

THE UNIVERSITY OF CHICAGO

An Architecture of Control:

Automatic Takedown Notices in Academic Digital
Publishing

By
Isadora Borges Monroy

August 2014

A paper submitted in partial fulfillment of the requirements for
the Master of Arts degree in
the Master of Arts Program in the Social Sciences.

Faculty Advisor: Adrian Johns
Preceptor: Alexandra Bass

Table of Contents

Figures and tables index	1
Abstract.....	2
Introduction: takedown notices as architectures of control.....	2
1 Takedowns: the architecture	16
2 Google's execution: the control.....	22
3 The data's evolution: trends in Google takedowns.....	27
4 Academic publishing's use of takedown notices	45
5 The Making of the Digital Millennium Copyright Act.....	56
Conclusion: Sixteen years of the DMCA's architecture of control	65
Works Cited	69

Figures and tables index

<i>Figure 1 Monthly averages of URL takedown requests (M2011-May2014) in four years</i>	12
<i>Figure 2 Top 10 copyright owners by requests (2011-2013)</i>	29
<i>Table 1 Comparison of notice frequency by type of sender (2002-2005, 2011-2013) ____</i> 32	32
<i>Figure 3 Top 10 reporting agencies by requests (2012-2013)</i>	30
<i>Figure 4 Reporting agents' concentration by notices (2011-2013)</i>	33
<i>Figure 5 Top 30 reporting organizations' frequency of notices (2011-2013)</i>	33
<i>Table 2 Longitudinal histogram of URL removal (M2011-May2014) by reporting organizations</i>	34
<i>Table 3 Longitudinal histogram of URL removal (M2011-May2014) by copyright owners</i>	35
<i>Figure 6 Four years of request by the top 10 reporting organizations (M2011-May2014)</i>	36
<i>Figure 7 Total requests by year (M2011-May2014)</i>	37
<i>Figure 8 Top 50 sending content owners' request percentage (2011-2013)</i>	39
<i>Table 4 Top removals by industry categories (2011-2013)</i>	38
<i>Table 5 GTR longitudinal data (M2011-May2014)</i>	42
<i>Table 6 Publishing industry reporting agents' longitudinal request rates and success indexes</i>	43
<i>Figure 9 Histogram of requests by agencies and owners (2011-2013)</i>	45
<i>Table 7 Numeric percentage of sum of requests by top 100 industry categories</i>	46

Acknowledgements: I would like to thank my thesis advisor Professor Adrian Johns for a year of insightful discussions and his guidance in making my arguments stronger, to Professor John Kelly for his introduction to Lawrence Lessig's literature and the Science, Technology and Society field, my preceptor Alexandra Bass for going over this paper with tremendous dedication, and the MAPSS faculty and staff for a year's worth of invaluable experiences. Allison Kallo for helping me figure out how to work with 17 million rows of data, to Garrett and Devon for their insight into the publishing world, and to CONACYT.

For Zuraya, Guilherme, and Rebeca. I hope this one is easier to read than the last one.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Abstract

Before portable media, piracy in printed content was the bane of censoring states and publishers' returns on investment. Nowadays, digital content can easily be distributed outside authorized channels. However, the internet also affords surveillance and access control mechanisms that can disrupt the transmission or distribution of media, or hamper search engines' result indexes.

This thesis explores different media industries' use of "takedown notices", a digital anti-piracy tool which removes access to copyright infringing content by notifying service providers of a claim to ownership. Service providers must disrupt its access or risk liability for abetting infringement. In the tradition of *quis custodiet ipsos custodiet* (who will guard the guards?), this paper examines prominent gatekeeping entities who issue takedown requests and Google—the biggest receiver of search-based takedowns. I particularly analyze why some publishers deploy automated takedowns for academic authors and whether it serves these authors' best interests. I make the case for distinguishing their use in academic production from other media because academic authors have additional production incentives which commercial artists and their producers do not. Finally, I connect the open access debate open access to digital piracy of academic works in the hopes of reconciling academic values with piracy concerns.

Introduction: takedown notices as architectures of control

A careful examination of the anti-piracy history reveals private policing of infringement is a practice consistently adopted as public policy by states in both sides of the Atlantic ever since modern copyright became a legal protection. As certain trades become increasingly germane to some national economies, these states take it upon themselves to legitimize the private practices that connect trade associations through voluntary anti-piracy groups. The ensuing union can then enforce, with the state's legal support, the standards of copyright behavior they deem appropriate upon the general public which uses their protected goods. The increasing clout of these associations combines with processes of international power asymmetries and globalization. By power asymmetries I am alluding to how in international negotiations the most powerful countries, defined in economic and military terms (i.e. "hard power") are able to internationalize concepts and standards of copyright and intellectual property that best defend their leading national industries (Seville 2006, 62). Globalization, in turn, has helped define piracy in universal terms, preferring uniform legal definitions across states over deference to historic national legal standards.

This homogenizing effort has not been without its difficulties, as national customs, economic disparities, and issues of fairness, equality, and freedom of speech have been raised (Karaganis 2011).

However, as internet researchers like Lawrence Lessig (2006), Nate Anderson (2008) and Jessica Litman (2006) have found the internet is not an anonymous safe haven. Surveillance and control of the flow of information can be more effective than through the analog mediums that preceded the internet—to the concern of freedom of speech advocates, anti-systemic dissidents, and defenders of noncommercial “leisure exchange” (Rojek 2005). Until the digital age, piracy could only be controlled through the law, and depended solely upon on judicial oversight for its enforcement to be effective. This changed when two mechanisms were codified into American law in 1998 via the Digital Millennium Copyright Act (DMCA). Title I outlaws circumventing digital locks on content called digital rights management (DRM), and proscribes distributing a service or device that can frustrate access controls¹. Most research on the DMCA has focused on whether Title I’s anti-circumvention legislation has come at a cost to innovation (see for example, Tarleton Gillespie, *Wired Shut*). Less systematic work on Title II has been conducted. Title II adds the On-Line Copyright Infringement Liability Limitation Act (OCILLA) section (§512) to the 1976 Copyright Act (US Code Title 17, or USCA)². The provision sought to address responsibility and liability of online material that infringed on copyright and to enforce access controls via “takedowns”. There are four kinds of takedown requests. The first chapter will explain the differences although this thesis focuses on search-based takedowns.

