org/issues/intellectual-property/guide-to-youtube-removals> (last visited Sept. 14, 2014).

⁵⁵ *Id.*

⁵⁶ Thabet Alfishawi, *Improving Content ID*, YOUTUBE (Oct. 3, 2012), <http://youtubecreator.blogspot.com/2012/10/improving-content-id.html>. See also Parker Higgins, *YouTube Upgrades Its Automated Copyright Enforcement System*, ELEC. FRONTIER FOUND. (Oct. 5, 2012), <https://www.eff.org/deeplinks/2012/10/youtube-upgrades-its-automated-copyright-enforcement-system>.

potentially works of nature,⁵⁷ public domain works,⁵⁸ or raise fair use issues,⁵⁹ their absence from the Chilling Effects repository is conspicuous.⁶⁰

In addition, while the Chilling Effects repository is not explicitly jurisdiction-specific, the bulk of the notices follow the DMCA, primarily because most of the contributing service providers are U.S. corporations. But as most jurisdictions around the world have adopted safe harbor laws that are based on the DMCA,⁶¹ the conclusions drawn from this repository remain highly relevant in these jurisdictions.

⁵⁷ See Mike Masnick, *Rumblefish CEO: Claiming Copyright on Your Incidental Recordings of Birds Was Merely a Series of Unfortunate Errors*, TECHDIRT (Feb. 27, 2012, 2:35 PM), <http://www.techdirt.com/articles/20120227/13044117890/rumblefish-ceo-claiming-copyright-your-incidental-recordings-birds-was-merely-series-unfortunate-errors.shtml>.

⁵⁸ See Timothy B. Lee, *How YouTube Lets Content Companies "Claim" NASA Mars Videos*, ARS TECHNICA (Aug. 8, 2012, 12:25 PM), <http://arstechnica.com/tech-policy/2012/08/how-youtube-lets-content-companies-claim-nasa-mars-videos>.

⁵⁹ See, e.g., Adam Holland, *Germans Unable to Watch Dashboard Cam Videos of Chelyabinsk Meteor*, CHILLING EFFECTS (Feb. 20, 2013), <http://chillingeffects.org/weather.cgi?WeatherID=712> (reporting that the vehicle dashboard cameras that captured the Chelyabinsk meteor also recorded background music playing in the vehicles, and in Germany, the Content ID system makes the assumption that the German music rights association, GEMA, controls access to these songs and has not granted the requisite publishing rights).

⁶⁰ A review of the first dataset confirms that the Chilling Effects repository only has twelve notices sent directly by content providers and reporters to YouTube. But a perusal of these notices confirms that none of these are sent as part of the Content ID system.

⁶¹ See SENG & FERNÁNDEZ-DÍEZ, *supra* note 17.

Finally, in the process of analyzing these notices, data formatting inconsistencies and corruption of the contents of a small number of notices (fewer than a thousand) in the Chilling Effects repository were detected. Some notices also have corrupted or scrambled metadata such as notice dates and sender and copyright owner identities. To deal with this problem, new metadata extracted from the notice contents was cross-referenced against and, where appropriate, used to supplement the suspect metadata.

III. TAKEDOWN NOTICES IN GENERAL

Section 512(b) of the DMCA exempts an online service provider from monetary damages for copyright infringement by reason of caching an unauthorized copy of the material on its system or network.⁶² Likewise, § 512(c) of the DMCA exempts an Internet intermediary, such as an online hosting service provider, from monetary damages for copyright infringement by reason of its users' storage of material that resides on the provider's system or network.⁶³ Similarly, § 512(d) of the DMCA exempts an Internet intermediary as a service provider of information location tools, such as search engines or portals, from monetary damages for referring or linking users to an online location containing infringing material or infringing activity.⁶⁴ To qualify for these safe harbors, service providers have to act expeditiously to remove or disable access to material or activity claimed to be infringing upon notification of claimed infringement.⁶⁵ This notification

⁶² 17 U.S.C. §§ 512(b)(1), (2)(E) (2012).

⁶³ *Id.* § 512(c)(1).

⁶⁴ *Id.* § 512(d).

⁶⁵ *Id.* §§ 512(b)(2)(E), (c)(1)(C), (d)(3).

served on the service provider's "designated agent"⁶⁶ is the takedown notice. Of course, the service provider is free to refuse to act on the notification. In such a case, the service provider loses the protection of the safe harbor provisions, and its liability will be decided without reference to § 512.⁶⁷

Analysis of the first dataset ($N_0=545,545$) shows that only a small number of service providers—Google and Twitter—consistently contribute their takedown notices to the Chilling Effects repository. The number of notices, summarized in Table 1,⁶⁸ confirms what is already widely known—Google receives a very large number of takedown notices. From 67,571 notices in 2011, this number jumped six-fold to 441,370 notices in 2012. Large numbers of takedown notices seem to reflect the maturation of the service provider's services. For instance, Digg and Twitter were founded in 2004 and 2006, respectively, but it was only in 2008 and 2010 that they started receiving significant numbers of takedown notices. Similarly, Google only saw its first thousand takedown notices in 2009, more than ten years after its incorporation.⁶⁹

Table 1 shows that the year-on-year increase in the number of Google's notices is 304 percent for 2010, 305 percent for 2011, and 524 percent for 2012. Google, however, is not the only service provider bearing the brunt of the increase. Twitter has also seen year-on-year increases in

⁶⁶ *Id.* § 512(c)(2).

⁶⁷ H.R. REP. NO. 105-551, pt. 2, at 54 (1998).

⁶⁸ See *infra* Part VIII, Table 1, available at http://vjolt.net/vol18/issue3/seng/Seng_Table01.pdf.

⁶⁹ The small number of notices it received in 2008 could be an aberration. It could be that Google failed to submit all its notices to Chilling Effects in 2008, or that they were dated erroneously.

takedown notices—1,248 percent for 2011 and 61 percent for 2012. [Figure 1](#)⁷⁰ illustrates this increase.

The substantial increase in takedown notices that Twitter has received since 2010 can be better seen by excluding Google notices from [Figure 1](#), as shown in [Figure 2](#).⁷¹

Are these increases related to the debacle pertinent to the Stop Online Piracy Act (SOPA) and the PROTECT IP Act (PIPA)⁷² and their failure to pass legislative muster?⁷³ There is some support for that view. Analyzed on a monthly basis, [Figure 3](#)⁷⁴ shows that there is a spike in the volume of notices directed at Google—the first post-SOPA month-on-month increase was in January 2012 at 88.7 percent (17,164 notices). Additionally, there was a second month-on-month increase, which was even more substantial, around April 2012 at 227 percent (45,551 notices). Perhaps this is indicative of ramped up enforcement post-SOPA/PIPA, because since then, Google has been steadily processing around 30,000 to 40,000 notices each month, hitting a peak of 60,114 notices in October 2012.

⁷⁰ See *infra* Part IX, [Figure 1](#), available at http://vjolt.net/vol18/issue3/seng/Seng_Figure01.pdf.

⁷¹ See *infra* Part IX, [Figure 2](#), available at http://vjolt.net/vol18/issue3/seng/Seng_Figure02.pdf.

⁷² See, e.g., John D. Sutter, *Why Wikipedia Went Down at Midnight*, CNN (Jan. 18, 2012, 4:59 PM), <http://edition.cnn.com/2012/01/17/tech/web/wikipedia-sopa-blackout-qa/index.html>.

⁷³ See Mike Masnick, *Funny How Copyright Holders Only Ramped Up Google DMCA Takedowns After SOPA Failed*, TECHDIRT (Dec. 13, 2012, 10:23 AM), <https://www.techdirt.com/articles/20121212/22445321369/funny-how-copyright-holders-only-ramped-up-google-dmca-takedowns-after-sopa-failed.shtml>.

⁷⁴ See *infra* Part IX, [Figure 3](#), available at http://vjolt.net/vol18/issue3/seng/Seng_Figure03.pdf.

Takedown notices received by Twitter show a similar trend of coordinated enforcement. As [Figure 4](#)⁷⁵ illustrates, there are two month-on-month increases in January 2012 at 252 percent (437 notices) and in April 2012 at 83.3 percent (700 notices).⁷⁶ But there is also a peak of 1394 notices received by Twitter in January 2011, representing an almost unprecedented 386 percent increase from December 2010. A detailed look into these notices shows that 90 percent of them came from the reporting organization Web Sheriff and were issued for two copyright owners—Magnolia Pictures for films and movies and Beggars Group & XL Recordings for sound recordings.⁷⁷ Since this represented a ten-fold increase in enforcement from the previous month, it strongly suggests that starting January 2011, copyright owners and reporters have decided to put Twitter squarely in their crosshairs.

⁷⁵ See *infra* Part IX, Figure 4, available at http://vjolt.net/vol18/issue3/seng/Seng_Figure04.pdf.

⁷⁶ These figures are taken from Twitter's own Transparency Report for the first half of 2012. See *Twitter Transparency Report 1*, *supra* note 43.

⁷⁷ For instance, notices that originated from Web Sheriff for Magnolia Pictures to Twitter were all sent on January 31, 2011 but while Web Sheriff sent 140 notices in December 2010, it sent 1098 notices to Twitter in the last two weeks of January 2011. See, e.g., *Web Sheriff DMCA (Copyright) Complaint to Twitter*, Chilling Effects (Jan. 31, 2011), <http://www.chillingeffects.org/dmca512/notice.cgi?NoticeID=59118> (example of notice sent on behalf of Magnolia Pictures); *Web Sheriff DMCA (Copyright) Complaint to Twitter*, Chilling Effects (Jan. 31, 2011), <http://www.chillingeffects.org/dmca512/notice.cgi?NoticeID=58731> (example of notice sent on behalf of Beggars Group & XL Recordings).

A. Content Providers

Who are the parties sending these takedown notices? And which industries⁷⁸ do they represent? The DMCA uses the term “complaining party” to refer to the party sending takedown notices.⁷⁹ The DMCA goes on to refer to the complaining party as either the “copyright owner” or “a person authorized to act on behalf of the copyright owner.”⁸⁰ So one way to study these questions is to ask who the complaining parties are—content providers, copyright owners, or their industry representatives (referred to collectively as “providers”).

A longitudinal study of the first dataset (N1=539,558) reveals some interesting facts, as set out in [Table 2](#).⁸¹

First, in terms of overall figures, the top provider of takedown notices is the British Phonographic Industry (BPI), the British record industry’s trade association. In contrast, the International Federation of the Phonographic Industry (IFPI) is placed fifth and the Recording Industry Association of America (RIAA) is placed thirteenth. In all, the music industry is well

⁷⁸ For purposes of this paper, the industries are classified as the following: music (musical compositions and sound recordings), movies (motion pictures, including soundtracks and music videos), broadcasts (television broadcast, cable and satellite programming, including sports programming), adult entertainment (works for adult entertainment purposes, usually audiovisual works), books (literary publications), software (computer programs, excluding video games), games (usually video games), and websites (where the source site cannot be classified into any of the above categories, e.g., instructional sites such as Lynda.com and information aggregation sites such as RipOff Reports).

⁷⁹ See 17 U.S.C. §§ 512(c)(3)(A)(iv)–(vi) (2012).

⁸⁰ *Id.* § 512(c)(3)(B)(i).

⁸¹ See *infra* Part VIII, Table 2, available at http://vjolt.net/vol18/issue3/seng/Seng_Table02.pdf.

represented on this list by three collective management organizations (CMOs). Representing the book industry, the Publishers Association is only placed twelfth by total notices issued.

In addition to music and books, the other spots in the top-twenty list are dominated by two other different industries: adult entertainment providers (e.g., Froytal, CA Co., RK NetMedia, Bang Brothers, GG Cash) and movie studios (e.g., 20th Century/Fox, HBO, Columbia Pictures, Sony Pictures, Magnolia Pictures, Lionsgate, Screen Gems, Paramount Pictures). In fact, the Intellectual Property Promotion Association (IPPA), the trade association for 250 Japanese adult entertainment movie production companies,⁸² only started its enforcement action in 2012 but quickly made its presence felt.⁸³ The music industry, however, is by far the most active issuer of takedown notices. Together, the BPI, IFPI, and RIAA account for 58.6 percent of all notices served between 2008 and 2012, whereas the aforementioned adult entertainment providers only account for 19.8 percent and their movie studio counterparts only account for 9.5 percent. (For a visual breakdown, see [Figure 5](#).⁸⁴) If notice numbers are used as a proxy for measuring enforcement activity, during this period, almost six in ten notices from the top-fifty content providers pertain to music infringement. And if notices from the movie industry (but excluding the adult entertainment industry) are included, 68.1 percent of all notices are sent by

⁸² See Kaoriko Okuda, *Huge Lawsuits Seek to Deter Sellers of Pirated DVDs*, ASAHI SHIMBUN (Sept. 30, 2011), http://ajw.asahi.com/article/behind_news/social_affairs/AJ2011093011409.

⁸³ The Intellectual Property Promotion Association placed sixth in terms of total notices issued.

⁸⁴ See *infra* Part IX, Figure 5, available at http://vjolt.net/vol18/issue3/seng/Seng_Figure05.pdf.

the music and movie industries. The fact that these two industries have, within six years, become the largest senders of takedown notices represents a stark change from the findings of the Urban and Quilter study in 2006.

The top fifty providers have also stepped up their online enforcement. While they accounted for only 23.9 percent of all the notices issued in 2010, this figure rose to 36.8 percent in 2011, and this figure reached 74.7 percent in 2012. Among the top fifty providers, the music industry providers accounted for almost half (45.9 percent) of *all* the notices issued in 2012. Not far behind lies the adult entertainment industry, at 15.3 percent in 2012, with the movie industry at a distant third with 6.6 percent.

However, lest the view be taken that only large industry associations and content providers use the takedown notice mechanism, [Table 3](#)⁸⁵ shows the maximum, average, and median number of notices issued by each identifiable content provider between 2008 and 2012.

The large number of notices issued by providers should come as no surprise by now. In 2012, one provider (BPI) was responsible for issuing close to 200,000 notices, which raised the average number of notices issued per provider to around twenty-five (see [Table 3](#)). In reality, however, the majority of providers issue just one notice—the median number of notices for providers between 2001 and 2012 is consistently one. Another way to examine this is by way of the cumulative distribution frequency (CDF) of number of notices per provider. The CDF chart in [Figure 6](#)⁸⁶ shows that in 2012, 62.9

⁸⁵ See *infra* Part VIII, [Table 3](#), available at http://vjolt.net/vol18/issue3/seng/Seng_Table03.pdf.

⁸⁶ See *infra* Part IX, [Figure 06](#), available at http://vjolt.net/vol18/issue3/seng/Seng_Figure06.pdf.

percent of all providers sent only one notice, 87.1 percent of all providers sent no more than five notices, and 91.7 percent of all providers sent no more than ten notices. In other words, cumulatively, more than nine out of ten notices are sent out by content providers who send fewer than ten notices each year. This suggests that content providers like BPI, Froytal, CA Co., and Microsoft, who do issue tens and hundreds of thousands of takedown notices yearly (the so-called “large content providers”) are really in the minority, although they do submit more takedown *requests* in total than the other content providers (“small content providers”), as is subsequently shown in this Article.

B. Reporters and Reporting Agents

As noted above, the DMCA uses the term “complaining party” to refer to the party sending takedown notices.⁸⁷ Because the DMCA allows “a person authorized to act on behalf of the copyright owner”⁸⁸ to be the complaining party, the party submitting the takedown notice need not be the content provider, its licensee, or its industry representative. This party can be an attorney of the content provider or a specialist entity known as a “reporting agent” who detects online infringement and sends takedown notices to the relevant service providers on behalf of the owners or licensees as clients. This part of the study will refer to the “complaining party” who is actually responsible for submitting the notice to the Internet intermediary as the “reporter.”⁸⁹ In contrast, the

⁸⁷ See 17 U.S.C. §§ 512(c)(3)(A)(iv)–(vi) (2012).

⁸⁸ *Id.* § 512(c)(3)(B)(i).

⁸⁹ The Google Transparency Report (GTR) refers to this party as the “reporting organization.” See *Google Transparency Report: Reporting Organizations*, GOOGLE,

DMCA refers to this party as “the person making the notification.”⁹⁰ As reporters, they need not be licensed to use the content referred to in the takedown notice but have to be authorized to serve the takedown notice on behalf of the copyright owner or its exclusive licensee.⁹¹

The rise of “reporting agents” is illustrated by a longitudinal study of reporters in [Table 4](#).⁹²

Between 1998 and 2004, the use of reporting agents to submit takedown notices was very limited.⁹³ Reporting agents appear to first arrive on the scene in 2005. Since then, there has been a proliferation of reporting agents. Between 2008 and 2012, reporting agents that make up the top thirty reporters account for between 36.8 percent and 59.6 percent of all notices issued yearly. Although BPI, IFPI, the Publishers Association, and RIAA still generate their own takedown notices (and their stepped up efforts in 2012 accounted for 45.7 percent of all notices served), the undeniable trend is for the continued use of reporting agents to detect infringement and issue takedown notices on behalf of copyright owners and their licensees.

<http://www.google.com/transparencyreport/removals/copyright/reporters>
(last visited Sept. 14, 2013).

⁹⁰ See 17 U.S.C. § 512(c)(3)(B)(ii).

⁹¹ *Id.* § 512(c)(3)(A)(vi).

⁹² See *infra* Part VIII, Table 4, available at http://vjolt.net/vol18/issue3/seng/Seng_Table04.pdf.

⁹³ See Urban & Quilter, *supra* note 21, at 654 (noting that 94 percent and 98.5 percent of all § 512(c) and § 512(d) notices, respectively, were sent by or on behalf of the rightsholders directly).

But reporting by small content providers and individuals⁹⁴ is arguably still alive and thriving.⁹⁵ For instance, the distribution spread of notices issued by reporters is very similar to that of content providers. The median number of notices issued by reporters between 2008 and 2012 is still one, even though the average number of notices issued by reporters has risen from 3.28 in 2008 to 44.46 in 2012. As illustrated in Table 5⁹⁶ and Figure 7,⁹⁷ the large spread (i.e., standard deviation) of the number of notices issued in 2012 suggests that notwithstanding the use of reporting agents and automated processes, there is considerable reporting other than by large content providers and reporting agents. Thus, more than 65 percent of all reporters have only issued one notice and almost 95 percent of all reporters have not issued more than ten notices. This is corroborated by the observation that in Table 4, individuals as reporters contribute as a group around 5.3 percent of all notices and are within the list of top thirty reporters. Table 4 also shows that one individual reporter (the Malaysian blogger who operates the technology site RaymondCC Tech) single-handedly issued 1189 notices in 2012.⁹⁸ This puts him in the same league as reporting agents such as Police Du Net and Ripoff.