A “takedown” notice is a document, usually sent electronically, where a copyright owner or a hired reporting agent contacts an online service provider to petition it to discontinue the unlawful access to its property. Each document is called a notice, and is composed of at least one claim (an item of property like

¹ The Audio Home Recording Act of 1992 pioneered the inclusion of a legal mandate to incorporate technological copyright protective measures like the Serial Copy Management System” (Gillespie 2007, 321). Title I of the DMCA replicated the spirit of the 1992 law.

² Therefore, this paper will use the terms OCILLA, Title II (of the DMCA), and section (symbolized as §) 512 as synonymous because their content is the same.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

a book title or album), and must point to at least one specific infringing link, or URL, per claim. Each URL is called a request, and must lead to the particular page with the claim's content. It cannot just point to a website like the now defunct www.megaupload.com; rather, it requires the full address that will take you to the particular piece of work. Service providers are required to respond in particular ways, depending on the service they provide, but they all must act towards disabling access to infringing content to the best of their ability. This thesis will outline in broad strokes the differences among service provider kinds, and focus on search-based takedowns as detailed in §512(d), which are tasked with removing from their search indexes results that lead to a request's URL.

This thesis seeks to gauge whether issuing automated takedown notices generates a tension or conflict of interests between publishers (even if they are non-profit) who contract out their copyright defense to agents using automated takedowns, and the interests of the academic authors they represent. An underlying premise of this work is that academic authors differ from trade authors and artists copyright also protects. I assert this because while academic authors are not exempt from worrying about their personal economies, they are embedded in a system of research and teaching, mainly through universities, where prestige is measured in citations, conference participation, tenure, etc. Scholars are not expected to become popular best-sellers, and their fortunes are not made or broken by how a monograph fares in the marketplace. Publication is certainly an essential part of an academic author's career and informs how far up the ivory tower they rise (impacting stipends, grants, fellowship offers, salaries, etcetera), but the success of her idea is not measured by how well it sold. Instead, the success of an idea is measured through quantitative impact indexes and qualitatively prestigious metrics (such as awards, being referred to as a leader in the field or having asserted a new concept or paradigm, etcetera).

Sometimes academic authors even produce journal articles at a cost to them by submitting to journals who charge to deliberate on publication. They do so not because knowledge is free (as in beer), but because their research institution pays them a salary to produce and transmit information and

knowledge. In turn, these research institutions expect their scholars to engage with their peers in a manner that implicitly requires thinking of their produced knowledge as a non-rival and non-exclusive goods. Academic authors do not compete for a reader's money the way mass market paperbacks compete at a convenience store. Thus, academic authors are set apart from other copyright subjects because their incentive to produce does not rely on the metric of commercial consumption (i.e. sales) which artists track.

In a sense, successful and efficient consumption of academic works means the adoption or re-appropriation of their innovative ideas by other researchers. So while some authors publish academic books that do make them money, the major incentive to produce these kinds of works are promotions, job offers, and university department titles which often translate into a higher salary than a publisher's advance and royalties would. These benefits are not exclusive for academic authors who become bestsellers—there are exceptions to the rule that academic authors do not make a fortune from their books alone, but these people still benefit from the same incentive structure most academic authors do when good ideas get passed around, contended and defended by their peers. The dynamics described above do not apply to other content media.

This distinction now established, the object of this thesis, to understand how the current state of automated takedown notices in academic authors' work impacts their interest, stems from a concern that these notices are being sent out without the content creators' full understanding of the effect this tool may have on her own intellectual career. Some musicians who depend on sales have publically voiced discontent when their fans are taken to court by their record labels. If record labels or the reporting agents they employ are carrying out copyright enforcement without the full consent of the artist they represent³, they may be issuing takedown notices under the assumption that all artists would want this protection. Likewise, academic publishers may be sending out takedown notices for authors who do not want to be

³ Under §512(c)(3)(v) and (d)(2) the complaining party must in good faith be under the "belief that the use of material in the manner complained of is not authorized by the copyright owner, its agent, or the law."

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

represented thus, or who do not know their work is being subjected to a system of perfect access control. This thesis seeks to contribute to academic authors' and their publishers' understanding of the system in order for its enforcement to derive from the full and knowing consent of the author. If search-based takedowns, a preemptive measure to obstruct access to copyright infringing works, are being carried out with uninformed permissions, then this area needs further exploring in the courts and in academia, because copyright representatives may be at fault. Finally, I will associate the debate on scholarly journals' open access to the scholarship on digital piracy of academic works in the hopes of reconciling academic goals with piracy concerns.

In order to understand of how takedowns work, particularly as deployed by online service providers (OSPs), the thesis will apply Lawrence Lessig's concept of "architecture of control" to the language of section (§) 512 subsection (d) of the USCA Title 17, also known as Title II of the 1998 Digital Millennium Copyright Act. Law and digital technology have a common characteristic. They both require clear rules, or code, to operate in a precise fashion. Like the law, digital code tells a software, program or protocol what behavior to identify to, and how to go about reacting to it. For both law and digital technology, an overarching "architecture of control" (Lessig, *Code: version 2.0*. Online version 2006) sets the principles by which particular code is developed. In law, for example, where two rights or protections conflict an architecture of control builds a hierarchy of values that is applied to solve the dilemma. In an example of the right to privacy versus the government's right to investigate crime, the current legal architecture of control allows a person's privacy rights to be extinguished if law enforcement has obtained a warrant sanctioned by a judge to search his home and correspondence.