⁹⁴ It is assumed here that individual reporters are those whose identities as reporters are redacted in the notices on the Chilling Effects repository.

⁹⁵ No observations are made about those content providers who do *not* issue any takedown notices.

⁹⁶ See *infra* Part VIII, Table 5, available at http://vjolt.net/vol18/issue3/seng/Seng_Table05.pdf.

⁹⁷ See *infra* Part IX, Figure 7, available at http://vjolt.net/vol18/issue3/seng/Seng_Figure07.pdf.

⁹⁸ See *Blog DMCA (Copyright) Complaint to Google*, CHILLING EFFECTS (May 28, 2012), <http://www.chillingeffects.org/dmca512c/notice.cgi?NoticeID=385210>.

C. Takedown Notice Formats

The DMCA provides that to be an effective notification, a takedown notice “must be a written communication.”⁹⁹ The DMCA clearly envisages electronic communications as written communications because it allows for these written communications to be signed by way of “a physical or electronic signature.”¹⁰⁰ So a very important issue with practical ramifications is to investigate the format in which DMCA takedown notices are communicated to the recipient. Using the meta-information available on the Chilling Effects repository, [Table 6](#)¹⁰¹ shows the proportion of notices sent to service providers in web or online form, in email form, by way of postal mail, by way of fax, or in some other form.

It is clear that even though notifications by post or fax count as “written communication,” in practice, complainants and service providers overwhelmingly prefer web form and email communications. The reasons for doing so are simple. Takedown requests that supply “information reasonably sufficient to permit the service provider to locate the material”¹⁰² can be long and complex.¹⁰³ Having takedown requests in electronic form ensures that the correct location is used and minimizes errors in processing such requests. There is anecdotal evidence to suggest that despite some disquiet on the part of reporters, service providers will give preferential

⁹⁹ 17 U.S.C. § 512(c)(3)(A) (2012).

¹⁰⁰ *Id.* § 512(c)(3)(A)(i).

¹⁰¹ See *infra* Part VIII, Table 6, available at http://vjolt.net/vol18/issue3/seng/Seng_Table06.pdf.

¹⁰² *Id.* § 512(c)(3)(A)(iii).

¹⁰³ In particular, requests that take the form of Uniform Resource Indicators (URIs) can be long, complex, and technical, as some of these URIs are computer generated to ensure that they are unique references to accessible content. See *infra* text accompanying note 117.

treatment to electronic notices or refuse to act on non-form and non-email communications, unless they are also presented with their electronic equivalents.¹⁰⁴ Such a demand does not appear to be unreasonable, especially where the service provider receives increasingly large numbers of takedown notices with very long takedown requests.¹⁰⁵ In fact, it is entirely logical for the service provider, as a response to the use of automated systems to detect and issue large numbers of takedown notices, to avoid interpretational ambiguities and streamline their own processes through the use of web form notices.¹⁰⁶

This anecdotal evidence is confirmed by an empirical analysis of the first dataset. As Table 6 shows, email notices dominate up until 2010. However, since 2011, their relative numbers have gradually decreased. Figure 8,¹⁰⁷ which presents the monthly figures of the notices by their formats, shows that starting from January 2011, the majority of notices were web form notices and they have since displaced all other notice formats, including emails.

¹⁰⁴ For examples of such anecdotal evidence, see, for example, Sara Hawkins, *How to File a DMCA Takedown Notice*, SARA F. HAWKINS (Oct. 4, 2012), <http://sarafhawkins.com/how-to-file-a-dmca-takedown-notice>. There are, however, online service providers who still prefer fax and letter-based takedown notices. See, e.g., *Creative Commons DMCA Notice & Takedown Procedure*, CREATIVE COMMONS, <http://creativecommons.org/dmca> (last visited Sept. 14, 2014).

¹⁰⁵ Analysis of the second dataset shows that the average length of a URI as a takedown request in the second dataset is 79.31 characters in 2011 and 75.88 characters in 2012.

¹⁰⁶ For instance, the parsers developed for this study had difficulties with email notices when breaks and hyphens were inserted (either manually or as part of an automated process) to make long URIs more “readable.” There may have also been confusion when some URIs had other URIs “embedded” in them. See *infra* text accompanying note 121.

¹⁰⁷ See *infra* Part IX, Figure 8, available at http://vjolt.net/vol18/issue3/seng/Seng_Figure08.pdf.

Nor is this only a trend with notices to Google. The numbers and proportion of web form notices to Twitter, Yahoo!, and Digg as service providers have also increased and, as [Figure 9](#)¹⁰⁸ and [Figure 10](#)¹⁰⁹ show, since January 2011, web form notices have prevailed over all other notice formats.

D. Takedown Requests in General

To be an effective takedown notice under the DMCA, it is a formal requirement for the notification to identify the copyrighted work claimed to have been infringed.¹¹⁰ The DMCA does not describe what constitutes the relevant information for this identification, but most reporters supply, at the very least, the title of the copyright work, its authorship, owner, performer, or publisher information. More detailed information will include a brief description of the copyright work and the nature or type of copyright work in question.¹¹¹ A failure to supply the copyrighted work information will render the notice deficient and ineffective.¹¹²

In practice, many service providers refer to this component of the takedown notice as a “claim,” and a takedown notice can have many claims. In turn, as takedown requests are associated with each copyrighted work, each claim

¹⁰⁸ See *infra* Part IX, Figure 9, available at http://vjolt.net/vol18/issue3/seng/Seng_Figure09.pdf.

¹⁰⁹ See *infra* Part IX, Figure 10, available at http://vjolt.net/vol18/issue3/seng/Seng_Figure10.pdf.

¹¹⁰ 17 U.S.C. § 512(3)(3)(A)(ii).

¹¹¹ Some online service providers may also require the reporter to supply the URI(s) where a licensed copy of the original copyright work could be located. See *Perfect 10, Inc. v. Google, Inc.*, No. CV04-9484 AHM (SHx), 2010 WL 9479059, at *8 (C.D. Cal. July 26, 2010) (referring to Uniform Resource Locators (URLs) that reference the copyrighted work).

¹¹² *Id.* (ruling that the Group A notices which “uniformly do not identify specifically which copyrighted works were infringed” were ineffective).

can have many takedown requests. For a § 512(c) notice (and a § 512(b) notice),¹¹³ the DMCA describes this takedown request as “[i]dentification of the *material* that is claimed to be infringing or to be the subject of infringing activity and that is to be removed or access to which is to be disabled, *and* information reasonably sufficient to permit the service provider to *locate the material*.”¹¹⁴ For a § 512(d) notice, the DMCA describes this as “identification of the *reference or link*, to material or activity claimed to be infringing, that is to be removed or access to which is to be disabled, *and* information reasonably sufficient to permit the service provider to *locate that reference or link*.”¹¹⁵ Thus for a § 512(b) notice, a § 512(c) notice and a § 512(d) notice, the location information is the *necessary* unit of information which the reporter has to supply and for which a recipient service provider is requested to act on.¹¹⁶

Most service providers equate the request with the location information, and the location information with a Uniform Resource Indicator (URI). A URI is an Internet address, and as an engineering unit, it unambiguously resolves to a specific location where online resources can be found.¹¹⁷

¹¹³ 17 U.S.C. § 512(b)(2)(E) (importing the requirements in § 512(c)(3) into a § 512(b) notice).

¹¹⁴ *Id.* § 512(c)(3)(A)(iii) (emphasis added).

¹¹⁵ *Id.* § 512(d)(3) (emphasis added).

¹¹⁶ In *Google*, the court seemed to initially take the view that the identification information for the infringing material is less onerous for a § 512(c) notice than a § 512(d) notice. 2010 WL 9479059, at *14. However, the court accepted Google’s argument as tenable that in both instances, the URL has to be supplied as information “reasonably sufficient” to permit the location of the material. *Id.*

¹¹⁷ The URI is a string of characters used to identify a name or an Internet resource. A URI can be classified as URL, otherwise known as a web

Not only does a URI include a Uniform Resource Locator (URL)¹¹⁸ and therefore allows for a web resource to be located, it also includes other resources like Usenet¹¹⁹—the worldwide distributed Internet discussion system—and therefore can be used to locate Usenet messages. Thus, issues faced by the owner/reporter in *Perfect 10, Inc. v. Giganews, Inc.* in failing to supply location information to locate the offending Usenet message can be dealt with by framing the Message-ID as a URI.¹²⁰ But that is not to say that URIs are perfect; reporters can supply incomplete or truncated URIs or non-work specific URIs.¹²¹

address, or a Uniform Resource Name (URN). A URI therefore includes a URL. See M. MEALLING & R. DENENBERG, W3C UNIFORM RESOURCE IDENTIFIER INTEREST GROUP, REPORT FROM THE JOINT W3C/IETF URI PLANNING INTEREST GROUP: UNIFORM RESOURCE IDENTIFIERS (URIs), URLs, AND UNIFORM RESOURCE NAMES (URNs): CLARIFICATIONS AND RECOMMENDATIONS 2 (2002), available at <http://tools.ietf.org/html/rfc3305.pdf> (explaining the differences between a URI and URL). In this study, for the most part, the two terms can be used interchangeably, since the resources referred to here are primarily web pages and resources.

¹¹⁸ *Id.*

¹¹⁹ See M. Horton & R. Adams, *Standard for Interchange of USENET Messages*, (Network Working Group, Request For Comments: 1036, 1987), available at <http://tools.ietf.org/html/rfc1036.pdf>; see also P. Hoffman, *The News and nntp URI Schemes* (VPN Consortium, Network Working Group, Internet Draft 4, 2005), available at <http://tools.ietf.org/html/draft-hoffman-news-nntp-uri-04.pdf>.

¹²⁰ *Perfect 10, Inc. v. Giganews, Inc.*, 993 F. Supp. 2d 1192, 1201 (C.D. Cal. 2014) (drawing the analogy between a Message-ID, which is the only unique identifier to locate Usenet messages, and the URL).

¹²¹ See *Google*, 2010 WL 9479059, at *8, *15 (stating, for instance, references that lack image-specific URLs); *Perfect 10 v. Google, Inc.*, 416 F. Supp. 2d 828, 854 (C.D. Cal. 2006).

Supplying some other location information in place of the URI may subject the service provider to an “untenable burden,”¹²² since “the burden of policing copyright infringement—identifying the potentially infringing material and adequately documenting infringement—[rests] squarely on the owners of the copyright [and reporters].”¹²³ Courts are conscious of the volume of takedown requests received by the service providers and the need for service providers to takedown the infringing materials expeditiously to stay within the safe harbors. Consequently, courts have frowned upon reporters who supply location information in the form of verbose search instructions, thumbnail images, and screen shots,¹²⁴ or supplying location information spread across separate notices or documents.¹²⁵ There is a very good reason for the use of URIs by both the reporter and the service provider.

But the limits of equating the location information with a URI must be understood. After all, the DMCA House and Senate Reports very carefully refer to “a copy or description of the allegedly infringing material *and* the URL address of the

¹²² *Giganews*, 993 F. Supp. 2d at 1201.

¹²³ *Perfect 10, Inc. v. CCBill LLC*, 488 F.3d 1102, 1113 (9th Cir. 2007).

¹²⁴ *Giganews*, 993 F. Supp. 2d at 1201 (rejecting this supplied information on the basis that they did not constitute unambiguous identification information and because extracting the requisite Message-ID would be an untenable burden on the service provider).

¹²⁵ See *CCBill*, 488 F.3d at 1113 (holding that an adequate notice must be constituted from “a written communication” and not separate communications); *Google*, 2010 WL 9479059, at *9, *11 (holding that Group B notices did not create an undue burden as all the information is provided in one file but that the Group C notices were defective because they did not contain all the required information in a single written communication).

location . . . alleged to contain the infringing material” as “an example of such *sufficient* information.”¹²⁶ For instance, allegedly infringing resources online may be described using search terms that link to the infringing materials on a search engine. Indeed, analysis of the second dataset shows that there are *only* sixty-nine URIs (out of more than fifty-six million) that take the form of search requests on Google.¹²⁷ However, as these URIs will *not* resolve to a specific location for the infringing resource but rather to search result pages that present a list of resources, not all of which are alleged to be infringing, Google has rejected almost all these takedown requests. As the court in *Giganews* pithily observed: “[W]hile a web search may ‘find’ a number of results, the search itself does not actually locate the items found; the search engine just presents its search results in a list, and any item in that list is not ‘locate’ until its URL is extracted.”¹²⁸

There is another reason why search terms phrased as URIs or, more broadly, as any location information, must be rejected.¹²⁹ While this is desirable from the copyright owner or

¹²⁶ S. REP. NO. 105-190, at 46 (1998) (emphasis added); H.R. REP. NO. 105-551, pt. 2, at 55 (1998) (referring to the URL address of the location (web page) alleged to contain the infringing material as an example of sufficient information to constitute an effective takedown request).

¹²⁷ These are URIs with google.com as the host domain but with the search terms embedded as queries to the URIs.

¹²⁸ *Giganews*, 993 F. Supp. 2d at 1201.

¹²⁹ In *Perfect 10, Inc. v. Cybernet Ventures, Inc.*, the court seemed to suggest that “a representative list” of infringing works will suffice as the location information for the takedown request. 213 F. Supp. 2d 1146, 1179–80 (C.D. Cal. 2002). A close reading however of this dictum suggests that the court has conflated the identification requirement for the infringed copyrighted work in 17 U.S.C. § 512(c)(3)(A)(ii), which permits the reporter to provide a representative list, rather than an exhaustive list, of copyrighted works infringed, with the location information for the

reporter's perspective, such a broad demand will ask too much from the service provider, and if legally sanctioned as being *de rigueur*, will have the effect of substituting the service provider's judgment for the owner/reporter's. It ought *not* to be forgotten that the "goal of this provision is to provide the service provider with *adequate* information to find and address the allegedly infringing material *expeditiously*."¹³⁰ If, *arguendo*, the takedown request is in the form of very specific search terms that yield links to infringing materials, the reporter cannot give the assurance that *all* the links may be to infringing materials.¹³¹ Not only does the reporter not have control over the search results based on the search terms, but to equate the takedown request with the search terms is to force the service provider to undertake self-censorship by invalidating, now and prospectively, entire websites on the basis of these search terms in order to stay within the contours of the safe harbor immunities. After all, the results that are returned via search terms are in a "constant state of flux. . . . [T]here is no certainty that any particular search will yield the exact same results at different times."¹³² Such an approach is only one step removed from the much criticized provision in SOPA, which empowers entire sites to be disabled because a few resources on that site are infringing. Thus, it is refreshing to see from the empirical evidence that these sixty-nine requests represent an aberration, and that this is unequivocally

infringing work in 17 U.S.C. § 512(c)(3)(A)(iii), which does not spell out this relaxed requirement as a substitute for the URI.

¹³⁰ S. REP. NO. 105-190, at 46 (1998) (emphasis added); *see also* *Viacom Int'l Inc. v. YouTube, Inc.*, 940 F. Supp. 2d 110, 115 (S.D.N.Y. Apr. 18, 2013).

¹³¹ *See Giganews*, 993 F. Supp. 2d at 1200 (noting that the search term supplied yielded some Usenet messages that were non-infringing).

¹³² *Id.*

not the practice adopted by the overwhelming majority of owners and reporters as regards takedown requests.

Hence, while a URI is not *sufficient* to constitute a takedown request, it remains a *necessary* component of a takedown request. To conduct this analysis of takedown requests, a census of Google DMCA form notices between 2008 and 2012 on the second dataset ($N_2=501,286$) was undertaken and the requests in the form of URIs were extracted for this case study.¹³³ Extracting all the requests and sorting their yearly totals by the copyright owners' names and the industries they represent yielded the breakdown in [Table 7](#).¹³⁴

The number of requests set out in [Table 7](#) by content provider/owner paints a rather extraordinary picture of the sheer magnitude of the operation of the DMCA takedown mechanism, more so than the breakdown of notices that is portrayed in [Table 2](#). While the top twenty content providers/owners issued an almost insignificant number of takedown requests between 2008 and 2010, the same providers/owners issued on average 93,596 takedown requests in 2011. In 2012, this figure rose to an average of 2,050,211 takedown requests. In fact, the bulk of the requests came in 2012, when these content providers/owners issued a total of 41,004,223 requests, representing a twenty-fold increase over 2011.

¹³³ As previously noted, the second dataset for the case study excluded Twitter notices because their submissions to the Chilling Effects repository were incomplete at the time of the study. *See supra* text accompanying note 43.

¹³⁴ *See infra* Part VIII, [Table 7](#), available at http://vjolt.net/vol18/issue3/seng/Seng_Table07.pdf.

In 2012, the top provider/owner by total requests is the RIAA at 7.7 million requests, with Froytal placing a close second at 7.0 million requests. BPI, the top provider by total takedown notices, is placed fourth by total requests (4.9 million requests in 2012). By and large, a top provider by total notices is likely to also be a top provider by total requests (nine of the top twenty providers by notices (marked by an asterisk in Table 7) are also top providers by requests, and all but two of the top twenty providers by requests are also top-fifty providers by notices). A notable exception is IFPI, which was placed fifth in Table 2 by total takedown notices. However, it is actually placed in 131st position by total takedown requests (17,143 requests). Conversely, Adobe is not even a top-hundred provider by total notices, as it only issued ninety-two notices in 2012, but it is actually placed twentieth by total requests made (437,252 requests). Therefore, if one seeks to map the scale of the takedown activity by a provider, the number of requests made would, in some respects, be a more accurate measure.

In this regard, when aggregated by number of takedown requests from the top-fifty providers, as shown in Figure 11,¹³⁵ the music industry, which accounts for 58.6 percent of total takedown notices,¹³⁶ still accounts for the largest share of total takedown requests (32.1 percent), while the adult entertainment industry remains in second place (30.1 percent). The movie industry, placed third by total number of notices issued (9.4 percent), is placed fourth at 13.4 percent by way of total takedown requests, while games/software rises to third at 15.2 percent. Broadcasts, websites, and books weigh in at 7.4 percent, 1.3 percent, and 0.5 percent respectively. Therefore, it seems that the adult entertainment industry is as aggressive as

¹³⁵ See *infra* Part IX, Figure 11, available at http://vjolt.net/vol18/issue3/seng/Seng_Figure11.pdf.

¹³⁶ See *infra* Part IX, Figure 5, available at http://vjolt.net/vol18/issue3/seng/Seng_Figure05.pdf.

the music industry in policing its digital content on the Internet, although the requests breakdown demonstrates that there is also aggressive enforcement activity by the movie and the games/software industries.

What relationship, if any, exists between the notices, claims, and requests? A longitudinal analysis of the statistics on the notices, claims, and requests of this second dataset yields the data included in [Table 8](#).¹³⁷

So across the years, the total number of notices, claims, and requests have each increased. In fact, there appears to be an exponential correlation between the number of claims in each notice and the total number of takedown requests per notice. The larger the number of claims in a notice, the larger will be the total number of takedown requests. From only a maximum of one claim in each notice in 2010, the figure rose to a maximum of 997 claims per notice in 2011 and 1484 claims per notice in 2011 and 2012, respectively. Likewise, from a maximum of 9515 takedown requests per notice in 2011, this number rose to a maximum of 25,050 requests per notice in 2012.

However, despite the huge numbers of total claims and takedown requests issued by the senders, such as the “mega” notice with 25,050 requests in 2012, most notices only average around 5 claims per notice and around 125 requests per notice. But even the requests figures are inflated, as explained by the standard deviation and median numbers. At 2.00 and 4.00 median requests per notice in 2011 and 2012, respectively, this demonstrates that as a whole, notwithstanding the issuance of many takedown notices with large numbers of requests per

¹³⁷ See *infra* Part VIII, Table 8, available at http://vjolt.net/vol18/issue3/seng/Seng_Table08.pdf.

notice, there also are many notices with small numbers of claims and takedown requests.

Who is responsible for issuing these small “claims/requests” notices? Could the reporting agents be responsible? What about the individual copyright owners? To investigate this phenomenon further, Table 9¹³⁸ breaks down the notices, claims, and requests of the top thirty reporting agents by total requests in 2012.

The analysis begins by noting that many of the top reporters in Table 4 by volume of notices like BPI, Degban, AudioLock, Irdeto, Muso, and Publishers Association are also top reporters by requests in this table. Since the second dataset ($N_2=501,286$) is a subset of the first dataset ($N_1=539,558$), if the total notice figures in the two tables are the same or very close, this confirms that many of the same reporters have targeted Google by way of their web form-based takedown notices and that some have done so exclusively. Together, these thirty reporters are responsible for fifty-three million takedown requests, or 97.8 percent of all takedown requests in 2012. This list of reporters includes content providers like 20th Century/Fox, NBCUniversal, Warner Brothers, and Microsoft in addition to CMOs like RIAA, BPI, BAF, Publishers Association, and IFPI. But like Table 4, reporting agents like Degban, Takedown Piracy, DtecNet, and Marketly make up the majority on the list. In fact, of the top thirty reporters, nineteen of them are reporting agents.

Additional insights into how these reporters work can be found by sorting them based on their average number of claims and requests per notice. Based on the claims per notice

¹³⁸ See *infra* Part VIII, Table 9, available at http://vjolt.net/vol18/issue3/seng/Seng_Table09.pdf.

figures, the top thirty reporters make, on average, 19.13 claims in each notice. However, RIAA's claims per notice figure (439.53 claims per notice) is definitely an outlier. Removing this from the figures, the average claims per notice figure for the remaining twenty-nine reporters drops to 4.64 claims per notice. Table 9 clearly shows that, unlike other reporters, RIAA is packing thousands of copyright claims and takedown requests into each takedown notice.¹³⁹ Even though RIAA issued only 2211 notices in 2012, by this strategy, it has issued 7,632,938 takedown requests against Google in 2012 and is placed second by way of total takedown requests. In contrast, Degban averages only 1.59 claims per notice but must issue 77,869 notices to accumulate a total of 11,741,358 requests in 2012.

Degban's approach of issuing more notices with fewer claims but averaging several hundred requests per notice is more in line with the practice of most reporters. In this regard, it is notable that Web Sheriff, DMCA Force, Warner Bros, and Unidam only make *one* claim in each takedown notice. But this has not prevented them from issuing more than 900,000 takedown requests in total. All in, twenty of the top thirty reporters—who only average 2.34 claims per notice—never make more than ten claims in each takedown notice. In other words, these twenty reporters pack two hundred times *fewer* claims into each of their notices than the RIAA.

A different perspective presents itself according to the terms of requests per notice figures. On this, the first fifteen reporters issue an average of 1652.08 requests per notice, but the other fifteen reporters issue an average of 115.20 requests

¹³⁹ RIAA's average requests per notice ratio is the highest among the thirty reporters at 3452.26

per notice or fourteen times less. The issuing of fewer requests per notice appears to translate into a decision to issue more takedown notices. Thus, the first fifteen reporters average 1815 notices, but the next fifteen reporters average 22,182 notices, or twelve times more notices.

While one may initially believe that reporting agents issue more takedown requests than content providers, data from Table 9 show otherwise. The four content providers/owners who are not CMOs only issued, on average, 846 notices but averaged 1141.81 requests per notice. In contrast, the six CMOs averaged 33,421 notices and 1235.28 requests per notice, while the nineteen reporting agents averaged 7706 notices but only 661.84 requests per notice. In other words, among the top thirty reporters, providers on average issued about one-tenth as many notices and about 50 percent more takedown requests than agents, suggesting that providers can be at least as proactive and aggressive, if not more so, than agents. Of the agents, MarkMonitor stands out for issuing the smallest number of notices (only eighty in 2012) with consistently some of the largest number of takedown requests per notice, as evident in its median, first, and third quartile requests per notice figures.

That is not to say that CMOs are not proactive and aggressive in issuing takedown notices and requests. The statistics above show at least two different strategies adopted by, say, RIAA, BAF, and Stichting BREIN on the one hand and, say, BPI, Publishers Association, and IFPI on the other. As can be seen from the median, first, and third quartile requests per notice figures, the former group of CMOs seem to favor the relatively infrequent use of notices with large numbers of takedown requests (their median requests/notice figures are 2743, 1733, and 776, respectively), whereas the latter seem to favor the more frequent use of notices with

smaller numbers of takedown requests (their median requests per notice figures are 3, 23, and 12 respectively). This demonstrates there are corporate reporters and reporting agents which have been issuing large numbers of notices with a small number of takedown requests that cumulatively add up to large numbers of requests.

On the other hand, issuing notices with large numbers of requests is no longer the exclusive province of corporate reporters and reporting agents. On the assumption that notices with redacted reporter names are individuals, individuals are collectively ranked twenty-sixth in Table 9 and have issued 9425 takedown notices. Unsurprisingly, they make only an average of 1.67 claims in each notice, but these accounted for 120,000 takedown requests in 2012. As a group, they also enforce very regularly, issuing notices every 1.5 days.¹⁴⁰ Given that approximately the same aggregate of individuals are also ranked fourth by total notices issued (see Table 4),¹⁴¹ this strongly suggests that individual providers or individual copyright owners are actively enforcing their rights. This is because it is presumably relatively easy for them to detect instances of infringement and because the mechanism for issuing takedown notices and requests to service providers is both accessible and easy to use.

¹⁴⁰ The accuracy of this estimate is based on the assumption that reporters with redacted names are individuals. Short of accessing the original unredacted notices, there is no sure way to accurately verify if a notice and its requests are made by reporters as individuals.

¹⁴¹ In Table 4, individuals are identified as such by the metadata supplied by Chilling Effects. In Table 9, they are identified by the information parsed from the notice contents, hence the slight discrepancies in total numbers of notices.

What is also salient is the regularity of the enforcement, as measured by the various statistics set out in the two rightmost columns in Table 9. In this regard, the reporters demonstrate a wide range of enforcement patterns. Top reporters like BPI, Degban, Takedown Piracy, and DtecNet issue takedown notices and requests almost every day in 2012, with BPI setting the blistering pace at 360 out of 366 days in 2012 and issuing at least one notice every 1.02 days. As a provider/CMO, BPI clearly matches and even exceeds the pace set by the top reporting agents like DtecNet and Degban at 1.18 and 1.20 days per notice, respectively. In contrast, some reporters like Armovore, MarkMonitor, Microsoft, and BAF issue notices and requests on fewer than ninety days of the year.

In sum, there is wide variation in the strategies adopted by providers, CMOs, and reporting organizations in the issuance of takedown requests. In spite of the different strategies adopted, these reporters are responsible for issuing a huge number of notices, claims, and requests, as shown by an analysis of the second dataset of Google takedown notices. But individuals as a whole also issue a number of notices, claims, and requests that is not insignificant. The critical issue becomes ensuring that all these reporters, regardless of their different strategies and backgrounds, receive equal treatment, including priority of processing, speed of response,¹⁴² and level of

¹⁴² GOOGLE, HOW GOOGLE FIGHTS PIRACY 14 (2013), available at <https://docs.google.com/file/d/0BwxyRPFduTN2dVFqYml5UENUeUE/edit> [hereinafter HOW GOOGLE FIGHTS PIRACY] (claiming that Google's average turnaround time for takedown notices is less than six hours, but it is unclear from report whether unsophisticated reporters' notices take more time or are lower in priority than other reporters).

scrutiny for their takedown notices and requests, despite the huge numbers of notices and requests involved.

E. Trusted Reporters and their Takedown Requests

The sheer volume of notices and takedown requests associated with Google, which are issued by a small number of content providers and reporting agents, has led Google to institute the Trusted Copyright Removal Program (TCRP). Providers and agents participate in the TCRP on the basis that they are “reliable, high accuracy submitters” whose notices Google did not want to delay the processing, as compared to “nonsophisticated submitters” who issue a lot of “incomplete or abusive” notices.¹⁴³

In contrast to Google’s avowed philosophy of transparency in takedown notices, the exact details of this program are shrouded in relative secrecy.¹⁴⁴ The TCRP appears to be an automated method for providers and agents to submit large numbers of notices and takedown requests electronically to Google, which Google processes rapidly via an automated process. But no TCRP submitter has publicly released information on how it checks the validity of the automated takedown requests.¹⁴⁵ Nor has Google released information on how it checks the validity of its system for processing these

¹⁴³ Rebecca Tushnet, *PTO/NTIA: Notice and Takedown*, REBECCA TUSHNET’S 43(B)LOG (Dec. 12, 2013),

<http://tushnet.blogspot.com/2013/12/ptontia-notice-and-takedown.html>.

¹⁴⁴ Mark Leiser, *The Copyright Issue and Censorship Threat Buried Within Google’s Transparency Report*, THE DRUM (Dec. 23, 2013, 7:30 AM), <http://www.thedrum.com/news/2013/12/23/copyright-issue-and-censorship-threat-buried-within-googles-transparency-report>.

¹⁴⁵ *Id.*

takedown requests, though it has hinted that these are built around machine-learning algorithms.¹⁴⁶

Information about who the TCRP participants are is also not in the public domain. However, it is possible to reconstitute the list of TCRP participants from their electronic “Trusted User” signatures in their notices as captured in the Chilling Effects repository. Based on this reconstituted information, the list of participants and the total number of notices and requests they have submitted as Trusted Users are set out below. The date of their first issued notice as a Trusted User, the date of their last issued notice for 2012, and their “takedown rate”¹⁴⁷ are also listed in [Table 10](#).¹⁴⁸

From Table 10, we can surmise that the TCRP began sometime in March 2011 and comprises about fifty participants. By the end of 2012, it was responsible for the issuance of 376,000 notices and fifty-four million takedown requests.¹⁴⁹ Participants include content providers like Fox and NBC, technology companies like Microsoft, CMOs like RIAA, BPI, and IFPI, and reporting agents. So if there is arguably a spike in online enforcement of takedown notices post-SOPA, the TCRP is the engine that makes it possible.

¹⁴⁶ HOW GOOGLE FIGHTS PIRACY, *supra* 142, at 17.

¹⁴⁷ The takedown rate is measured by the ratio between a count of the URIs that are successfully removed and the total URIs that are *processed* by Google as part of its web search takedown notices. The total number of URIs processed is different from the total URIs *submitted*, because the former is extracted from the GTR but the latter is extracted from the Chilling Effects repository. Many of the notices listed in the repository are not listed in the GTR raw data.

¹⁴⁸ See *infra* Part VIII, Table 10, available at http://vjolt.net/vol18/issue3/seng/Seng_Table10.pdf.

¹⁴⁹ In 2012 alone, TCRP participants issued 346,548 notices and 49,652,888 takedown requests.

To prevent runaway programs from overwhelming their systems, Google has placed daily caps on the maximum number of requests that can be submitted by each participant under the TCRP.¹⁵⁰ The TCRP appears to operate on a tiered system. Going by the maximum number of requests per notice for each participant, some participants (e.g., Warner Brothers, Unidam and Contributor) have a cap of 1000 requests per notice placed on them, while for others, the cap is 10,000 requests (e.g., Takedown Piracy, Fox Group Legal, and Remove Your Media). In the highest tier are participants like Degban, RIAA, Morganelli, BPI, and NBCUniversal, which can issue more than 10,000 takedown requests in each notice. Some content providers and reporting agents have complained about this 10,000 request limit, which appears to be not only a notice limit but also a daily limit.¹⁵¹ However, in April 2013, after the data for this study was finalized, it was reported that Google consented to lifting this limit.¹⁵² It remains to be seen whether Google's lifting of the limit applies across the board to all reporters or only to the top reporters. Given that some reporters already issue more than 10,000 requests in each notice, a complete lifting of the limit would be welcomed by providers and reporters alike, provided, of course, that notices and requests submitted via the TCRP are reliable.

Just how trustworthy is the TCRP? Google has to date not published any figures. But according to the information in

¹⁵⁰ Ernesto, *Google Relaxes DMCA Takedown Restrictions, Eyes Abuse*, TORRENTFREAK (Apr. 11, 2013), <http://torrentfreak.com/google-relaxes-dmca-takedown-restrictions-eyes-abuse-130411/>.

¹⁵¹ Ernesto, *Anti-Piracy Groups Want Google to Lift DMCA Takedown Cap*, TORRENTFREAK (Feb. 19, 2013), <http://torrentfreak.com/anti-piracy-groups-want-google-to-lift-dmca-takedown-cap-130219/>.

¹⁵² Ernesto, *supra* note 150.

the second dataset (which, as noted above, is drawn from and incorporates Google's own GTR raw data), it is possible to calculate the "takedown rate" of each TCRP participant. The last column of the above table shows that the computed average success rate of takedown requests is 96.2 percent, excluding the much lower figures for Vobile and BayTSP as outliers. This is a high success rate; TCRP participants are to be commended for their efforts. However, this takedown rate is lower than the general rate of 97.5 percent publicly noted by Google.¹⁵³ In contrast, drawing from the same dataset, individual reporters (identified as "redacted" in the reporter field) have a collective takedown rate of 85.2 percent, which, while lower than the average takedown rate of the TCRP participants, is at least the same as or higher than the takedown rate of three TCRP participants in Table 10: Degban, Muso, and Peer Media Technologies.

In addition, as part of its efforts to uphold the trustworthiness of the TCRP, Google referred to its 2012 disbarring of two participants from the TCRP for their "repeated failure to submit accurate notices."¹⁵⁴ While it did not identify these participants, from examining Table 10, those disbarred could be the reporting agents Armovore and LeakID, whose last Trusted User submissions were in February and August 2012, respectively. However, the takedown rates of Armovore (97.3 percent) and LeakID (99.4 percent) are in line with and actually higher than the average failure rates of the TCRP participants (3.8 percent).

This may prompt questions as to whether the TCRP has promoted a higher quality of takedown submissions. On the

¹⁵³ HOW GOOGLE FIGHTS PIRACY, *supra* note 142, at 14.

¹⁵⁴ Tushnet, *supra* note 143.

one hand, placing caps on the number of takedown requests allowed by each TCRP participant may force them to be selective and first target egregious piracy sites with their takedown requests. The converse question is whether the TCRP's large-scale automated submission and takedown scheme has enabled inaccurate notices and encouraged sloppy takedown requests. In line with Google's moves toward greater transparency in its takedown process (and indeed, its release of the GTR raw data makes this analysis possible), perhaps more transparency as to the operational details of the TCRP will help to answer some of these vexing questions.

IV. GOOGLE SERVICES' TAKEDOWN NOTICES AND REQUESTS BY DMCA SECTION

A. Notices and Requests for Google Services

Initially, a key objective of this study was to classify the Chilling Effects repository notices by their DMCA categories, as this could provide a solid empirical indication of the use and reliance on each DMCA safe harbor.¹⁵⁵ However, the utility of this exercise is limited. Since notice submissions to the Chilling Effects repository are voluntary, the categories of DMCA notices are greatly affected: first, by the selection bias of the providers as contributors, and second, by the reporters' choice of the service providers' departments to address the notices. As it turns out, the repository is dominated by Google and Twitter notices. And as these providers primarily offer services built around information location tools (Google by way of its Web Search services and Twitter by way

¹⁵⁵ The same classification exercise was done under the Urban and Quilter study. *See* Urban & Quilter, *supra* note 21, at 642.

of links to resources embedded in its users' tweets),¹⁵⁶ any classification will show a predominance of § 512(d) notices. This holds true for the second dataset (N₂=501,286). Nonetheless, this part of the study attempts to quantify the notices and requests received by various Google services as a measure of the enforcement activity against each service and also to permit a better understanding of the relationship between the different Google services and their takedown notices. Twitter notices were not classified for this part of the exercise because they were excluded from the second dataset.¹⁵⁷

Assigning DMCA section numbers based on the nature of the services provided,¹⁵⁸ [Table 11](#)¹⁵⁹ sets out the notices and requests statistics for 2011 and 2012.¹⁶⁰

¹⁵⁶ There is a difficult question as to whether Twitter's 140 character service is better classified as an "information location tool," or whether it is a hosting service for pithy content, or if it is a hybrid of both services (e.g., where hashtags are used for posts). Twitter itself seems to recognize this in its online DMCA forms for submitting takedown notices. My thanks to Professor Urban for sharing this observation with me.

¹⁵⁷ See *supra* text accompanying note 43.

¹⁵⁸ Google's ad-based services, such as AdSense, AdWords, and DoubleClick, do not fall neatly into the existing DMCA sections, since they involve neither hosting of third party content nor linking to infringing material. They are arguably *sui generis* and will be classified as such in this table. As for services which involve some hybrid hosting and referral activity where the hosted file can be located and downloaded, such as Picasa and Groups, the service will be classified based on its predominant activity (and the secondary activity will not be held to preclude the service provider from the predominant safe harbor category). In this case, Picasa is classified as involving the provision of hosting services. See *UMG Recordings, Inc. v. Shelter Capital Partners LLC*, 718 F.3d 1006, 1018–19 (9th Cir. 2013); *Capitol Records, LLC v. Vimeo, LLC*, Nos. 09 Civ. 10101(RA), 09 Civ. 10105(RA), 2013 WL 5272932, at *16 (S.D.N.Y. Sept.