Conceptual alternatives for the study of technological regimes exist. The science anthropologist Bruno Latour studies "the making of the law". His focus, however, is on ethnographies of "enunciation regimes", studying the "knitting of legal reasoning" (Latour 2009, viii). This framework works well for those who, like him, have access to "behind the closed door where the cases were discussed" because legal

thinking (enunciation, when discussed out loud) is the making of the law. Latour's intent on the production of truth by jurists ("the essence of the law" (Latour 2009, x)) requires studying in person the practice of lawmaking. This endeavor, while interesting, is incompatible with the subject-matter for two reasons, ironically related to access. First, the negotiations, lobbying and "knitting of legal reasoning" already happened. An ethnography of a past event is history, not anthropology. Second, while the doors might have been open to him to study French administrative law, few makers and of American law and stakeholders of copyright law in particular would grant the same level of access to a researcher. For example, since 2013 the Executive's Trade Representative has been asking Congress for the power to negotiate without Senate ratification (and public debate) the Trans-Pacific Partnership with 11 other countries. The treaty deals in a range of commercial subjects, including intellectual property, but has "been conducted largely in secret" (DePillis 2013). If the production of the truth is Latour's objective, the subjects being studied have to be willing to act in a transparent manner.

The concept of "architecture of control", on the other hand, has the ability to aid in understanding not just how and why a law was made, but how and why it is actually being implemented in a certain way. Coined by Lawrence Lessig and explained in multiple essays like *The Laws of Cyberspace* (1988) and *The Architecture of Innovation* (2001), it is especially developed in his book *Code and other Laws of Cyberspace* (1999), later updated as *Code 2.0* (2009). The concept arises from the notion that offline and online behavior is regulated by constraints: laws, social norms (communal understandings and shared expectations), the market (regulating through price), and "what some might call nature, but which I want to call 'architecture'", "the world as I find it" (Lessig, *The Laws of Cyberspace* 1988, 3). Computers execute through code the standards and protocols (like laws do) of the digital world. In creating these rules, coders build an architecture that defines the terms in which we inhabit digital space (cyberspace, in slightly dated terms). His objective was to rebut the idea that cyberspace is inherently free, since regulation through code is part of the way computers work—and can thus be transformed into regulation for control. Regulation for

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

freedom is possible, but it has to be a deliberate policy choice, since “control versus freedom will be the debate of the twenty-first century” (Lessig, *The Architecture of Innovation* 2001, 1785). Thus, an architecture is a collection of structures and constraints that seek to protect fundamental values. In this case, the architecture leans towards control in an effort to protect the rights of content owners to copy.

The particular architecture of control laid out in the Digital Millennium Copyright Act of 1998 (DMCA) aspired to regulate a *sui generis* internet dubbed the “National Information Infrastructure” (NII) by laying out how it should be built in the future. The NII would afford a pay-per-use licensing system for media of all kinds. The legislators’ assumption was that content owners would only digitalize their media if they could be assured that their implementation of locks against content piracy would be secured through the law. The architecture of control they devised through the DMCA had the goal of creating a digital environment where copyright holders would be able to enforce a compulsory license system, which could even include micropayment structures suggested at the time (Nelson 2004). Furthermore, some parties fantasized, because digital technology might afford content owners perfect control of the use of their material, fair use and the second hand sale of content could one day be extinguished.

Constructing the DMCA required establishing legal protections to maintain the immutability of digital rights management (DRM). The legal protections derived from issuing takedown notices will be the subject of chapter 1 for users, Google, and copyright holders and their reporting agents as a means to understand their behavior in the application of takedowns and the implications of their high volume of use. The takedown mechanism was meant to combine with DRMs in order to remove any unauthorized, unprotected copies of content. It is worth noting that individual takedowns are designed to be executed without initial court review. This innovation in copyright disputes is meant to make takedowns a low-cost mechanism for content providers to make their claims. Only when an alleged infringer is notified of the takedown, which does not happen often for search takedowns can she have the option of filing a counter-notice. If this takes place, the service provider who carried out the takedown is legally required to inform the

copyright holder, who can then sue the alleged infringer. It is then that a judicial process, in the formal courtly sense, is triggered. Takedowns thus are a part of the private-handling tradition of copyright policing (Johns 2010), but they are different from other enforcement mechanisms because of the automated systems that are enforcing perfect control.

As discussed above, the DMCA's architecture aspires to perfect access control in order to make the digital market compatible with compulsory licensing. This is a second innovation because copyright law has traditionally not aspired to perfect control. Anti-piracy measures were previously targeted at making unlawful copying (at times limited to for-profit infringement) risky enough that the average person would not engage in such activities on a commercial level. Some critics of the DMCA think copyright law should go back to this architecture of control, incentivizing copyright holders to make their work cheaper and easier to get legally (Google 2013, 4) (Kelty 2012), instead of constructing an architecture which can technologically enforce a compulsory licensing system on every piece of data online (Lessig, Code: version 2.0. Online version 2006).

The thesis seeks to establish if there is a discernible trend in the use of takedown notices, and if the book industry has followed other content industry practices in search-based takedown requests. The rate of takedown requests, or success index, was derived from Google's Transparency Report (GTR). Since July 2011 Google has begun compiling into a Transparency Report⁴ the search-based takedown requests it receives from content owners and reporting agencies. The database logs every notice Google has received from March 2011 into a continually updated website⁵ and a downloadable .zip organized into three comma separated value files. Other Google services like Picasa and YouTube have their own reporting mechanism which does not count towards the GTR. Of the three GTR files, one displays information by domains, focusing on how many unique links each website gathered per notice. For

⁴ Google had also previously sent the notices it received to the Chilling Effects repository and continues to do so, adding a link to the complaint on Chilling Effects in its own compilation. See footnote 6.

⁵ The searchable website is <https://www.google.com/transparencyreport/>.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

example, if a notice contained 5 URLs from megavideo, Google condenses the information by stating that for the notice with an ID number x, they took down 3 URLs from megavideo, 1 requested URL resulted in no action taken, and 1 URL is pending action. The file also covers whether Google received a counter-notice for any of the reported links within the notice. A counter-notice is filed by the website's owner or a service's subscriber as a means to refute a claim of copyright infringement and request the file be put back online or listed again in a search.