It is evident that notices and requests directed to Google's Search services dominate. In 2011, Google's Search services team processed 31,441 notices comprising 2.6 million requests across ten search services (54.4 percent of all notices and 97.4 percent of all requests). In 2012, this number rose to 411,721 notices comprising fifty-four million requests (94.4 percent of all notices and 99.9 percent of all requests). In comparison, Google's twenty-seven hosting services, which include Blogger/Blogspot, Cloud, Play, Plus, and Sites, processed 26,041 notices comprising 70,333 requests in 2011 (45.1 percent of all notices and 2.6 percent of all requests). In 2012, absolute numbers remained about the same at 24,115 notices comprising 72,085 requests, but relative to the Search services, they fell to 5.5 percent of all notices and made up only 0.1 percent of all requests. So by absolute numbers, the level of enforcement for Google's hosting services remained about the same, but the level of enforcement for Google's Search services definitely escalated.

As an aside, there are a small number of email notices targeting Google Wifi services.¹⁶¹ Because technically, there is

18, 2013). Google Groups is also classified as the provision of hosting services, where Usenet servers exchange newsfeeds (Usenet postings from their users) with each other. *See* Perfect 10, Inc. v. Giganews, Inc., 993 F. Supp. 2d 1192, 1198–1202 (C.D. Cal. 2014) (holding that Perfect 10 had not established that its notices comply with § 512(c) against Giganews, a Usenet service provider).

¹⁵⁹ *See infra* Part VIII, Table 11, available at http://vjolt.net/vol18/issue3/seng/Seng_Table11.pdf.

¹⁶⁰ As far as possible, these figures exclude counter notices, whose figures are set out in Part IV.B.

¹⁶¹ *See Removing Content from Google*, GOOGLE, <https://support.google.com/legal/troubleshooter/1114905?rd=2> (last visited Sept. 17, 2014). Many of these notices appear to relate to users who use Google Wifi services to share torrents and download files. *See, e.g.,*

no takedown mechanism for § 512(a) notices under the DMCA,¹⁶² no further reference will be made to them in this study.

Likewise, no web form notices directed to Google's caching services were detected in this study, even though there is a notification for takedown mechanism in § 512(b) of the DMCA.¹⁶³ Perhaps this is related to the caselaw in Google's favor on this matter.¹⁶⁴ But a better explanation is that once the link to an unauthorized online resource, which is incidentally cached, is successfully removed by way of a § 512(d) takedown request, the cached copy is no longer accessible.¹⁶⁵ And if the complaint is with respect to Google's caching of an owner-authorized resource which the owner wants removed, if the owner adjusts the cache settings for that resource or prevents Google's robots from accessing that resource,¹⁶⁶

Download DMCA (Copyright) Complaint to Google, CHILLING EFFECTS (Jan. 2, 2013), <https://www.chillingeffects.org/notices/783295>.

¹⁶² See Urban & Quilter, *supra* note 21, at 649 (discussing the absence of any takedown mechanism under DMCA § 512(a) for online service providers acting as conduits because unlike the other safe harbor provisions, the § 512(a) safe harbor is unqualified, i.e., it is not contingent on responding to a takedown notice).

¹⁶³ 17 U.S.C. § 512(b)(2)(E) (2012).

¹⁶⁴ See, e.g., *Field v. Google, Inc.*, 412 F. Supp. 2d 1106, 1109 (D. Nev. 2006) (granting summary judgment to Google in copyright infringement suit over Google's cached copies of plaintiff's copyrighted material).

¹⁶⁵ See *Perfect 10, Inc. v. Google, Inc.*, No. CV 04-9484 AHM (SHx), 2010 WL 9479059, at *13 (C.D. Cal. July 26, 2010) (citing Google's declaration that it is its practice to automatically remove a cached link when it suppresses the corresponding web page URL).

¹⁶⁶ This is based on the Robots Exclusion Standard. See *Robots.txt Specifications*, GOOGLE (Aug. 2, 2012), https://developers.google.com/webmasters/control-crawl-index/docs/robots_txt.

industry-standard caching protocols¹⁶⁷ require Google to drop that resource from its own cache in order to comply with § 512(b).¹⁶⁸ Given the frequency with which Google's robots access websites and their pervasiveness,¹⁶⁹ this will happen quickly. For these practical reasons, there is little room for the use of a § 512(b) takedown notice to remove cached copies of unauthorized resources, particularly where it is preceded by a § 512(c) or § 512(d) notice.

A very different story exists for § 512(d) takedown notices. The trend of increasing numbers of § 512(d) notices was first noted in the 2006 Urban and Quilter study.¹⁷⁰ But it still remains surprising to see how Google's general Search notices have come to dominate even notices directed at other search services. This is also reflected in the high requests per notice ratio for Google's general Search (132.53 requests per notice), compared with that for Imagesearch (4.64 requests per notice) in 2012. Likewise, in 2012, Picasa, Plus and Blogger/Blogspot as non-search services have 7.09, 3.43, and 3.02 requests per notice and received 5222, 3213, and 58,052 requests, respectively. Together with general Search and Imagesearch notices, they make up the top five services by takedown requests. Thus, it can be argued that services with strong search features better lend themselves to copyright

¹⁶⁷ This is the HTTP/1.1 protocol for caching. See R. Fielding et al., *Hypertext Transfer Protocol: Caching in HTTP* (Network Working Group, Request for Comments: 2616, 1999), available at <http://www.w3.org/Protocols/rfc2616/rfc2616-sec13.html>.

¹⁶⁸ 17 U.S.C. § 512(b)(2)(B).

¹⁶⁹ *Googlebot*, GOOGLE, <https://support.google.com/webmasters/answer/182072> (last visited Sept. 17, 2014) (describing how Google's robot, Googlebot, accesses "your site [no] more than once every few seconds on average").

¹⁷⁰ See Urban & Quilter, *supra* note 21, at 647.

enforcement, in that content providers and reporters will use the very same search features alleged to facilitate access to infringing material to detect and take down or disable access to that infringing material.¹⁷¹

Google's Blogger service is notably number two, at a certainly substantial 19,220 § 512(c) notices with 58,052 requests in 2012. What is surprising is how heavily targeted Blogger/Blogspot is by way of takedown notices. Even though it is an operational division in Google, analysis of the dataset shows that Blogger still receives more takedown notices each year than Twitter, Yahoo!, and Digg combined. In fact, in 2009 and 2010, Google Blogger/Blogspot received more takedown notices than Google Search. One possible reason why Blogger/Blogspot notices and requests have remained relatively constant between 2011 and 2012 while general Search notices have exploded is that providers and reporters are using Google Search as a first port of call to locate allegedly infringing material, including materials hosted on Blogger/Blogspot, thereby inflating the Google Search results.¹⁷² Furthermore, Google has confirmed that if Google Search results point to infringing materials on Google's own

¹⁷¹ Even though there are only a small number of YouTube notices, this hosting service with strong search features also has a relatively high request per notice takedown ratio (17.95 in 2012) relative to the other Google services. The small number of such notices is explained by the fact that, as previously noted, YouTube operates the Content ID takedown system, whose notices do not appear in the Chilling Effects repository. *See supra* text accompanying note 53.

¹⁷² Analysis of the second dataset shows that there were 6458 general Search requests targeting materials hosted on Blogger and Blogspot in 2011, and in 2012, this figure rose to 36,030 requests. And the takedown rate of these requests is 97.7 percent and 97.6 percent in 2011 and 2012 respectively.

hosting services such as Blogger, these materials are also removed together with links to these materials.¹⁷³ Without these measures, the number of *direct* Blogger takedown notices and requests would arguably be higher.

Google's Play service is also quite heavily targeted, receiving 2054 § 512(c) notices and 3097 requests in 2012. This is surprising at first glance, because Play is a distribution platform for licensed music, books, movies, TV programs, and applications. A quick examination suggests that a large proportion of these notices and requests are related to complaints about applications and games on Google Play/Android. This could be because of the relatively open ecosystem for Android applications, with no or minimal preapproval required for submissions, thereby encouraging a "submit first, takedown later" approach.¹⁷⁴ No doubt the relatively transparent access which developers have to track and download each other's applications also facilitates the reasonably active self-policing by the developers themselves.

It is also interesting to see how Google has created a special category of takedown notices for its advertising services—AdWords and AdSense. Though technically not DMCA notices, Google has modeled a special class of notices for them on the basis of the DMCA takedown mechanism. AdWords takedown notices enable complaints to request the removal of paid advertising paired to certain keywords that are

¹⁷³ *Visible Changes: Google Transparency Report*, GOOGLE, <http://www.google.com/transparencyreport/removals/copyright/changes> (last visited Sept. 17, 2014).

¹⁷⁴ See Robin Wauters, *37% of Published Android Apps were Later Removed, Compared to 24% of iOS Apps*, TECHCRUNCH (Oct. 21, 2011), <http://techcrunch.com/2011/10/21/37-of-published-android-apps-were-later-removed-compared-to-24-of-ios-apps>.

shown alongside search results that support infringing activities.¹⁷⁵ And AdSense takedown notices enable complaints to request the removal of AdSense advertising shown on sites with unauthorized copyrighted material.¹⁷⁶ It is presumed that by putting in place a mechanism for disabling its monetizing services, Google could better manage its indirect liability risks¹⁷⁷ if subscribers are found to have paid for infringing advertising for AdWords or third party sites are found to be hosting infringing or illicit material or conducting infringing activity for AdSense. As reported to the Chilling Effects repository, their numbers are currently small, as they currently make up less than 0.05 percent of all notices and less than

¹⁷⁵ *AdWords: Copyright—What's the Policy?*, GOOGLE, https://support.google.com/adwordspolicy/answer/176015?hl=en&ref_topic=1346941 (last visited Sept. 14, 2014) (allowing AdWords for sites that promote video trailers, audio guides, e-books, ringtones, anime, and works of independent artists and record labels and otherwise disallowing AdWords for sites that make unlicensed use of copyrighted content, although these seem to be a description of Google's policy for AdSense); *Why Was My Ad Disapproved for Legal Reasons?*, GOOGLE, <https://support.google.com/adwords/answer/1209109?hl=en> (last visited Sept. 17, 2014).

¹⁷⁶ *Content Policies: Copyright Infringement*, GOOGLE, https://support.google.com/adsense/answer/9892?hl=en&ref_topic=1271507 (last visited Sept. 20, 2014); *Policy Tips: Avoiding Copyright Infringement*, GOOGLE (Aug. 27, 2012), <http://adsense.blogspot.com/2012/08/policy-tips-avoiding-copyright.html>.

¹⁷⁷ See *The Six Business Models for Copyright Infringement: A Data-Driven Study of Websites Considered to Be Infringing Copyright*, PRS FOR MUSIC (June 27, 2012), <http://www.prsformusic.com/aboutus/policyandresearch/researchandconomics/Documents/TheSixBusinessModelsofCopyrightInfringement.pdf>; *USC Annenberg Lab Ad Transparency Report*, USC ANNEBERG INNOVATION LAB (Jan. 5, 2013), http://www.annenberglab.com/sites/default/files/uploads/USCAnnenbergLab_AdReport_Jan2013.pdf.

0.0001 percent of all requests in 2012. But there is no doubting the importance of this class of notices to Google and to content providers and their reporters— after all, over 90 percent of Google’s revenue comes from advertising via its AdSense and AdWords programs.¹⁷⁸

B. Counter Notices

The DMCA also provides that service providers like Google who receive § 512(c) takedown notices¹⁷⁹ notify their “subscribers”¹⁸⁰ of the removal or disabling of access to the subscribers’ material. This protects subscribers by giving them an opportunity to contest the original takedown notices on grounds of mistake or misidentification of the material as infringing.¹⁸¹ If a subscriber provides a valid “counter notice,” the service provider must promptly notify the reporter of the subscriber’s objections.¹⁸² If between 10 and 14 business days after receiving the counter notice, the reporter has not sought a court order to restrain the subscriber with regard to the material in question, the service provider is legally required to place the

¹⁷⁸ U.S. SEC. AND EXCH. COMM’N, FORM 10-Q: GOOGLE 32 (2013), available at <http://www.sec.gov/Archives/edgar/data/1288776/000128877613000068/google10-qq32013.htm> (showing Google advertising revenues for nine months ended on September 30, 2013 for Google, excluding revenues from the Motorola division).

¹⁷⁹ 17 U.S.C. § 512(g)(2) (2012).

¹⁸⁰ A “subscriber” is a user of hosting services offered by a service provider. See, e.g., *id.* §§ 512(g)(3)(C), (j)(1)(A)(ii), (j)(1)(B)(i).

¹⁸¹ *Id.* § 512(g)(3)(C).

¹⁸² *Id.* § 512(g)(2)(B).

subscriber's material back online or cease disabling access to it ("put back").¹⁸³

The "put back" mechanism by way of a counter notice is an important balancing mechanism built into the DMCA to prevent the abuse of takedown notices.¹⁸⁴ On its own, it is an important subject of study because it offers insights into how subscribers have responded to takedown notices, and evidence as to the overall health of the takedown notice mechanism. In 2006, the Urban and Quilter study looked into this and only found seven counter notices in the Chilling Effects repository for all notices up to 2006.¹⁸⁵ Since then, there has been little written about counter notices, because anecdotally, their numbers have remained very small.

Not surprisingly, the metafields provided on the Chilling Effects repository are of little assistance. To detect counter notices in the dataset, the parsers had to be configured to look for keywords and key fields within each notice. Some of these keywords and fields are specific to the varying internal formats which Google uses for handling counter notices. Since no counter notices were detected between 2008 and 2010, the information in [Table 12](#)¹⁸⁶ contains the number of counter notices detected in the Google form notices between 2011 and 2012.

These small numbers of counter notices come as a surprise, relative to the large numbers of notices served in 2011

¹⁸³ *Id.* § 512(g)(2)(C).

¹⁸⁴ See *Lenz v. Universal Music Corp.*, 572 F. Supp. 2d 1150, 1156 (N.D. Cal. 2008) (explaining that § 512(f) with § 512(g)(2) are mechanisms to prevent the abuse of takedown notices).

¹⁸⁵ See Urban & Quilter, *supra* note 21, at 679.

¹⁸⁶ See *infra* Part VIII, Table 12, available at http://vjolt.net/vol18/issue3/seng/Seng_Table12.pdf.

and 2012. Of the sixty-eight notices in 2011, fifty-nine of them relate to a § 512(c) takedown notice, and eight of them relate to a § 512(d) notice. Likewise, in 2012, eighty-two of the counter notices relate to a § 512(c) takedown notice, and only two relate to a § 512(d) notice.

There is at least one legal explanation for the relatively larger numbers of counter notices stemming from a § 512(c) takedown notice rather than a § 512(d) takedown notice. Given that a service provider like Google has a much closer relationship with a subscriber who chooses to host her content with Google, such a subscriber could be readily identified for notification purposes. On the other hand, if Google disables a link that otherwise points to content hosted by a third party, it is not clear if there is any legal obligation for an information referral service provider to inform the affected third party. The counter-notification provision, § 512(g), refers to notifying a “*subscriber of the service provider*” whose “*material [is] residing at the direction of [the] subscriber . . . on a system or network controlled or operated by or for the service provider . . . , or to which access is disabled by the service provider.*”¹⁸⁷ The language of it comports with a takedown notice issued under § 512(c) rather than § 512(d), because it would be a stretch to refer to the owner of a site whose content has been linked or indexed by the service provider as the provider’s “*subscriber.*”¹⁸⁸ In addition, the information location links

¹⁸⁷ 17 U.S.C. § 512(g)(2) (2012) (emphasis added).

¹⁸⁸ See, e.g., *Perfect 10, Inc. v. Google, Inc.*, No. CV04-9484 AHM (SHx), 2010 WL 9479059, at *4 (C.D. Cal. July 26, 2010) (noting that search and caching services do not have account holders or subscribers and so, Google was not required to implement repeat infringer policies for these services); *Urban & Quilter*, *supra* note 21, at 626 (noting that search providers likely have no service relationships with the alleged infringers).

provided by a § 512(d) service provider would not be as “material residing at the direction” of the subscriber, since these are generated by the service provider.¹⁸⁹ Nonetheless, Google has gone beyond the DMCA formal requirement and made it possible to file a § 512(d) counter notice;¹⁹⁰ the fact that at least ten aggrieved parties responded by way of a counter notice in relation to a search takedown request in 2011 and 2012 shows that these parties feel that they have a strong case.

Urban and Quilter suggested that one reason for the small numbers of counter notices issued in response to § 512(c) takedown notices is that affected parties can easily move to another service provider to host the disabled content within the ten to fourteen business-day deadline for the complainant to respond.¹⁹¹ But this option is not available for parties de-listed from Google search results pursuant to § 512(d) notices. Anecdotal evidence suggests that when sites that rely almost exclusively for search results to drive traffic to them are de-listed from Google, there is a calamitous effect on the businesses of these sites.¹⁹² This is because search traffic is

¹⁸⁹ On the other hand, it could be argued that it is the originating site which has “made available” its content for indexing and subsequent linking by the service provider (through, for instance, settings on the robots.txt file on its site), though such a site could not be easily described as the “subscriber.”

¹⁹⁰ *DMCA Counter-Notification Form*, GOOGLE, https://support.google.com/legal/contact/lr_countersnotice?product=websearch (last visited Sept. 21, 2014) (accessing the page may require the user to first log on to Google).

¹⁹¹ See Urban & Quilter, *supra* note 21, at 679–80.

¹⁹² See, e.g., Josh Constine, *Google Destroys Rap Genius’ Search Rankings as Punishment for SEO Spam, but Resolution in Progress*, TECHCRUNCH (Dec. 25, 2013), <http://techcrunch.com/2013/12/25/google-rap-genius> (noting the recent fracas relating to Rap Genius).

indigenous to the operations of any search provider. If the affected party moves to another hosting service provider, even if the search provider has rebuilt its indexes to that content, the fact it is on a new host and will not necessarily have the same incoming links may result in the provider's search engine ranking the same content more poorly than the original content.¹⁹³

The relative lack of counter notices issued by affected parties does give cause for concern. One explanation often advanced is that it would be incongruous for an infringer to challenge the takedown notice by a counter notice, disclose her identity, submit herself to U.S. jurisdiction, and expose herself to a possible legal claim.¹⁹⁴ In this regard, pundits point to the high takedown rates achieved by reporters, assert that takedown notices are highly accurate, and conclude that most takedowns are legitimate.¹⁹⁵ But it is unclear if the failure to mount a counter notice is because of subscriber intimidation by the language of takedown notices¹⁹⁶ or simply ignorance or unawareness of the possible responses to a § 512(c) notice or a § 512(d) notice. While Google has taken additional steps, such as flagging URIs as search results removed by takedown

¹⁹³ See Danny Sullivan, *What Is Google PageRank? A Guide for Searchers & Webmasters*, SEARCHENGINELAND (Apr. 26, 2007, 1:18 AM), <http://searchengineland.com/what-is-google-pagerank-a-guide-for-searchers-webmasters-11068>.