The second file the GRT provides organizes takedown notices by requesting party. Google tracks the notice ID, the date and time of its submission, the corresponding Chilling Effects⁶ URL where the notice was archived under, the copyright owner's name and the ID Google gave it, the reporting organization name and ID if an agent is acting on a copyright owner's behalf, and a count of how many URLs were removed in regards to the particular notice, for how many Google took no action, how many URLs are pending review and if any of these takedowns elicited a counter-notice. The third file contains detailed information about the particular URLs where no action was taken, listing the complete URLs for each notice and whether a counter-notice was filed.

The first database was downloaded on December 13, 2013 and goes back to March 2011. It contains 292,275,017 links requested by 30,981 copyright owners and 27,213 reporting agents to be taken down⁷. This database (2011-2013) was used to conduct in depth analysis of the top reporting agencies and copyright holders. The instances of URLs removed, no action taken and action pending were added into a

⁶ Chilling Effects is a website which invites internet users (individual and corporate alike) who have received "cease and desist" letters or DMCA takedown notices, to submit them to their organization to make the request easily understood "in plain English" (Chilling Effects Clearinghouse 2014) in order to "support lawful online activity against the chill of unwarranted legal threats" (Chilling Effects Clearinghouse 2011). Initiated by Harvard's Berkman Center for Internet & Society, law clinics of diverse universities and the Electronic Frontier Foundation since 2002 have had a standing request for "cease and desists" triggered by online content, including DMCA takedown notices. As a consequence, these private documents are turned over to the public for their scrutiny. The collaborating universities are Harvard, Berkeley, Stanford, University of San Francisco, University of Maine, George Washington School of Law, and Santa Clara University School of Law (Chilling Effects Clearinghouse 2011).

⁷ Google has assigned 30,981 unique IDs for copyright holders and 27,213 unique IDs for reporting agencies in the 2011-2013 GTR. The numbers will not reflect if an entity has sought out multiple IDs or if they have switched from reporting themselves to hiring a reporting agency,

column labeled “sum attempts” field, and a success index was created (with the ratio removed URLs / sum attempts). The findings, detailed in chapter 3, were compared to the first takedown database article Urban and Quilter’s 2005 article, *Efficient Process or “Chilling Effects”? Takedown Notices under Section 512 of the Digital Millennium Copyright Act* and Daniel Seng’s *The State of the Discordant Union: An Empirical Analysis of DMCA Takedown Notices*. Their work helps contextualize more recent trends in search-based takedowns because they contain notices Google submitted to Chilling Effects before 2011. A second database downloaded from Google includes data from March 2011 to May 19, 2014 (M2011-May2014). It was used for longitudinal metrics as well as to verify that the sharper increase in search-based takedowns is a trend of the past three years.

Studying Google’s handling of search result takedowns, the subject of chapter 2, is a relevant exercise in understanding architectures of control because this company is currently such a significant gatekeeper to digital information and a powerful enforcer of the DMCA’s Title II. As Google has grown, so have its responsibilities under different subsections of the DMCA. But as a search engine, it is in charge of a large amount of §512(d) takedowns. Furthermore, while DRMs still bind most e-books, publishers and other content owners are increasingly issuing takedowns (see Figure 1). Search-based takedowns have become attractive to copyright holders as the concentrated use of Google searches has created an effective set of conditions to cut off access to casual downloaders. By dividing the Google Transparency Report’s yearly data from March 2011 to May 2014 for all industries into monthly averages for search-based takedown requests, we can see a clear and growing trend to flag URLs, and Google’s accompanying removal of the requested links from their search results. The particular growth in the deployment of takedowns by the book industry will be detailed in chapter 4.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

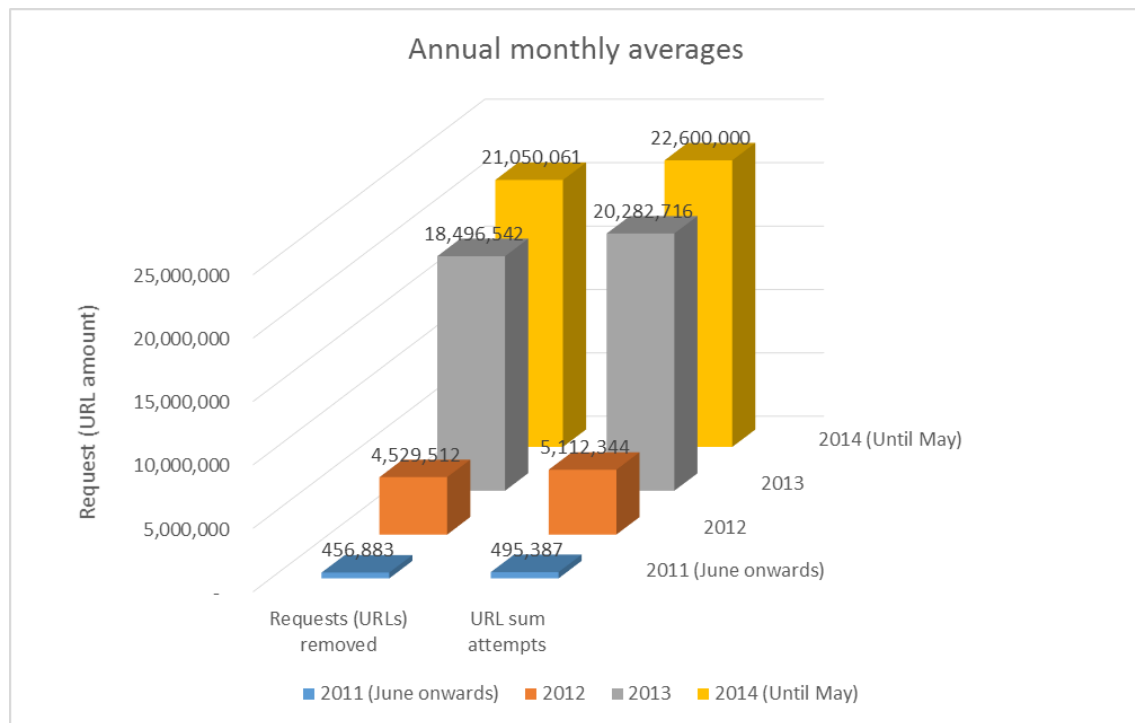


Figure 1 Monthly averages of URL takedown requests (M2011-May2014) in four years

The first reason why this thesis is concentrated on Google is that although this corporation shows results to and from many countries, as an American company it is beholden to following American law. The result is that searching for content and delisting it according to the 1998 DMCA effectively internationalizes American copyright standards. This is a recurrent dynamic in copyright history, where the economically stronger states drive the international agenda on copyright to secure their technology and production advantages by homogenizing intellectual property norms to reflect their national standards and interests in other countries. For example, the European Union's E-Commerce Directive adopted in 2000 imitates the US' safe harbor provisions for OSPs that take down content accused of infringement (Google 2013, 14). In the meantime, because the practice of issuing takedowns is modeled to happen without triggering a court review⁸. Furthermore, most search-based takedown notices do not elicit a counter-notice from an uploader

⁸ Discussion over how to describe the takedown practice underlies the two big takedown papers. Urban and Quilter refer to it as an extra-judicial process (Urban and Quilter 2005, 623, 634, 635, 639), Daniel Seng discusses it as bypassing judicial oversight (Seng 2014, 3). The ambiguity in terminology relies on the fact that while the general process is legally sanctioned through the DMCA, only when some disputes end up in court as a product of getting sued over placing a counter-notice, is the particular

whose material is delisted from Google because there is no obligation for Google to find her and inform her of the request they received. These two processes normalize the practice of privately issuing and enforcing takedowns. Because content owners have come to expect takedowns to be an uncontested part of American copyright, the government has sought to introduce the mechanism in international trade agreements (Seng 2014, 2, 3), (Litman 2006).

This process has been described as *policy creep*, since almost imperceptible changes in copyright, enforced through technological measures “eventually mature into the policy equivalent of a *fait accompli*” (Neave 2012, 19). One example of policy creep is the DMCA’s requirement in that ISPs (and arguably, subscription-based OSPs) develop policies to terminate repeat infringers (§512(i)(1)). The concept has been attempted as standalone legislation through different numbered “strikes” laws, or “graduated response”, but has faced popular opposition and legal obstacles from countries where internet access is considered a right and not a privilege, or where identifiable information is questionably obtained (Anderson 2008). However, the legal groundwork for termination of internet services exists in the DMCA’s Title II and it can become standard practice for private companies to enforce in the United States without fear of being sued⁹.

claim judicially scrutinized. But individual takedowns are designed to be executed without initial court review, as a low-cost mechanism for content providers to make their claims and service providers to collaborate. Only when an alleged infringer is notified of the takedown (which does not happen often for search takedowns) and is *then* sued by the copyright holder, does a judicial process (in the formal courtly sense) get triggered. Thus, calling it “extra-judicial” is a quick and offhand way to allude to this complex set of circumstances. But while takedowns take from the private-handling tradition of copyright policing (Johns 2010), we must not forget that it is still within the scope of private *law*, publically discussed and passed by the American Congress (despite content owner’s heavy hand in its redaction), signed by the Executive and upheld by court cases of all levels, and thus publically sanctioned by the State.

⁹ Great Britain’s implementation of a warning program for peer-to-peer file-sharing called Vcap is an example of how policy creep internationalizes American copyright standards. The Intellectual Property advisor to the Prime Minister believes a purely educational program that targets casual users, and not more unrelenting file-sharers will be a failure of the program (Ernesto, UK ISPs and copyright holders praise new piracy warning system 2014). However, the government is so far not requiring ISPs to disconnect users after four warnings. This suggests that expectations of privacy and the right to internet access abroad are coming at a heads with the United States’ favored tactics. Monitoring is nonetheless expected to be carried out by an American company (MarkMonitor, one of the largest Google-reporting agencies) (Ernesto, This is How the UK Piracy Warnings Will Work 2014) which may help enforce American standards of liability.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

The second reason why this work concentrates on Google is that this service provider is the most used search service provider (Meyer 2013) (Sullivan February), and has been so for over a decade (Google 2000). For example, in August 2012 Google publically announced that it had been filtering and removing from its instant search service terms related to torrent downloading since 2011. While Google framed the matter as unrelated to censorship and as a distinct measure from the DMCA takedown notices, the process meant no longer suggesting cyber-locker domains in their auto-complete search bar, which significantly changed search results (Ernesto, Google Now Censors the Pirate Bay, IsoHunt, 4Shared and More 2011). Even though the infringing content may exist, removing from auto-complete and taking down Google search results can be as effective as getting the content pulled down from the actual hosting site because all but the most dedicated users rely on this search engine to be directed to allegedly infringing content.

Google's grip over users' habits is so pervasive some blogs have suggested just using a different search engine like Bing can be enough to pull up pirated content Google filters out. Severing the link between searches and accused content can thus be an effective bottleneck for some consumers of dubiously legal content. Furthermore, the low rate of redundant searches via other indexing websites is enough to keep issuing automated takedowns at a low cost. In this sense, Google's large market presence in online searches works in favor of an architecture of control that, unlike previous copyright law, seeks to stop *all* unauthorized reproductions.

A third reason why Google's practices should be scrutinized is because after the DMCA was passed, Google takedowns have functioned mostly outside the scope of policy discussions of intellectual property. While private law is pursued between private parties, it is still legitimized and ultimately enforced by the state. So for example if two parties enter into a contract and one party breaches it, the other takes the offending party to *public* court. But takedown legislation does not require notifying the uploader that his or her content has been pulled from search results. The effect is that while content and its availability via

search results are being continually and increasingly altered by those alleging copyright infringement, public discussion of its effect has been left to few academics because most internet users simply do not perceive the removal of content or of its indexing. This is unlike the DMCA's mechanism required for service providers who store information from subscribers, like YouTube or cyber-lockers (§512(c)) and must notify their subscribers of the takedown. Indexing removals are also less visible than content removals because Google informs the user that a result has *not* been displayed at the bottom of their results. Content removals, on the other hand, maintain the particular website up and once a user is connected to it and expectant consume the content she is then notified of the material's unavailability.