¹⁹⁴ 17 U.S.C. §§ 512(g)(2)(B), (3)(D) (2012).

¹⁹⁵ See, e.g., Declaration of David Kaplan in support of Warner Bros. Entertainment's Motion for Summary Judgment at 5, *Disney Enterprises, Inc. v. Hotfile Corp.*, (No. 11-20427) (S.D. Fla. Feb. 8, 2012).

¹⁹⁶ See Marjorie Heins & Tricia Beckles, BRENNAN CENTER FOR JUSTICE, NYU SCH. OF LAW, *WILL FAIR USE SURVIVE?: FREE EXPRESSION IN THE AGE OF COPYRIGHT CONTROL 36* (2005), available at <http://fepproject.org/policyreports/WillFairUseSurvive.pdf>.

notices or notifying site owners who have registered their sites with Google via the Webmaster program,¹⁹⁷ it is unclear how many site owners actually check the search results for their own URIs. Nor is information about the rate of subscription of the Webmaster program or the number of successful notifications sent by Google to Webmaster site owners available in the public domain. If this information is available, it will be possible to more accurately ascertain whether the lack of counter notices is because of the site owner's apathy, ignorance, or culpability.

V. PROPOSALS FOR REFORM

In the fifteen years since its enactment, the one component of the DMCA that has been subjected to the test of time considerably is the takedown notice mechanism. It is the linchpin that underlies the delicate balance of responsibilities between the copyright owner and the online service provider by placing the primary responsibility for policing the Internet for allegedly infringing activity on the owner. Takedown notice mechanisms have provided large content providers and individual copyright owners with an easily accessible, cost-effective, and expeditious process to seek the removal of infringing materials online. And it has arguably worked to reduce, though not completely eliminate, the egregious piracy of copyright content online. If the takedown mechanism is an exercise in empowerment, democratization, and cooperation, it has been a textbook triumph. If the numbers of takedown notices and requests are a measure of its relevance, it has been a sensational success.

¹⁹⁷ See HOW GOOGLE FIGHTS PIRACY, *supra* note 142, at 16.

A. Electronic Notices

As this paper has shown, however, many of the problems arise from the exponential use of takedown notices and requests. From a mere 7374 notices and 27,035 takedown requests in 2010, by the end of 2012, a total of 435,063 notices and more than fifty-four million takedown requests have been issued, as recorded in the second dataset. The lead in issuing almost six in ten notices was not undertaken by some esoteric copyright industry but rather by the music and movie industries. Consequently, the reporting and service provider industries have moved towards an increasing use of electronic formats for submitting takedown notices. Electronic format notices in web forms, which currently account for more than 98.0 percent of all notices, make it far easier to manage the takedown notices and avoid errors when processing the takedown requests. Well-implemented web forms with validation scripts can catch notices that do not have any provider or reporter information, lack proper descriptions of the copyright work to support the claims, or are missing the requisite takedown request for the allegedly infringing material. Such changes are clearly feasible and can easily be industry driven. This change should be welcomed.

Any disquiet from content providers, reporters, or even service providers who still insist on physical notices can be overcome by confirming in § 512(c) that the “written communication” to the designated agent of the service provider¹⁹⁸ can be in a “reasonable format prescribed by the agent.” This paper has shown that as of 2012, 99.6 percent of all notices in the first dataset are in the form of web forms or emails, and empowering the service provider to choose the

¹⁹⁸ 17 U.S.C. § 512(c)(3)(A).

electronic format for submission of the takedown notice reflects the actual practice today. Just as the DMCA provides that the reporter may be contacted via email,¹⁹⁹ it is reasonable to have a service provider accept email notices in lieu of web form notices, since it cannot be assumed that all service providers have the ability to implement the latter, but the format selected has to be objectively reasonable. Just as in this day and age of the Internet, it would be patently *unreasonable* for the service provider to *only* require takedown notices to be sent by letter or by fax, it would also be *unreasonable* for the service provider to require reporters to supply takedown notices in esoteric electronic formats. The use of electronic takedown notices may raise valid considerations about forgeries and impersonation²⁰⁰ but these can be dealt with by requiring other forms of validation such as prior registration of the reporter or implementing an email callback verification scheme to validate the reporter's e-mail address and other particulars.

B. The Growth of Notices with Mega Claims and Mega Requests

It is disturbing to see the trend where more claims and more takedown requests are packed into each takedown notice. Up until 2010, each notice contained only one claim, but in 2011, the average number of claims per notice was 2.18, and in 2012, this average was 5.05. Consequently, the average number of requests per notice had risen to 47.79 in 2011 and 124.75 in 2012. These increasing averages, however, paint a slightly

¹⁹⁹ *Id.* § 512(c)(3)(A)(iv).

²⁰⁰ See Enigmax, *Anti-Piracy Co. Blames Hack for Bogus DMCA's, but They're Just Sloppy*, TORRENTFREAK, (Mar. 7, 2012), <http://torrentfreak.com/anti-piracy-co-blames-hack-for-bogus-dmcas-but-theyre-just-sloppy-120307>.

misleading picture. More than 65 percent of all reporters have only issued one notice, and almost 95 percent of all reporters have issued no more than ten notices in 2012. Of these, individual reporters account for 5.3 percent of all takedown notices, although they average only 1.67 claims and 13.69 requests per notice. Naturally, questions will arise as to between a “mega” notice with thousands of takedown requests and a “micro” notice by an unsophisticated reporter, whether takedown notices of individual reporters will receive lower priority for processing purposes.

With Google lifting the cap on the number of requests per notice and enabling more TCRP participants to pack even more claims and takedown requests into one notice, the individual or the small copyright provider is right to be concerned that her single, monolithic takedown notice will receive less attention. Likewise, as between TCRP participants, there may be concerns that if all notices are processed in the order they are received, the participant who packs more claims and takedown requests into her notice will have more of her claims and requests handled first. Under such an arrangement, it could lead to an unhealthy race between participants to see who can put in the largest number of claims and takedown requests in each notice. After all, if notices are processed sequentially and in the order they are received, a “mega” notice will arguably give the reporter a better chance to have her takedown requests attended to first.

Some may consider § 512(c)(3)(A)(ii) of the DMCA, which allows a takedown notice to cover “multiple copyrighted works at a single online site,” as allowing a notice to be constituted with multiple claims of infringed works. But this provision is intended to have a much narrower reach. Both the Senate and House Reports gave the example of a multi-claim notice that was submitted to target an unauthorized Internet

jukebox operating from a “single online site” site hosted by that service provider. In such a case, a single § 512(c) notice with a representative list of infringed compositions or recordings suffices, “so that the service provider can understand the nature and scope of the infringement being claimed.”²⁰¹ It would be outside the legislative object of this provision to interpret § 512(c)(3)(A)(ii) to allow for a notice to be partitioned into multiple claims which target an indiscriminate number of subdomains which are under the operational control of different subscribers (as is the case with § 512(c) hosting services such as Blogger) or thousands of links to unrelated sites (as is the case with § 512(d) Search notices).

Even if the DMCA does not disqualify a notice if it is partitioned into claims and takedown requests are made within each claim, the concern here is to ensure that every takedown notice that is submitted for processing by a service provider receives roughly equal visibility, so that there is equality of treatment for all takedown requests. “Bundling” multiple claims that do *not* relate to a “single online site” cannot impute to the service provider the requisite knowledge as to the nature and scope of the infringement claimed.²⁰² Where “bundling” through multiple unrelated claims is done to unfairly secure priority processing of the notices, this must be strongly discouraged, particularly where the service provider has to handle large quantities of notices and the takedown requests are processed on a “first come first serve” basis in the order in which notices are received.

²⁰¹ S. REP. NO. 105-190, at 46 (1998); H.R. REP. NO. 105-551, pt. 2, at 55 (1998).

²⁰² S. REP. NO. 105-190, at 46; H.R. REP. NO. 105-551, pt. 2, at 55.

The solution therefore is to ban this practice of “bundling” multiple unrelated claims altogether. This could be done by inserting the following italicized words to § 512(c)(3)(A)(ii), to read “identification of the copyrighted work claimed to have been infringed, or, *only* if multiple copyrighted works at a single online site *shown to be under the same operational control* are covered by a single notification, a representative list of such works at that site, *but not otherwise*” to clarify that multiple claim notices are the exception. It is also clear that the above language has no reference to § 512(d) notices, thus multiple claim notices for § 512(d) notices should be completely disallowed.

Disallowing multiple claim notices does not mean that the parties cannot file multiple claims. They can still do so but not by way of filing all of them in one notice. The correct way would be to file separate notices for this purpose. Any inconvenience from re-entering the required formalities data can be solved by using input convenience tools for this purpose, and not by conflating multiple notices.

C. URIs as Takedown Requests

As noted above, while the language of the DMCA does not require the reporter to necessarily supply the URI to the infringing material as part of the takedown request, this has overwhelmingly been the practice of reporters, both sophisticated and “nonsophisticated.”²⁰³ The second dataset shows that more than fifty-four million URIs as takedown requests have been submitted in 2012. Formally sanctioning this practice in the DMCA is the next logical step. By stating

²⁰³ See Tushnet, *supra* note 143.

that information reasonably sufficient to permit the service provider to “*unambiguously*” locate the material “*includes* the Uniform Resource Indicator of the material,”²⁰⁴ this amendment will foreclose on the possibility that reporters can simply supply search terms as takedown requests, even search terms that masquerade as URIs. It will preserve the requisite precision that reporters are required to demonstrate when locating infringing material before service providers will act on their requests, reduce the burden on the service provider, and improve on its ability to act on the requests expeditiously. And it will also make it feasible for service providers to consistently explain to “nonsophisticated” reporters what a URI is and request it from all reporters to ensure that their notices are successfully processed.

D. Greater Transparency about the Trusted Copyright Removal Program

In a sense, the TCRP is a victim of its own success. Designed as a scheme to enable “sophisticated” reporters like large content providers, CMOs, and reporting agents to quickly and efficiently make takedown submissions, it is now responsible for 79.65 percent of all takedown notices and 91.50 percent of all takedown requests received and processed by Google in 2012. Operating by discriminating between reporters during the takedown process, Google ought to be more transparent towards the Internet community as to how the TCRP works. It is particularly troubling to learn that while the average takedown rate is 97.5 percent,²⁰⁵ the average takedown rate of the TCRP participants as more “sophisticated” reporters

²⁰⁴ 17 U.S.C. § 512(c)(3)(A)(iii) (emphasis added).

²⁰⁵ See HOW GOOGLE FIGHTS PIRACY, *supra* note 142, at 14.

is worse at 96.2 percent. Participants to the TCRP also deserve more information as to how they perform *inter se*. Publication of this information with, perhaps, prescribed takedown rates that participants have to meet to stay on the TCRP will encourage a “race-to-the-top,” with improved detection methods that will further cut down on the false positives that erroneously target legitimate content sites.²⁰⁶ Greater transparency will therefore serve as a “report card” for reporters and give content providers and owners better insight into the work of reporters. The only losers will be the shoddy and inaccurate reporters who do not deserve to be issuing takedown notices in the first place. That way, the Internet community and the market can better decide who the quality reporters are.

²⁰⁶ See, e.g., Ernesto, *Movie Studios Ask Google to Censor Their Own Films, Facebook and Wikipedia*, TORRENTFREAK (Dec. 3, 2012), <https://torrentfreak.com/movie-studios-ask-google-to-censor-their-own-films-facebook-and-wikipedia-121203> (reporting that a reporting organization known as “Yes It Is – No Piracy!” ostensibly sought to takedown legitimate copies of films on behalf of the movie studios on Verizon, Amazon, iTunes, and so on, and even remove film reviews published by The Guardian, The Independent, The Mirror, and the Daily Mail); Ernesto, *HBO Wants Google to Censor HBO.com*, TORRENTFREAK (Feb. 3, 2013), <http://torrentfreak.com/hbo-wants-google-to-censor-hbo-com-130203>; see also Leigh Beadon, *You’re All the Weakest Link: Bad Law Permits Bad Takedowns, Which Google Handles Badly*, TECHDIRT (Apr. 22, 2013, 2:21 PM), <http://www.techdirt.com/articles/20130422/09303922801/youre-all-weakest-link-bad-law-permits-bad-takedowns-which-google-handles-badly.shtml>.

E. Section 512(d) Notices and Section 512(g) Counter-Notifications

Finally, § 512(g) counter notifications should serve as the counterfoil to the excesses of takedowns. As the DMCA Senate Report explained, the takedown and counter notice provisions “balance the need for rapid response to potential infringement with the end-users legitimate interests in not having material removed without recourse.”²⁰⁷ Perhaps the datasets used in this study do not accurately capture all the counter notices submitted. Perhaps the bulk of the takedowns are legitimate and there is little room for the alleged infringers to dispute otherwise. Or perhaps both the alleged and innocent infringers have eluded the takedown by moving their content to another site. Whatever the reason, the mechanism for a counter notice in response to a §512(c) takedown is clear, but the mechanism for a response to a § 512(d) takedown is not. And that lack of “procedural protection”²⁰⁸ is believed to contribute to the paucity of counter notices. Surely, the existence of only *two* § 512(d) counter notices in response to *fifty-four million* takedown requests notices in 2012 is cause for concern.

Here, there are practical difficulties in requiring an information location service provider to contact the affected site owner in response to a § 512(d) takedown. The GTR published by Google is a good first step and publishing the takedown notice in the search results as well as notifying the site owner who happens to be a Google Webmaster user are all steps in the right direction. But perhaps more can be done to bring this information to the attention of the site owner. Currently, the GTR does not have any information regarding

²⁰⁷ S. REP. NO. 105-190, at 21.

²⁰⁸ *Id.*

the URIs that have been removed; it only has information about the domains which have received takedown URIs and the URIs which have *not* been removed. One potential solution is that Google can build a service in which a user can query for URIs that are removed pursuant to takedown notices and publicize its availability. While alleged infringers can indeed use such a tool to detect enforcement activity by reporters, given the highly asymmetrical imbalance between counter notices and takedown requests, innocent infringers will likely benefit more from such a tool. The language of the DMCA can also be clarified by requiring an information location service provider to undertake its best efforts to contact the owner or operator of the URIs taken down, or in the alternative, to publish information or tools which enable the owner or operator to know if their online resources have been targeted. This, together with the accompanying information advising the aggrieved infringer of her rights to serve a § 512(g) counter notice, will enable users to better manage any excessive takedowns and preserve the dynamic content of the Internet.

VI. CONCLUSIONS

The Executive Chairman of Google, Eric Schmidt, once said: “The [I]nternet is the first thing that humanity has built that humanity doesn’t understand, the largest experiment in anarchy we’ve ever had.”²⁰⁹ The same could also be said about the DMCA takedown mechanism. As a bold and innovative experiment that allows content owners, service providers and Internet users to arbitrate their differences, it has worked

²⁰⁹ Jerome Taylor, *Google Chief: My Fears for Generation Facebook*, THE INDEPENDENT (Aug. 18, 2010), <http://www.independent.co.uk/life-style/gadgets-and-tech/news/google-chief-my-fears-for-generation-facebook-2055390.html>.

remarkably well. And its success is due in no small part to its ability to temper our individual initiative and creativity with our commitment to participate in a positive way in this worldwide conversation.

But the only thing that is static about the Internet is change. We have changed from having individual copyright owners and their attorneys to having CMOs and reporting agents serve takedown notices. We have changed from sending notices with one or two takedown requests to sending thousands of notices with millions of requests to service providers. And we have changed from a manual system with individual review of notices to an automated system where both reporters and service providers use computers to process huge numbers of notices and requests with very short turnaround times. And with change comes the unavoidable consequences of owners and reporters having to grapple with unfamiliar electronic forms, of “micro” notices by individual reporters getting overwhelmed by the “mega” notices with thousands of takedown requests, of ill-informed copyright owners and reporters submitting vague, ambiguous, and abusive takedown requests, of lack of transparency about arrangements for automated submission and processing of “mega” notices, and of incessant user failure to respond to and correct abusive and erroneous takedowns.

What improvements can we make? For a start, we can enrich ourselves from the lessons we have learned in the sixteen years since the DMCA. We know that we can improve on the communication of takedown notices and make them more reliable and more trustworthy. We know that we can treat all notices by all reporters fairly and equally. We know that we can be accurate about infringing content that we have detected, and we can be precise about removing that content. We know that we want the quality of takedown notices to improve even

further. And we know that we have to preserve the online voices of the small and disenfranchised, who may otherwise be silenced forever.

Because it is only through constructive change that one can continue to preserve and uphold the very values that make the Internet special: its unfathomable richness, its encompassing inclusiveness, and its liberating freedom.

List of Extracted Fields in Takedown Notices

1. Notice ID
2. Google's Search Notice ID
3. Notice Date
4. Chilling Effects Notice Topic
5. URI of Notice on Chilling Effects Repository
6. Copyright Owner (Sender)
7. Copyright Owner Designation
8. Copyright Owner's Country
9. Reporter
10. Reporter's Country
11. Recipient
12. Recipient Service/Division
13. Recipient Country
14. Subject
15. Notice Form
16. Nature of Notice
17. Format of Notice
18. Description of Copyright Work
19. URI of Copyright Work
20. Count of URIs for Copyright Work
21. Description of Infringing Material or Activity
22. URI of Infringing Material or Activity/Takedown Request
23. Count of URIs for Infringing Material
24. URI of Infringing Material Hosted by Third Party
25. Count of URIs for Infringing Material Hosted by Third Party
26. Action taken in relation to URI/Takedown Request
27. Description of Response to Takedown Request
28. Flag: Does notice have acknowledgment of accuracy?
29. Flag: Does notice have statement of good faith belief?
30. Flag: Does notice have statement of authority to act?
31. Flag: Does notice have signature?
32. Flag: Is signature a trusted user signature?
33. Flag: Does notice have a response to takedown request?
34. Flag: Is notice a Counter notice?
35. Respondent to Counter notice
36. Response of Respondent to Counter notice

Table 1: Takedown Notices by Recipient and Year (N₁* = 543,778)¹

Year	Google	Twitter	Yahoo	Digg	Others	Total
2001	2		0		10	12
2002	67		0		33	100
2003	145		1		48	194
2004	249		0	0	54	303
2005	435		0	0	51	486
2006	217		3	0	34	254
2007	177	0	0	1	57	235
2008	62	0	4	27	52	145
2009	4,275	0	23	33	56	4,387
2010	16,827	307	508	21	65	17,728
2011	67,571	4,138	0	13	76	71,798
2012	441,370	6,646 ²	0	2	120	448,138

¹ Because the number of Twitter notices was adjusted upwards, the total number of notices in the dataset N₁ as reflected in Table 1 is 543,778 notices.