Thus, as the practice of search-based takedowns goes unchallenged it may become cemented as public policy and creep onto new proposed legislations in other countries as a result of bilateral agreements (Seng 2014, 1, 2) and proposed international treaties like the Trans Pacific Partnership¹⁰. Yet the onus of removing links from searches could be attenuated by moving closer to the "don't shoot the messenger" model of fully waived liability the Post Office enjoys when distributing packages, as suggested by research by Yen, Lemnley & Reese, and Urban & Quilter (Urban and Quilter 2005, 690). Instead lobbyists in 2012 sought to renegotiate the balance of work and liability OSPs have through the DMCA. Once their legislative proposals¹¹ failed, reporting parties were unable to transfer their policing onto OSPs and began using the existing takedown mechanisms in the DMCA's Title II at unprecedented levels (Masnick 2012).

As a means to understand the present digital policing structure of copyright infringement we require also looking into the process that led to the Digital Millennium Act. The historical analysis of Congress' legislative intent in chapter 5 helps explain *why* the architecture of control they crafted aspired to build a market with perfect access control. This extends to considering the motivations of diverging interest groups

¹⁰ See Mike Palmedo, *Mexican Economy Minister on TPP Provisions on Copyright Term, Camcording, and Notice-and-Takedown*, in the Program on Information Justice and Intellectual Property blog's <http://infojustice.org/archives/32001> (accessed March 30, 2014) for current opposition to the takedown mechanism abroad.

¹¹ Stop Online Piracy Act (SOPA) and the Preventing Real Online Threats to Economic Creativity and Theft of Intellectual Property Act (PIPA).

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

lobbying the lawmakers¹² and the process of negotiation within Congress' two chambers as a means of explaining the particularities of the DMCA's architecture of control. The brief history will help contextualize how takedowns institutionalize a relationship of trust between the reporting parties (content owners and their representatives) and the access gatekeepers (Google in this paper's case). It will also help explain the proliferation of reporting agencies, which in turn allows us to understand why academic publishers have hired them to issue automated takedown requests.

1 Takedowns: the architecture

Since the Digital Copyright Millennium Act was passed as US law in 1998 to amend the US Copyright Act of 1976, online copyright infringement has been dealt with through "takedown notices", as section 512 of the USCA now describes. The takedown mechanism compromise between copyright holders, ISPs and OSPs¹³ will now be described, in order to gauge how the law constructs private copyright policing. When someone uploads content whose copyright is contended for by a second party, it may be uploaded to a cyber-locker, embedded via a website's servers, or exchanged through a Peer-to-Peer (P2P) network. In

¹² Copyright owners, particularly artistic associations like the Recording Industry Association of America and the Motion Picture Association of America have had mixed results with other legal mechanisms like suing individuals, mass action lawsuits and digital rights management (DRM) locks on devices in the majority of media. DRM has tended to be more successful in e-books because it is so heavily tied to the way the material is consumed, downloaded and distributed in integrated platforms so that the applications are heavily tied to the devices that can display them. iTunes, for example, first drew criticism when users wanted to be able to listen to their music on non-Apple devices, their cars for example, and found out DRM was keeping them from copying the music they had bought onto a CD. The largest e-book sellers have preempted this issue, and taken advantage of advances in smartphone technology and mobile internet by making sure their applications are available on multiple platforms so that one can read the material anywhere like one can a paper book. This ubiquity does not equate into the same unrestricted access a paper book affords. Some vendors do not allow highlighting for the same book others do, some do not allow citing directly because this would require copying and pasting the material, and not all applications can produce the purchased copy in all devices, so that switching from your computer to your phone can be outright impossible, particularly if using smaller e-book vendors' applications.

¹³ The DMCA defines Title II's obligations in terms of "service providers", the collective for "a provider of online services or network access" (§512(k)(1)(B)). Urban and Quilter's 2005 article and the ensuing literature takes pains to distinguish among service provider types because each of the law's subsections applies differently for internet service providers (a company which provides internet access), and online service providers (an internet company that provides services from storing documents on the cloud (Dropbox), to connecting a customer to a cab driver (Uber), to showing cat videos (YouTube) to listing any of the aforementioned companies in a search (Google, Bing, etc.)). This paper will maintain the literature's tradition for rigor's sake, as the legislation makes important, material distinctions based on the diverging functions and capacities of internet and online service providers. For example ISPs are liable for carrying out the cut-off provisions of repeating infringers that OSPs cannot enforce (§512(a)), while ISPs do not host content OSPs can be held liable for if they do not take them down. ISPs are granted "blanket safe harbor" (Urban and Quilter 2005, 690) for their routing, transmission and storage activities, unlike OSPs. Categorizing OSPs as distinct service providers from ISPs helps refine our understanding of the responsibilities of each.

the first case, the cyber-locker (an OSP) stores the material (§512(b)(1)), in the second, the website (another OSP) hosts the content, and in the third the uploader has sole control over the material's availability. The law will make the cyber-locker and the website liable for monetary, injunctive or equitable relief unless the OSP complies with the conditions of §512. This subsection curtails liability for parties running "system caching and copy retention services", meaning those OSPs who require storing copies of content to increase downloading speeds "for a limited time, [...] so that subsequent requests for the same material can be fulfilled by transmitting the retained copy, rather than retrieving the material from the original source on the network" (US Library of Congress 1998, 10). In order to manage bandwidth efficiently, websites who are often visited maintain cached copies of popular content in order to load it at a faster rate. A cyber-locker, would have a hard case proving their services fall within the caching restrictions unless they follow the other rules including limiting access to materials through password protection set by the poster (US Library of Congress 1998, 11).

Before copyright holders reach the accused user, they will want to remove access to the content from the internet. This can either be petitioned for directly to the hosting OSP, and/or indirectly by removing indexed search results of it. Subsection (c) explains out how to carry the procedure out, creating incentives for the "expeditious" removal or disabling of access to the material (subsection(c)(1)(C)). In this section, both cyber-lockers and other hosting websites are in the same situation as indexing OSPs like Google or torrent listing websites like The Pirate Bay. However, the outcome of a §512(c) request is substantially different than a §512(d) request. The former removes the content from the internet, although it can leave particular (empty) page online, indicating to whomever reaches it that the content used to be there, but is no longer shown because of a copyright dispute. The latter removes the search result from a query, but does not tackle the online existence of possibly infringing material. An integrated approach to copyright

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

removals would include both kinds of takedowns, decreasing “linkrot”¹⁴, and removing the infringing content from the internet.

The compromise between copyright holders and OSPs legislated through the DMCA requires an inherent attribution of “good faith” from copyright holders that OSPs are from the onset ignorant of infringement because OSPs are willing to make it easy for copyright holders to send DMCA complaints to them (subsection(c)(2)), and promise to act expeditiously in resolving their request—although no quantifiable parameter is laid out in the law as to what constitutes “expeditious” compliance. This is what contributes to their safe harbor from initial liability. Furthermore, the courts have tended to side with OSPs in the matter of prior ignorance of copyright infringement when they do not incentivize users to share files and when they must deal with a great volume of content being disseminated through their network (Westley 2014). While this particular paper does not deal with the truthfulness, representativeness or specificity of individual takedown notices (unlike Urban and Quilter 2005), it is worth pointing out that if any of the parties abuses the good faith representation (including purported representatives of copyright owners’ interests), and material is disabled or removed because of a materially misrepresentative request, they may be liable for damages and attorneys’ fees (subsection (f)). Google has disabled two reporting agents for repeated misleading requests (Google 2013, 18), and Microsoft faced the inconvenience of having its own hired help suggest taking down Microsoft’s own websites (Ernesto, Microsoft Ditches Anti-Piracy Partner after Embarrassing DMCA Takedowns 2013). Previous works and present empirical data shows, however, that this provision may be underused because takedown executions do not require notification to alleged infringers who would put it upon themselves to make sure the request was fully compliant (Seng 2014, 53).

A specific list of the infringing work must be clearly identified by the reporting party with “reasonably sufficient” information so that the OSP may locate it, and remove the material itself or access to it ((c)(3)(iii)). Any request that does not comply with such details can be disregarded by OSPs because it

¹⁴ The existence of listed links that lead to expired content.

allows them to keep claiming ignorance of the specific copyright infringement allegation ((c)(3)(B)(i)) unless the request includes all but the exact URL information ((c)(3)(B)(ii)). This means that if a notice is incomplete regarding an OSP's ability to execute their takedown request, but the copyright owner or representative has made his or herself available to the OSP, the OSP is legally required to reach out and clarify the matter, a requirement that grants a lot of accommodation to the claiming party. Only if the OSP promptly attempts to contact the notifying party can it remain under compliance for non-monetary liability (((c)(3)(A)). Thus, a provision to state that the information is accurate, including a statement that the complaining party is acting in good faith on behalf of a disgruntled copyright owner, is required in all takedowns, but a great onus is placed on OSPs to clear themselves of any wrongdoing when suspicion is triggered by copyright owners' or representatives' complaints.

Subsection (c) and (d) of §512 applies to OSPs that host and link to content (called "information location tools"), respectively. Subsection (d) states that once an indexing or linking OSPs is no longer ignorant of a particular copyright infringement, they can maintain their status as not liable for relief if they act expeditiously to remove or disable access to the material. Linking OSPs are further protected from liability from the uploader if they take down material that is later deemed not to be infringing (subsection (g)). In contrast, subscription-based OSPs can be held liable by the uploader unless they notify that subscriber that they are removing or disabling access to their material (subsection (d)(2)(A)) or if access remains down while the copyright owner takes the subscriber to court (subsection (d)(2)(C)). Google's liability was decided in the 2010 case *Perfect 10, Inc. v. Google* (Seng 2014, 47, footnote 162), where it was determined that its search and caching services are not subscription based, and hence Google is not required to implement repeat infringer policies. By this logic they are also not required to issue notifications of requests to the uploaders they are showing search results of. In contrast, the OSP has the responsibility to notify the party requesting a takedown that it has received a counter-notification from their subscriber

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

(subsection (d)(2)(B)), so that it may decide whether to sue the subscriber or not. Only a response to a takedown challenge prompts the content to go back up.

We can conclude that the DMCA's architecture of control is inherently distrustful of uploaders. It also adds pressure on OSPs and ISPs to be receptive to any entity that claims copyright ownership of the flagged content, or which claims to represent the copyright holder. The law's systemic preference for content holders' claims over individual uploaders skews copyright allegations towards the biggest media companies and incentivizes the growth of a reporting agency industry. The pressure is evident in the subsections that outline how OSPs must reply to the complaining party within certain time limits and must keep in constant communication if a request is subjected to a counter-notice. Google is more likely to divert its takedown resources towards companies submit a high-volume of requests. Their Trusted Copyright Removal Program, which streamlines the process for the biggest requesters, seems to suggest the law's imbalance has real consequences on the execution of access controls. Chapter 5 will allow us to see that this asymmetry is no accident, but rather a direct reflection of big content owners' capacity to lobby Congress as a coherent group.

Furthermore, subsection (h) turns this private law mechanism into a public affair by providing a procedure whereby a copyright owner can obtain from a federal court a subpoena to compel a service provider to disclose the identity of an alleged infringer, which must be complied with "expeditiously". The section probably refers to an ISP, but this distinction is not made, so a subscriber of a cyber-locker could be intercepted via the OSP. Subsection (m) nonetheless states that these service providers should not monitor their users, as this would still be in violation of the provisions of the Electronic Communications Privacy Act (US Library of Congress 1998, 9). Subsection (i)(B) does, however, force service providers to accommodate and not interfere with third parties'—copyright owners—use of whatever "standard technical measures" it may develop to "identify or protect copyrighted works" (US Library of Congress 1998, 9). In turn, the definition of "standard" for these technical measures is defined in terms of what has been deemed

appropriate through “broad consensus” (one assumes informally since no mechanism otherwise is stated) between copyright owners and service providers. While stipulations of an “open, fair and voluntary multi-industry process” is a stated value in the technological measures produced to identify copyrighted works, no objection mechanisms or punishment for breaching these values is stated.