² See *supra* text accompanying note 43 of article.

Table 2: Numbers of Notices Issued by Top-20 Content Providers/Owners Between 2008 and 2012 (N₁ = 539,558).³

Rank	Content Provider/Owner	2008	2009	2010	2011	2012	Total	Industry
1.	BPI	0	0	87	3,620	187,385	191,092	Music
2.	Froytal	0	0	0	2,341	22,758	25,099	Adult Entertainment
3.	CA Co.	0	0	0	0	15,483	15,483	Adult Entertainment
4.	Microsoft	1	8	111	3,804	10,547	14,471	Software/Games
5.	IFPI	12	215	1,277	1,339	8,599	11,442	Music
6.	Intellectual Property Promotion Association	0	0	0	0	11,177	11,177	Adult Entertainment
7.	Sony Pictures	0	12	33	59	8,320	8,424	Movies
8.	RK NetMedia	0	4	89	16	6,371	6,840	Adult Entertainment
9.	20 th Century/Fox	1	27	414	218	5,358	6,018	Movies
10.	BangBros	1	6	36	48	5,243	5,334	Adult Entertainment

³ Note that there may be duplicate counting of notices in instances where a reporting agent represents more than one content provider. This is quite typical with notices submitted on behalf of the major movie studios like Columbia, Buena Vista/Disney, Paramount, and Sony.

⁴ This total is the total of all notices issued by the same content providers/owners from 2001 to 2012.

11.	HBO	0	0	19	243	4,861	5,123	Movies
12.	Publishers Association	0	0	0	427	4,561	4,988	Books
13.	RIAA	0	76	212	846	2,302	3,436	Music
14.	Magnolia Pictures	0	83	346	1,928	678	3,035	Movies
15.	Lionsgate	0	4	11	23	2,350	2,388	Movies
16.	Screen Gems	0	0	1	2	2,244	2,247	Movies
17.	Cricket Australia	0	0	16	562	1,598	2,176	Broadcasts
18.	GGCash	0	0	0	0	2,170	2,170	Adult Entertainment
19.	Paramount Pictures	2	11	55	178	1,845	2,091	Movies
20.	Take Two Interactive	0	0	0	1	2,076	2,077	Software/Games

**Table 3: Average and Median Number of Notices Submitted by Each
Provider Between 2008 and 2012**

Year	Maximum	Average	Standard Deviation	Median
2008	12	1.66	1.81	1.00
2009	215	2.03	7.69	1.00
2010	1,277	3.42	22.71	1.00
2011	3,804	4.33	55.55	1.00
2012	187,385	24.49	1,418.32	1.00

Table 4: Top 30 Reporters by Total Volume of Notices (2008-2012) (N=539,558)⁵

Rank	Reporter	Agent Type	Notices					
			2008	2009	2010	2011	2012	Total (2008-2012)
1.	BPI	CMO	0	0	0	4,394	187,346	191,740
2.	Degban	Agent	0	0	0	6,205	77,868	84,073
3.	No Information	Unclear	0	508	68	13,394	7,755	21,725
4.	AudioLock	Agent	0	0	0	0	21,134	21,134
5.	[Private]	Individual ⁶	78	1,272	5,223	4,816	9,027	20,416
6.	Peer Media Technologies*	Agent	0	39	15	25	14,556	14,635
7.	Marketly	Agent	0	0	0	3,398	10,245	13,643
8.	IFPI	CMO	12	180	1,261	1,565	8,520	11,538
9.	Irdeto*	Agent	0	0	0	68	10,452	10,520
10.	Copyright Integrity International*	Agent	0	45	123	2,884	4,170	7,222
11.	Muso	Agent	0	0	0	2	6,595	6,597
12.	Watch-Over*	Agent	0	0	0	44	5,285	5,329
13.	Removeyourcontent	Agent	3	150	1,815	1,583	1,608	5,159

⁵ Reporters marked with "*" are trusted reporters under the Google Trusted Copyright Removal Program. See *supra* text accompanying note 124.

⁶ Note that the Chilling Effects database redacts private, personally identifiable information from the notices. In most instances, these individual reporters are identified as "redacted" or "private." For purposes of Table 4, it will be assumed that reporters with redacted information or identified as "redacted" or "private" are individuals rather than corporations.

14.	Web Sheriff*	Agent	0	360	834	2,810	1,109	5,113
15.	Publishers Association	CMO	0	0	0	423	4,566	4,989
16.	Takedown Piracy	Agent	0	0	24	1,649	2,989	4,662
17.	DtecNet	Agent	0	0	22	343	4,153	4,518
18.	RIAA*	CMO	0	73	211	1,489	2,305	4,078
19.	Ripblock	Agent	0	7	608	183	3,039	3,837
20.	NetResult*	Agent	1	103	434	1,259	1,684	3,481
21.	Digital Copyright Consultancy	Agent	0	0	0	0	2,636	2,636
22.	Morganelli	Agent	0	0	0	0	2,510	2,510
23.	Contributor*	Agent	2	7	281	1,073	698	2,061
24.	MiMTiD	Agent	0	0	158	817	849	1,824
25.	Remove Your Media*	Agent	0	0	0	11	1,491	1,502
26.	Ripoff*	Agent	0	0	0	75	1,267	1,342
27.	Police Du Net	Agent	0	0	0	18	1,254	1,272
28.	RaymondCC Tech	Individual	0	0	0	0	1,189	1,189
29.	Warner Bros. Entertainment	Industry	0	0	0	163	1,001	1,164
30.	Random Walk Trading	Industry	0	0	0	0	1,136	1,136

Table 5: Average and Median Number of Notices Served by Each Reporter Between 2008 and 2012

Year	Maximum	Total	Average	Standard Deviation	Median
2008	78	141	3.28	11.80	1.00
2009	1,272	4,387	5.00	48.67	1.00
2010	5,223	17,727	7.61	120.51	1.00
2011	13,394	71,800	9.36	199.27	1.00
2012	187,346	443,937	44.46	2,061.06	1.00

**Table 6: Formats of Takedown Notices Between 2008 and 2012
(N₁=539,558)**

Year	Form		Email		Mail		Fax		Others		Total
2008	6	4.3%	109	77.3%	6	4.3%	8	5.7%	12	8.5%	141
2009	2,527	57.6%	1,495	34.1%	22	0.5%	138	3.1%	205	4.7%	4,387
2010	7,647	43.1%	8,803	49.7%	90	0.5%	581	3.3%	606	3.4%	17,727
2011	61,378	85.5%	8,907	12.4%	22	0.0%	126	0.2%	1,356	1.9%	71,789
2012	438,542	98.8%	3,723	0.8%	9	0.0%	1	0.0%	1,636	0.4%	443,911

Table 7: Number of Requests Issued by Top-20 (by Total) Content Providers/Owners Between 2008 and 2012 (N₂=501,286)

Rank	Content Provider/Owner	Requests					Total	Industry
		2008	2009	2010	2011	2012		
1.	RIAA*	0	0	0	37,508	7,632,934	7,670,442	Music
2.	Froytal*	0	0	0	298,568	7,037,258	7,335,826	Adult Entertainment
3.	Microsoft*	0	7	72	806,500	6,059,122	6,856,701	Games/Software
4.	BPI*	0	0	3	74,451	4,918,268	4,992,722	Music
5.	20 th Century/Fox*	0	2	63	20	2,697,728	2,697,813	Movies
6.	NBCUniversal	0	0	0	257,645	2,174,825	2,432,470	Broadcasts
7.	RK NetMedia*	0	0	0	0	1,760,800	1,760,800	Adult Entertainment
8.	Warner Music Group	0	0	0	17,541	1,445,664	1,463,205	Music
9.	Universal Music Group	0	323	61	112,311	708,613	821,308	Music
10.	FUNimation Entertainment	0	0	0	3,129	750,636	753,765	Movies
11.	CA Co*	0	0	0	0	739,818	739,818	Adult Entertainment
12.	Evil Angel	0	0	12	60,019	575,441	635,472	Adult Entertainment
13.	Warner Bros. Entertainment	0	0	8	12,376	615,620	628,004	Movies
14.	Lionsgate*	0	0	2	3,532	622,162	625,696	Movies
15.	Lynda.com	0	0	0	35,173	587,871	623,044	Website
16.	Hydentra	0	0	0	37,364	573,699	611,063	Adult Entertainment

17.	XFC	0	0	0	45,955	557,700	603,655	Broadcasts
18.	Sony Music	0	3	0	52,273	550,842	603,118	Music
19.	Magnolia Pictures*	0	0	0	17,556	557,970	575,526	Movies
20.	Adobe	0	0	0	0	437,252	437,252	Games/Software

Table 8: Statistics of Notices, Claims and Requests Between 2008 and 2012 (N₂=501,286)

Year	Total Notices	Total Claims	Min Claim/Notice	Max Claim/Notice	Avg Claim/Notice	SD Claim/Notice	Total Requests	Min Request/Notice	Max Request/Notice	Avg Request/Notice	SD Request/Notice	Median Request/Notice
2008	6	6	1	1	1.00	0.00	14	1	5	2.33	1.37	2.00
2009	2,434	2,434	1	1	1.00	0.00	15,274	1	2,276	6.27	48.50	1.00
2010	7,374	7,374	1	1	1.00	0.00	27,035	1	269	3.67	9.83	1.00
2011	56,404	124,794	1	997	2.18	6.99	2,685,474	1	9,515	47.79	169.83	2.00
2012	435,063	2,200,143	1	1,484	5.05	34.44	54,263,600	1	25,050	124.75	718.23	4.00

Table 9: Top 30 Reporters by Total Requests in 2012 (N2=501,286)⁷

Rank	Reporter	Type	Claims/Notice				Requests/Notice				1 st Quartile	3 rd Quartile	Days Enforced	Avg Days/Notice			
			Total Claims	Min	Max	Avg	Total Requests	Min	Max	Avg					SD	Median	
1.	Degban+	Agent	77,869	123,920	1	101	1.59	11,741,358	1	25,050	150.78	856.85	8.00	1.00	102.00	306	1.20
2.	RIAA+	CMO	2,211	971,372	1	1,484	439.53	7,632,938	1	18,012	3,452.26	2,687.82	2,743.00	1,085.56	5,489.00	182	1.88
3.	BPI+	CMO	187,345	674,069	1	100	3.60	6,927,042	1	11,920	36.97	191.83	3.00	1.00	31.00	360	1.02
4.	Takedown Piracy+	Agent	2.89	8,151	1	10	2.73	5,825,883	1	10,000	1,949.11	3,305.36	301.00	24.00	4,993.00	307	1.19
5.	DtecNet+	Agent	3,345	6,327	1	10	1.89	3,669,072	1	9,994	1,096.88	2,255.87	212.00	33.00	1,523.30	309	1.18
6.	Marketly+	Agent	9,644	69,781	1	10	7.24	3,514,496	1	2,595	364.42	424.25	111.00	31.00	998.00	246	1.49
7.	20 th Century/Fox+	Owner	821	3,464	1	10	4.20	2,572,993	1	10,000	3,133.97	3,502.97	1,306.00	50.86	7,811.96	146	2.44
8.	NBCUniversal+	Owner	762	4,186	1	19	5.49	2,174,846	1	11,221	2,854.13	2,854.39	1,779.00	269.68	6,304.00	109	3.35
9.	BAF+	CMO	537	6,453	1	99	11.95	1,479,035	1	9,998	2,754.26	2,667.64	1,733.00	68.10	4,998.00	84	1.72
10.	Remove Your Media+	Agent	1,492	3,463	1	10	2.32	1,425,304	1	10,000	955.30	1,564.28	361.00	32.00	1,633.78	233	1.56
11.	Morganell+	Agent	2,510	4,808	1	73	1.92	1,001,911	1	12,743	399.17	1,037.66	49.50	2.00	500.00	112	1.61
12.	Removeyourcontent+	Agent	1,342	2,515	1	10	1.87	860,231	1	6,641	641.01	765.61	407.00	89.06	1,000.00	238	1.53

⁷ Reporters who are also members of the Google Trusted Copyright Removal Program (TCRP) are marked with a “+”.

13.	LeakID+	Agent	727	2,983	1	85	4.10	686,736	1	9,843	944.62	797.51	994.00	525.74	1,000.00	115	2.46
14.	AudioLock+	Agent	21,134	34,918	1	10	1.65	472,266	1	4,413	22.35	80.21	4.00	1.00	25.00	116	1.07
15.	Warner Bros. Entertainment+	Owner	1,001	1,001	1	1	1.00	361,925	1	1,000	361.56	415.54	127.00	4.00	995.00	180	1.92
16.	Publishers Association+	CMO	4,566	42,396	1	100	9.29	263,895	1	825	57.80	92.39	23.00	9.00	102.94	197	1.80
17.	MarkMonitor Inc.+	Agent	80	297	1	10	3.71	245,551	1	9,999	3,069.39	3,752.85	3,530.50	956.66	9,896.08	56	3.76
18.	Irdeto+	Agent	10,448	11,066	1	10	1.06	239,497	1	9,016	22.93	255.31	1.00	1.00	3.00	253	1.44
19.	Unidam Inc.+	Agent	554	554	1	1	1.00	232,179	1	1,000	419.10	350.28	328.00	48.00	915.86	74	1.29
20.	Ripblock+	Agent	3,039	3,385	1	9	1.11	227,485	1	2,219	74.86	144.42	28.00	6.00	106.26	227	1.41
21.	Digital Copyright Consultancy+	Agent	2,636	2,653	1	5	1.01	218,388	1	3,500	82.85	168.01	29.00	6.00	124.00	228	1.11
22.	DMCA Force	Agent	553	553	1	1	1.00	191,961	1	1,000	347.13	322.39	220.00	41.98	804.00	180	1.25
23.	Armovore+	Agent	100	255	1	10	2.55	166,191	1	9,997	1,661.91	2,838.80	425.50	27.50	2,773.00	24	8.52
24.	MiMTID+	Agent	823	868	1	6	1.05	161,255	1	5,502	195.94	472.39	63.00	10.00	299.82	251	1.45
25.	Microsoft+	Owner	796	5,911	1	10	7.43	129,799	1	1,904	163.06	257.82	72.00	17.86	228.56	57	6.14
26.	(redacted)	Indy?	9,425	15,684	1	10	1.67	127,743	1	1,000	13.69	65.49	1.00	1.00	10.00	242	1.50
27.	IFPI+	CMO	5,767	36,805	1	69	6.38	120,664	1	567	20.93	36.12	12.00	2.00	30.00	134	2.21
28.	Web Sheriff+	Agent	730	730	1	1	1.00	117,072	1	996	160.37	227.28	60.50	7.00	365.20	252	1.45
29.	Muso.com+	Agent	6,597	8,126	1	10	1.23	111,190	1	1,320	16.85	48.86	4.00	1.00	21.00	156	1.54
30.	Stichting BREIN+	CMO	102	4,427	1	100	43.40	110,962	1	4,644	1,087.88	1,016.69	776.00	159.20	2,168.04	50	2.51

Table 10: List of Trusted Users of the TCRP Between 2011 and 2012 (N₂=501,286)

Rank	Reporter as TCRP Trusted User	Type	Total Notices as Trusted User	Total URIs as Trusted User	Maximum No of URIs per Notice	Start Date of Reporting as Trusted User	Last Date of Reporting as Trusted User (within dataset)	Takedown Rate
1.	Degban	Agent	83,520	12,232,454	25,050	Jul 22, 2011	Dec 31, 2012	85.0%
2.	RIAA	CMO	2,484	7,707,854	18,012	Jun 9, 2011	Dec 31, 2012	99.1%
3.	BPI	CMO	191,467	7,158,867	11,920	Jun 6, 2011	Dec 31, 2012	97.7%
4.	Takedown Piracy	Agent	4,199	6,214,027	10,000	Aug 29, 2011	Dec 31, 2012	99.6%
5.	Markedly	Agent	13,039	4,146,070	2,595	Jun 3, 2011	Dec 31, 2012	99.4%
6.	DrecNet	Agent	3,447	3,792,698	9,994	Aug 31, 2011	Dec 31, 2012	98.5%
7.	20 th Century/Fox	Owner	811	2,572,979	10,000	Jul 24, 2012	Dec 31, 2012	96.8%
8.	NBCUniversal	Owner	927	2,306,527	11,221	Aug 2, 2011	Dec 31, 2012	94.2%
9.	BAF	CMO	537	1,479,035	9,998	Aug 9, 2012	Dec 30, 2012	99.5%
10.	Remove Your Media	Agent	1,247	1,350,984	10,000	Jul 3, 2012	Dec 31, 2012	98.3%
11.	Morganelli	Agent	2,509	1,001,847	12,743	Aug 19, 2012	Dec 31, 2012	97.9%
12.	Removeyourcontent	Agent	963	675,954	6,641	Jul 3, 2012	Dec 31, 2012	99.1%
13.	Audiolock	Agent	20,284	457,950	4,413	Sep 27, 2012	Dec 31, 2012	95.9%
14.	Warner Bros. Entertainment	Owner	1,138	374,159	1,000	Mar 4, 2011	Dec 21, 2012	92.4%
15.	Publishers Association	CMO	4,988	266,074	825	Jul 18, 2011	Dec 21, 2012	91.9%
16.	Irdeto	Agent	10,502	241,353	9,016	Nov 9, 2011	Dec 31, 2012	99.4%
17.	MarkMonitor Inc.	Agent	46	240,009	9,999	Nov 9, 2012	Dec 31, 2012	99.9%
18.	Unidam, Inc.	Agent	550	231,965	1,000	Sep 25, 2012	Dec 28, 2012	96.3%
19.	Digital Copyright Consultancy	Agent	2,181	189,569	3,500	Jun 29, 2012	Dec 31, 2012	87.5%
20.	MiMTID	Agent	1,072	173,279	5,502	Jul 1, 2011	Dec 29, 2012	97.2%
21.	Armovore	Agent	88	164,236	9,997	Jan 19, 2012	Feb 26, 2012	97.3%
22.	Web Sheriff	Agent	875	156,754	996	Jul 18, 2011	Dec 31, 2012	97.5%
23.	IFPI	CMO	6,088	134,805	567	Nov 3, 2011	Dec 31, 2012	92.4%

24.	Leakid	Agent	85	123,912	9,843	Nov 4, 2011	Aug 22, 2012	99.4%
25.	Microsoft	Owner	602	112,757	1,904	Jun 7, 2011	May 8, 2012	99.5%
26.	Muso	Agent	6,598	111,199	1,320	Dec 9, 2011	Dec 31, 2012	85.2%
27.	RipBlock	Agent	841	108,733	2,219	Oct 30, 2012	Dec 31, 2012	95.6%
28.	Stichting BREIN	CMO	64	98,537	4,644	Nov 1, 2012	Dec 28, 2012	99.9%
29.	Attributor	Agent	607	62,211	1,000	Nov 8, 2012	Dec 25, 2012	95.8%
30.	Peer Media Technologies	Agent	12,364	54,852	20	Sep 26, 2012	Dec 31, 2012	84.5%
31.	Entura International LTD	Agent	40	40,836	5,040	Nov 1, 2012	Dec 25, 2012	99.5%
32.	Walt Disney	Owner	125	28,436	4,069	May 7, 2012	Dec 21, 2012	99.2%
33.	Entertainment Software Association	CMO	178	20,443	1,322	Jan 15, 2012	Dec 18, 2012	99.2%
34.	ACIUJF	Agent	13	16,293	2,818	Aug 13, 2012	Dec 20, 2012	99.8%
35.	Vobile Inc.	Agent	1,093	13,645	65	Oct 12, 2012	Dec 31, 2012	64.6%
36.	The Adaptor	Agent	92	4,582	634	Aug 16, 2012	Dec 21, 2012	97.7%
37.	BayTSP	Agent	79	1,659	303	Jul 31, 2011	Feb 1, 2012	50.0%
38.	International Needlework Designers Association	CMO	31	1,312	81	Jun 26, 2012	Dec 14, 2012	-
39.	TEST ⁸	-	5	108	104	Nov 7, 2012	Dec 21, 2012	-
40.	We Delete	Agent	3	86	55	Mar 8, 2012	Mar 15, 2012	-
41.	IP-Echelon	Agent	1	26	26	Jul 25, 2012	Jul 25, 2012	-
42.	MÃœYAP (MÜYAP)	CMO	22	22	1	Nov 29, 2012	Dec 5, 2012	-
43.	Piracy Patrol	Agent	1	16	16	Nov 28, 2012	Nov 28, 2012	-
44.	Digital Platform İletisim Hiz.A.Ş.	Agent	1	11	11	Nov 7, 2012	Nov 7, 2012	-

⁸ The User “TEST” appears to be a test account for TCRP participants to make test submissions of takedown notices and requests. However, the user “Google Inc” also appears to be a test account for Google to make test submissions.

45.	Google Inc.	-	4	10	3	Aug 31, 2012	Nov 14, 2012	-
46.	Recording Industry Association of New Zealand (RIANZ)	CMO	1	3	3	Dec 28, 2012	Dec 28, 2012	-
Total			375,812	54,069,138				

Table 11: Notices and Requests for Google Services in 2011 and 2012 (N₂=501,286)

Google Service	DMCA Section	2011					2012				
		Notices	Requests	Avg Requests/Notice	Providers	Reporters	Notices	Requests	Avg Requests/Notice	Providers	Reporters
Ads/AdSense	Ads	281	317	1.13	204	143	178	0	0.00	10	9
Ads/Adwords	Ads	4	4	1.00	3	2	40	63	1.58	33	2
Blogger/Blogspot	s512(c)	21,656	50,362	2.33	9,343	6,015	19,220	58,052	3.02	6,117	3,973
Cloud/Docs	s512(c)	120	218	1.82	82	56	107	295	2.76	83	64
Cloud/Storage	s512(c)	2	4	2.00	2	2	1	1	1.00	1	1
Development/Code	s512(c)	42	74	1.76	35	29	45	93	2.07	34	26
Development/Sketchup	s512(c)	0	0	0.00	0	0	4	4	1.00	2	2
Gmail	s512(c)	20	8	0.40	20	16	10	0	0.00	9	8
Maps/Earth	s512(c)	1	1	1.00	1	1	1	1	1.00	1	1
Maps/Maps	s512(c)	154	275	1.79	147	133	71	96	1.35	67	62
Maps/Panoramio	s512(c)	53	57	1.08	46	10	33	52	1.58	21	7
Picasa	s512(c)	1,987	16,451	8.28	1,033	378	737	5,222	7.09	445	228
Play	s512(c)	0	0	0.00	0	0	1,380	2,144	1.55	479	419
Play/Android	s512(c)	799	490	0.61	417	356	563	575	1.02	284	252
Play/Artists	s512(c)	1	3	3.00	1	1	5	8	1.60	3	3
Play/Books	s512(c)	68	75	1.10	41	30	105	370	3.52	58	32
Play/Chrome	s512(c)	3	12	4.00	3	3	0	0	0.00	0	0
Play/Music	s512(c)	1	1	1.00	1	1	0	0	0.00	0	0
Play/Newsstand	s512(c)	0	0	0.00	0	0	1	0	0.00	1	1
Plus	s512(c)	139	302	2.17	110	65	937	3,213	3.43	609	392
Plus/Buzz (closed)	s512(c)	6	4	0.67	6	5	2	2	1.00	1	1
Plus/Google (closed)	s512(c)	6	5	0.83	5	6	0	0	0.00	0	0

Plus/Orkut	s512(c)	219	313	1.43	188	106	272	402	1.48	198	127
Plus/Profiles	s512(c)	116	242	2.09	93	55	28	14	0.50	24	19
Search	s512(d)	28,937	2,608,021	90.13	3,400	2,125	408,803	54,177,860	132.53	8,911	4,760
Search/Code (closed)	s512(d)	15	81	5.40	11	9	0	0	0.00	0	0
Search/Finance	s512(d)	2	2	1.00	2	2	1	1	1.00	1	1
Search/Groups	s512(c)	19	72	3.79	15	7	31	81	2.61	22	10
Search/ImageSearch	s512(d)	1,803	5,054	2.80	737	386	2,782	12,921	4.64	1,144	527
Search/News	s512(d)	8	11	1.38	8	7	21	38	1.81	14	14
Search/Places	s512(d)	13	13	1.00	13	12	2	2	1.00	2	2
Search/Product	s512(d)	660	1,081	1.64	48	46	107	481	4.50	23	20
Search/Schemer	s512(d)	0	0	0.00	0	0	3	11	3.67	3	3
Search/Video	s512(d)	3	97	32.33	3	3	2	2	1.00	1	1
Sites	s512(c)	538	1,158	2.15	369	279	525	1,073	2.04	397	302
Sites/Appengine	s512(c)	29	42	1.45	28	24	14	40	2.86	12	10
Video	s512(c)	62	164	2.65	50	34	19	341	17.95	16	16
Video/YouTube	s512(c)	0	0	0.00	0	0	4	6	1.50	2	2

Table 12: Counter Notices Between 2008 and 2012 (N₂=501,286)

Google Service	Originating DMCA Section	2011 Counter Notices	2012 Counter Notices
Unknown	N/A		5
Ads/Adsense	Ads	10	
Blogger	s512(c)	36	58
Picasa	s512(c)	2	1
Play	s512(c)		21
Play/Android	s512(c)	21	
Plus	s512(c)		1
Plus/Orkut	s512(c)		1
Search	s512(d)	6	2
Search/Product	s512(d)	2	
Total		77 (0.131%)	89 (0.020%)

Figure 1: Notices by Recipient Between 2001 and 2012

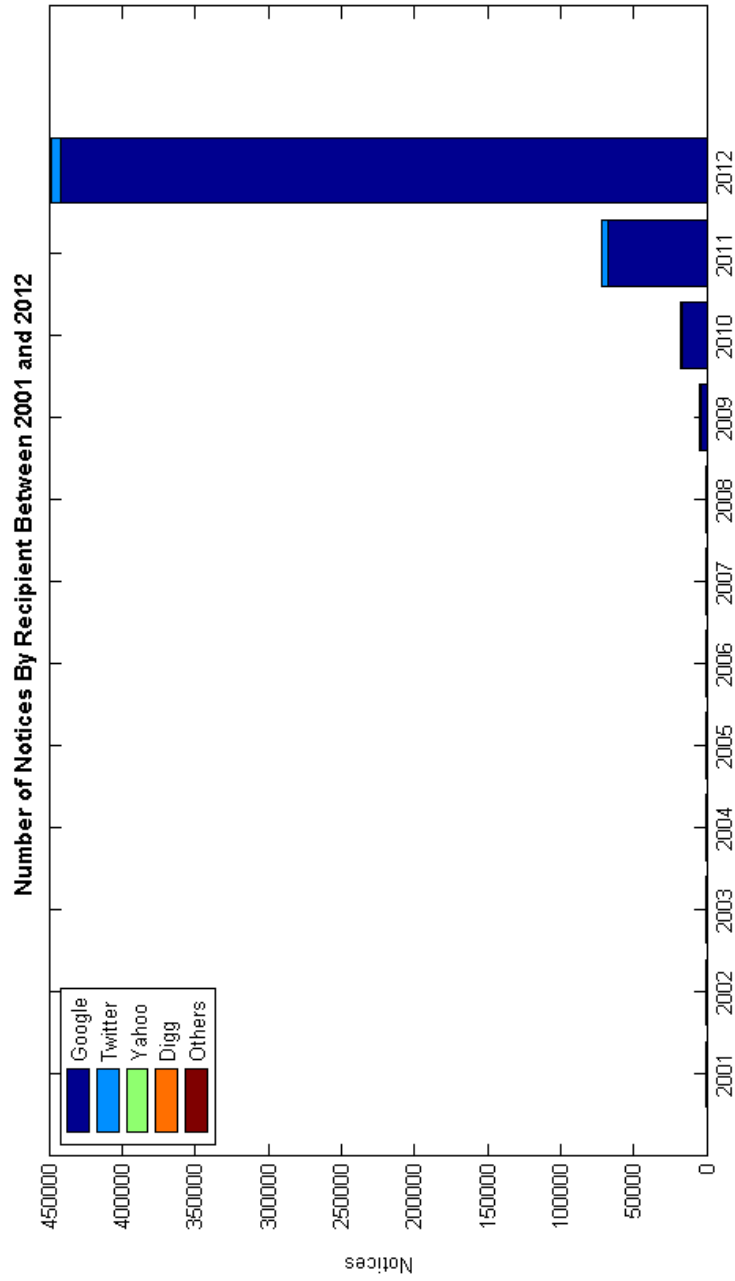


Figure 2: Notices by Recipient (Excluding Google) Between 2001 and 2012

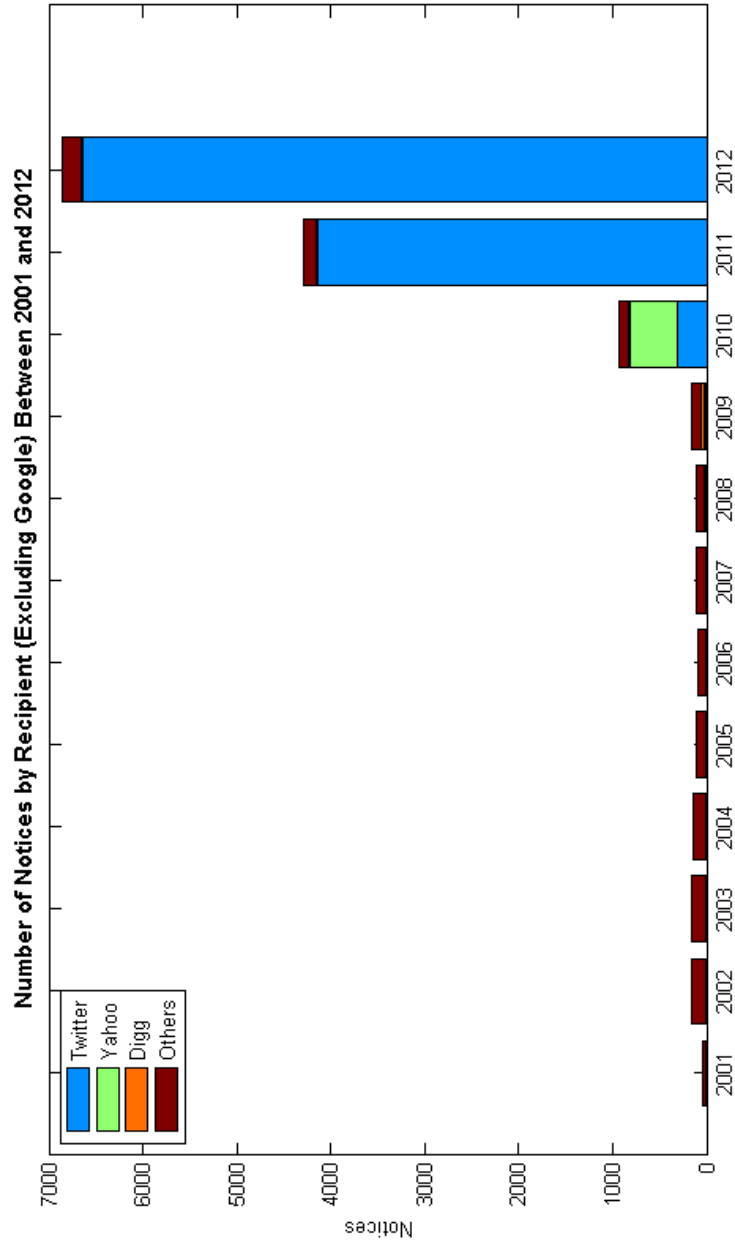


Figure 3: Notices by Recipient Between Jan. 2001 and Dec. 2012

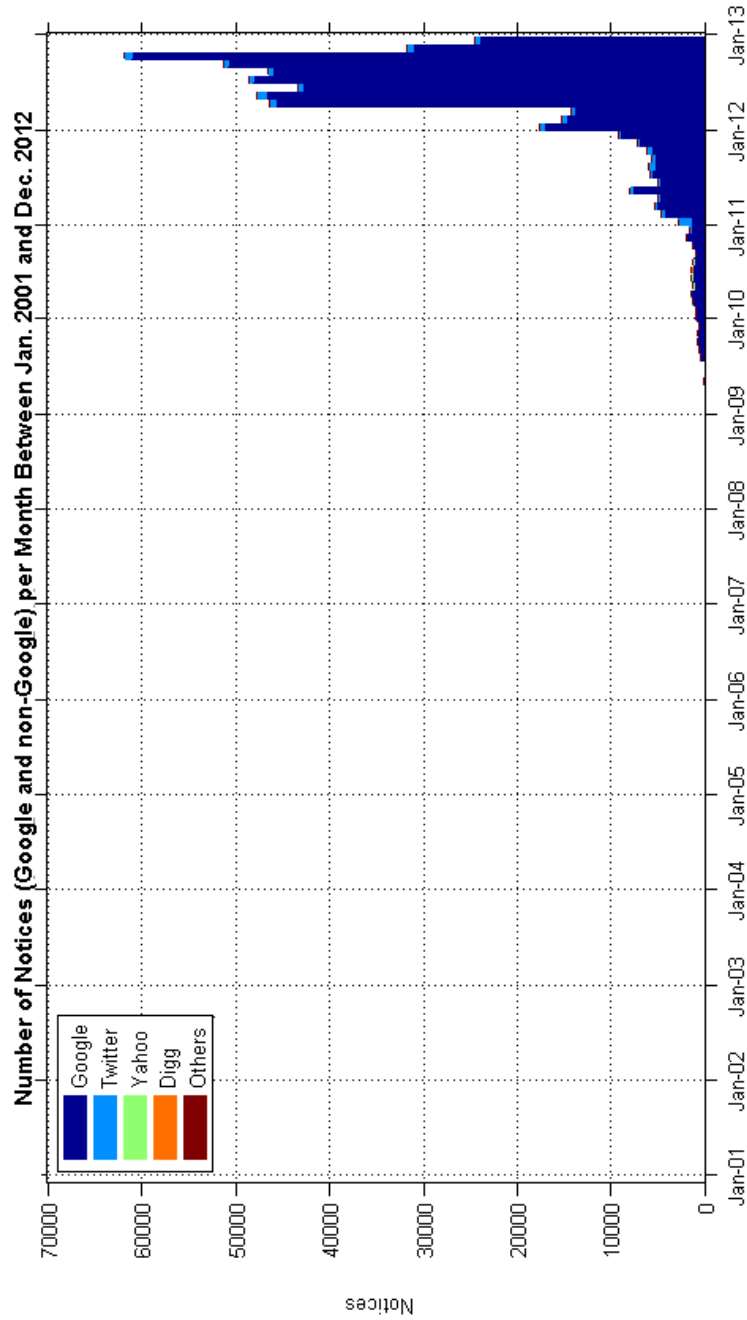


Figure 4: Notices by Recipient (Excluding Google) Between Jan. 2001 and Dec. 2012

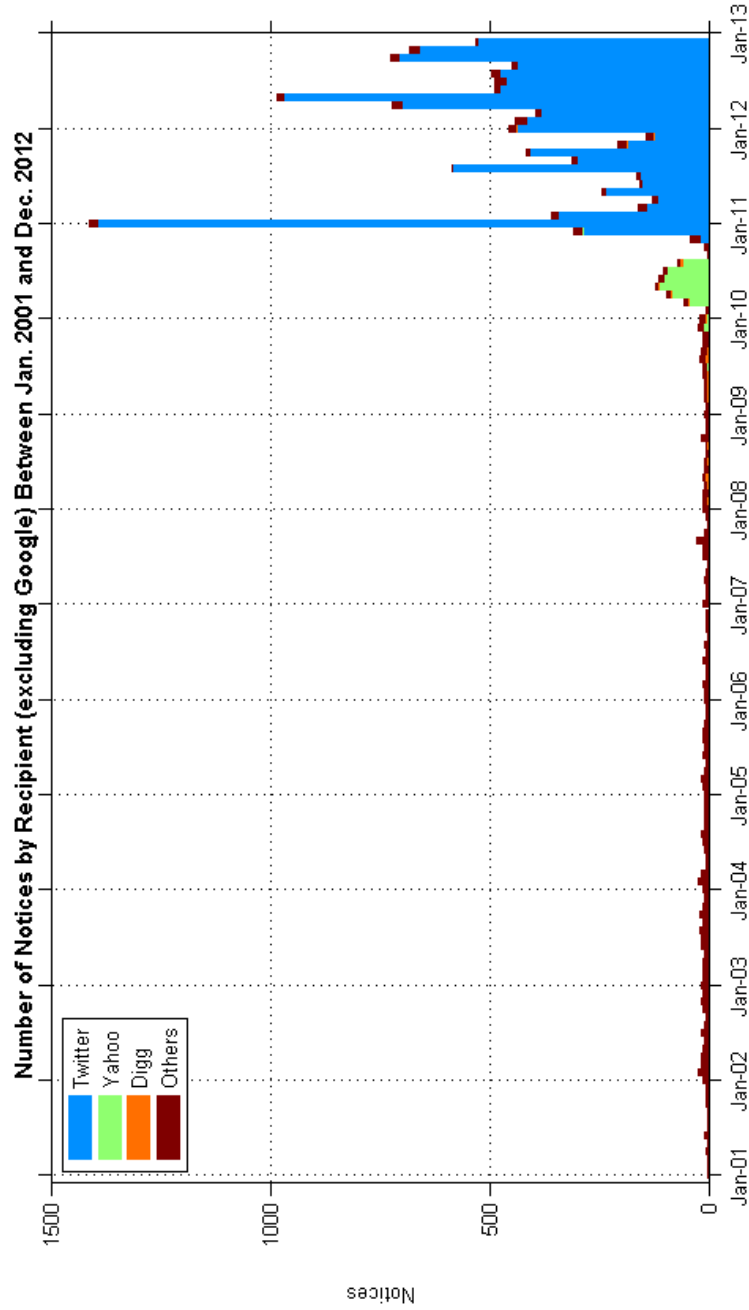
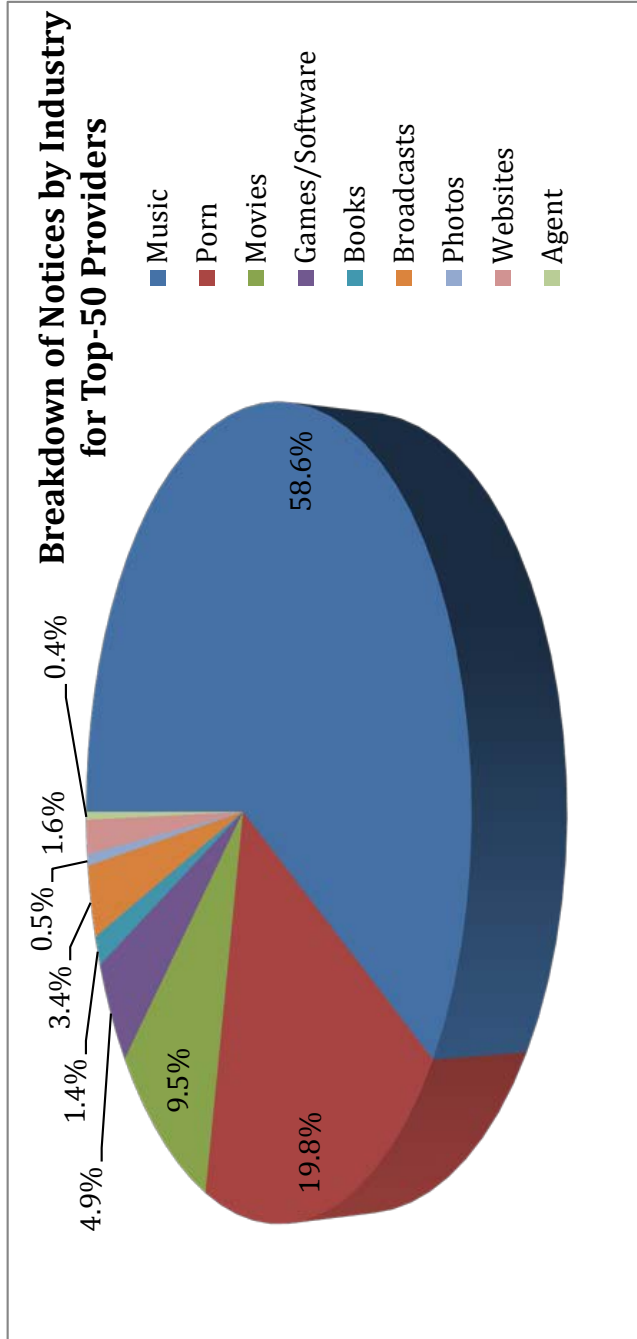
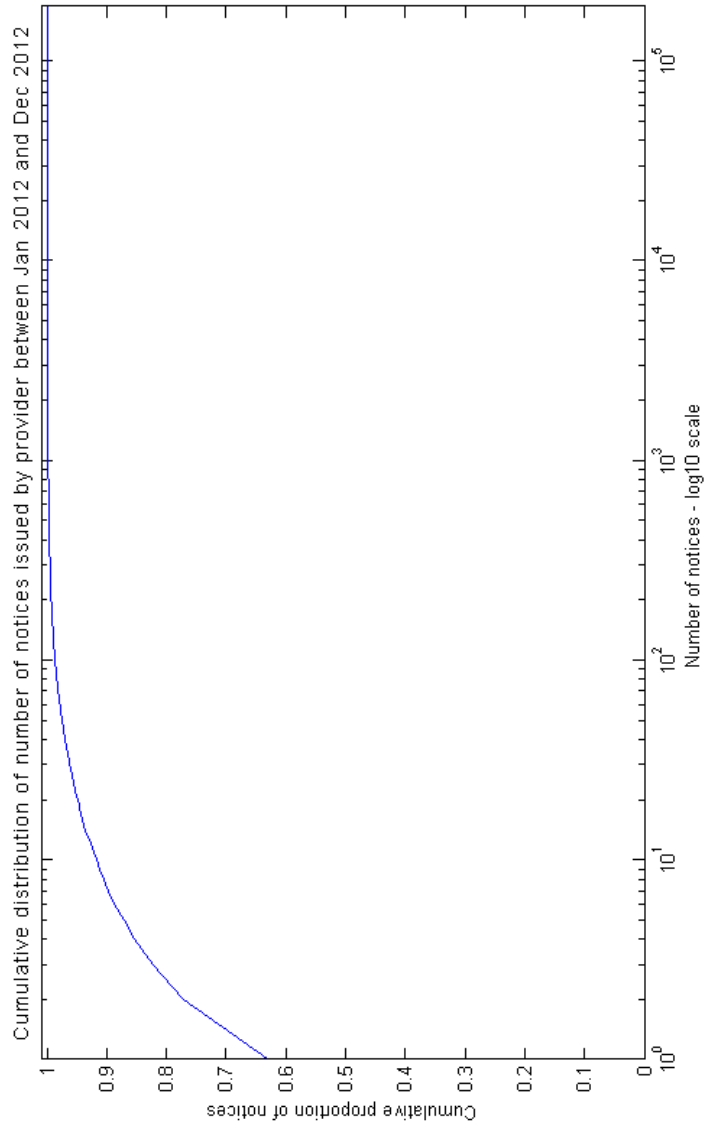


Figure 5: Breakdown of Notices by Industry for Top 50 Providers Between 2008 and 2012 (N=362,926).¹



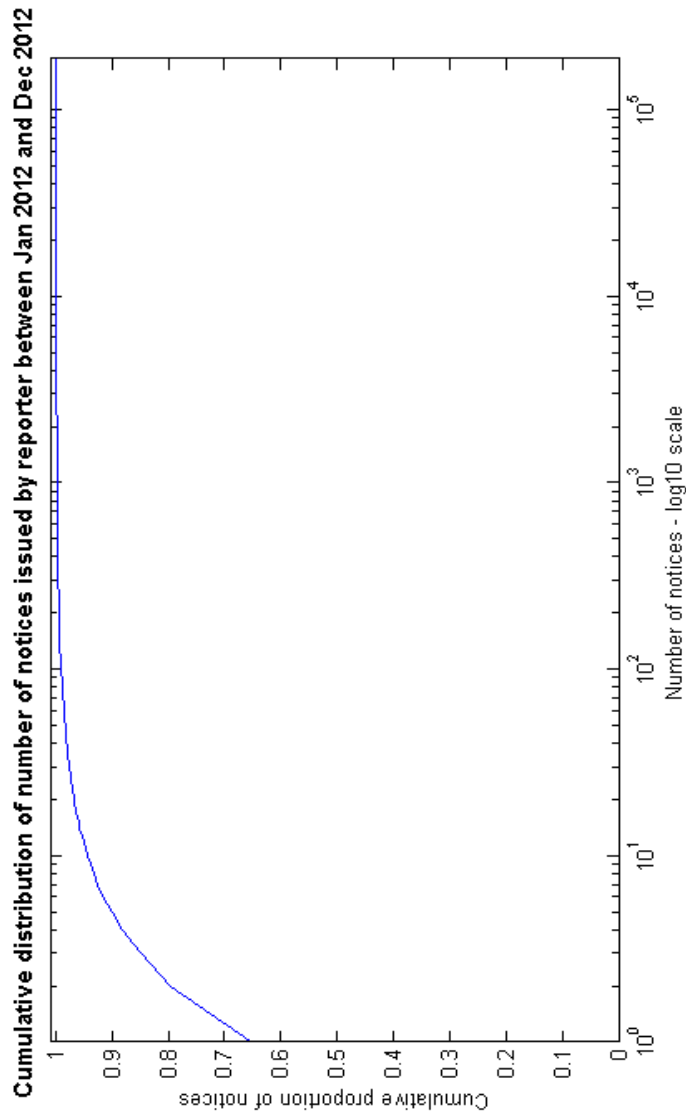
¹ The numbers represent the proportion of notices for the top fifty senders by industry vis-à-vis all notices sent by the top fifty senders. The Chilling Effects metadata identify several reporting agents as owners—these make up 0.4 percent of all senders of notices in the period under review.

Figure 6: Cumulative Distribution Frequency of Number of Notices by Each Provider in 2012



No. of Notices	CDF
1	62.85%
5	87.15%
10	91.74%
50	97.50%
100	98.65%
1,000	99.84%
10,000	99.97%
100,000	99.99%

Figure 7: Cumulative Distribution Frequency of Number of Notices by Each Reporter in 2012



No. of Notices	CDF
1	65.14%
5	90.02%
10	94.23%
50	98.30%
100	99.03%
1,000	99.71%
10,000	99.94%
100,000	99.99%

Figure 8: Proportion of Notices by Format Between 2001 and 2012 (Monthly)

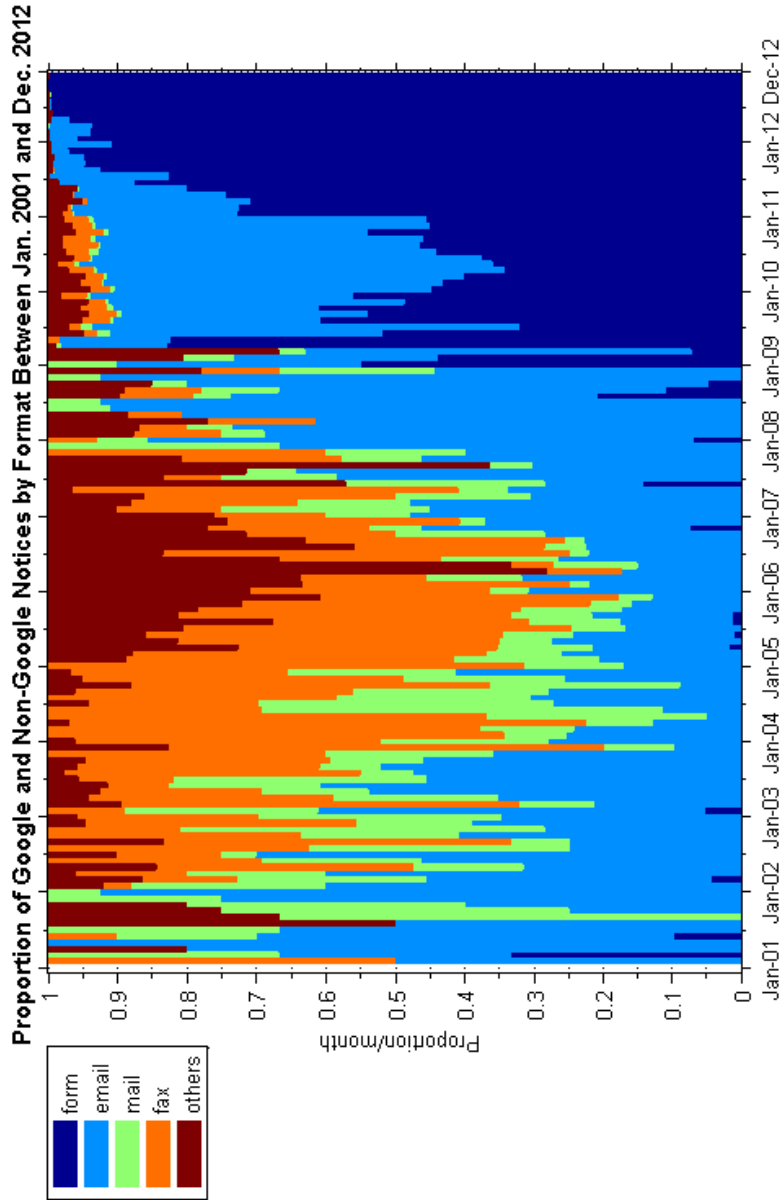


Figure 9: Number of Notices (Excluding Google) by Format Between 2001 and 2012

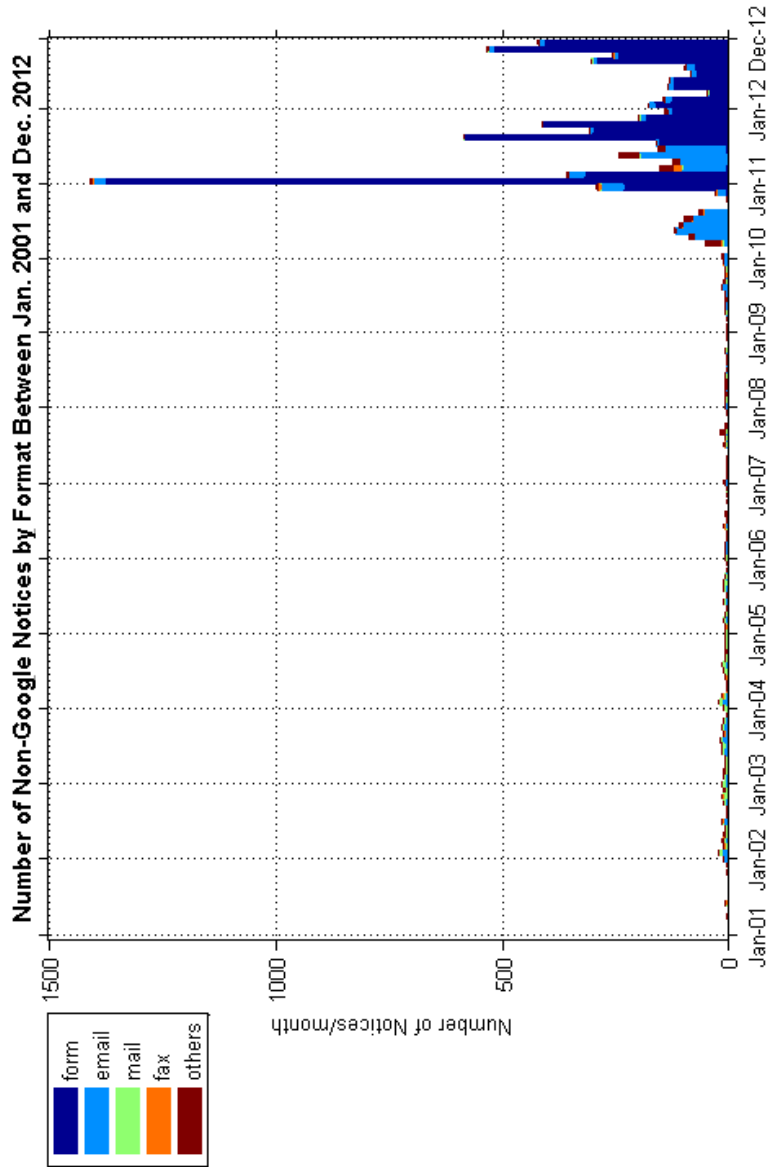


Figure 10: Proportion of Notices (Excluding Google) by Format Between 2001 and 2012

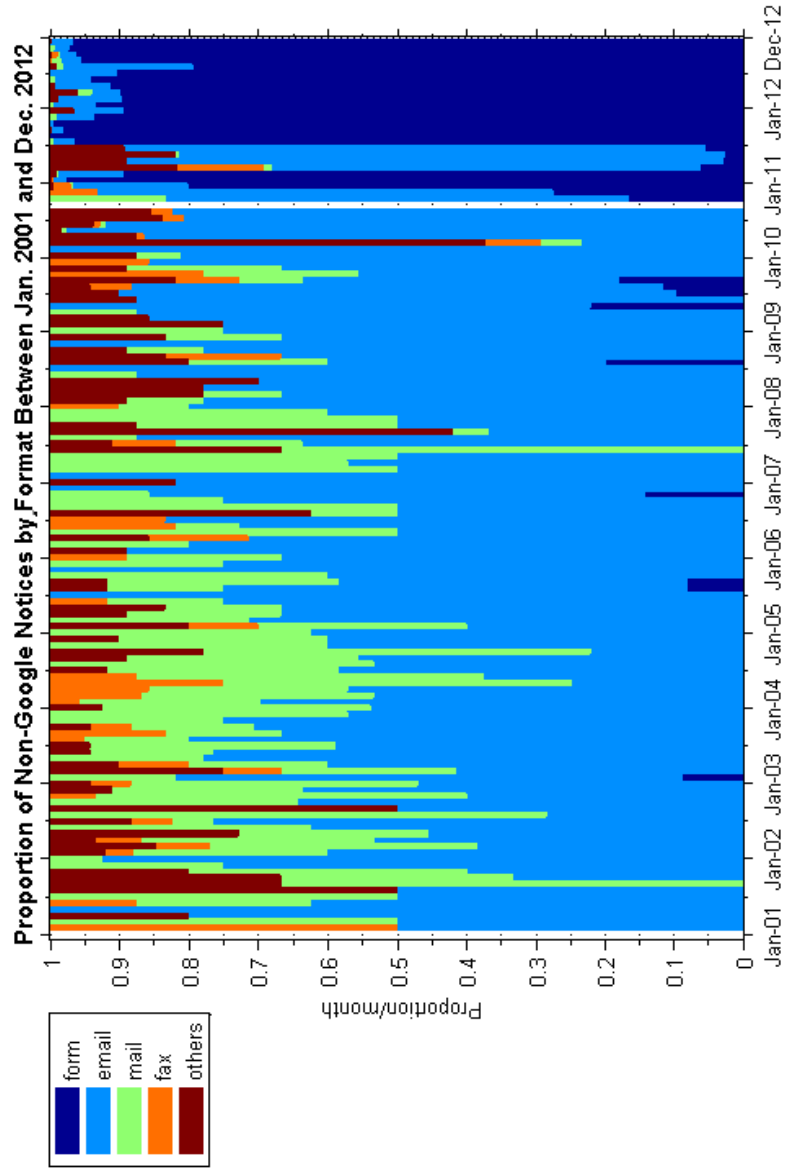


Figure 11: Breakdown of Requests by Industry for Top-Fifty Providers Between 2008 and 2012 (N₂=501,286)