So far, none of the literature on takedown notices has addressed this amalgamation of DMCA provisions: digital rights management surveillance *combined* with takedown notices, as a joint anti-pirating effort. Subsection (i) is crucial for the subpoenas mentioned in subsection (h) to be technically feasible, but has received minimal, if any, scrutiny as compared to the attention spent on the demands for payment and trials this legal but private tracking lead to.

To summarize the foregoing information, before turning to the analysis of Google’s role in takedown notices, I will condense the DMCA’s liability tiers depending on each service provider’s range of activities. For service providers, subsection (a) applies to ISPs and subsections (b), (c) and (d) to OSPs. Within OSPs, subsection (b)’s caching rules are part of Google’s services, occurring at a rate of “once every few seconds on average” (Google 2014). Takedowns of search results remove access to cached copies. Subsection (c) concerns OSPs which house information on their systems or networks, such as cyber-lockers or other hosting websites. In this regard, Google services like Blogger can be considered potentially liable if they don’t follow the procedures laid out in the subsection. According to Daniel Seng’s research through the Chilling Effects database, in 2012 Picasa, Google Plus (its social network), and Blogger received 5,222, 3,213 and 58,052 requests respectively (Seng 2014, 44). Subsection (d) of §512 deals with limitations for information location tools—indexing and/or searching of hyperlinks. As with the other provisions, an OSP like Google must be ignorant of infringement, must not receive financial benefits from copyright infringement if it is “directly attributable to the activity” (US Library of Congress 1998, 13), and must block access to the material (removal is also stated, but is in practice outside the scope of indexing OSPs). Unlike other subsections, because their services are unrelated to a subscription base

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

these OSPs do not have to keep track of who is accused of infringement to disable their account after an amount of warnings or “strikes” are issued, nor do they have to notify users that material they may have posted is deemed infringing of copyrights by a purported owner.

2 Google’s execution: the control

As a result of the DMCA’s Title II and Google’s rising position as information gatekeeper, the company has begun receiving a number of §512(d) requests on a regular basis (Google 2014) starting in 2009 (Seng 2014, 11 Table 1). Google’s Frequently Asked Questions (FAQ) website explains copyright in terms of the DMCA, and recommends the use of their web form as opposed to fax, email or written notices, which makes it easier for the company to apply automated processes to take down accused links. As mentioned in the introduction, each document Google receives is called a notice, and Google marks each with a unique request ID. Each request has at least one item (a claim) of allegedly infringing material found online. Seng’s longitudinal data shows that until 2011 there was a 1 to 1 relationship between claims and notices (Seng 2014, 31). Each claim must contain at least one URL (a request) which the reporting party wants to be taken down from Google’s search services. There also seems to be an exponential growth in requests per claim (Seng 2014, 31). Per §512(c)(3)(B)(i) complaining parties sending notices can be the copyright owners or an intermediary representative or agent who is authorized by the copyright owner to act on its behalf. Either one of these entities can qualify to become a trusted partner in Google’s Trusted Copyright Removal Program (see further on). Seng’s longitudinal study of Chilling Effects reveals an increase in reporting agents (Seng 2014, 17), a position that was nonexistent in the takedown notice records before 2005 (Seng 2014, 20). Not only have their numbers swelled, but their participation has too, as a percent of total requests and notices, as the figure 6 mosaic shows.

As takedown notices have become increasingly deployed, Google has developed its own automated processes to issue and deal with the high-volume of requests. One can comprehend Google’s

implementation of a web based form as a way to help copyright owners and their agents to submit easy to execute requests as non-compliance can risk OSPs “contributory copyright infringement” (Urban and Quilter 2005, 626) as a consequence of not taking down material “expeditiously”. As a consequence, copyright owners and their agents are themselves generating automated requests with web crawling technologies (MarkMonitor AntiPiracy 2014) (Remove Your Media Clients 2013). This sets off a chain reaction which has established reporters’ expectations of fast response times, pressuring Google to that effect at the expense of a cautioned review of petitions.

Section 512(d) search notices have increased dramatically because content owners and their representatives are using Google Search as a first step to locate allegedly infringing material. The benefit to such a strategy is twofold. Websites used for repeatedly infringing posts are indexed lower in search results, and are deemed as untrustworthy sources as a whole. But, as mentioned in chapter 1, implementing search result takedowns bottlenecks information but does not actually remove access to it. For example, someone may find a link delisted by Google while browsing a forum, or may hear of other indexing websites from another user. As complaining parties have come to depend on issuing automated search-based takedown notices, their algorithms will still find the content online but when they report it to Google, they may find the notice goes unattended to. However, the URLs will have already been delisted from search results because of a previous notice. After the sharp 2013 increase (see figures 1 and 7) many reporting agencies are finding their notices rebuffed by Google (Andy 2014) not because the OSP is unwilling to act on them, but because these requests have already been delisted. Google has decided to mark the duplicate requests as “no action taken” instead of “URL removed”, which is lowering the measure for success index.

When Quilter and Urban conducted their study on individual notices they concluded that OSPs receiving them applied little scrutiny to the requests, which led to takedowns of dubious merit. Verifying the persistence of poorly executed requests is now a harder feat, as the March 2011-December 2013 Google

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Transparency Report contemplates 912,659 notices which request to take down 292,275,017 links. Google reported that between July to and December 2011, it removed 97% of all URLs flagged in the requests they received (Google 2011, How many of these requests did you comply with?), and their response time was six hours (Google 2011, How quickly do you remove search results after a request is made?). Considering the volume and the speed, its compliance requires automation (Google 2013, 17). Seng's anecdotal evidence suggests preferential treatment is given to electronic notices, which may well be a logical response of OSPs to automated detection and large batch requests (Seng 2014, 22). Google has established this practice because it makes it possible to process millions of weekly demands through its own automated processes, purportedly with engineering and machine learning tools to filter inaccurate requests (Google 2013, 18).

Yet, automated takedown notices pose a dilemma to the goals of American intellectual property theory. According to scholars of digital technologies like Lawrence Lessig, automated takedowns are “latent ambiguities” because the technological fix endows copyright holders with implied perfect control. He argues, “[a]s the Supreme Court has noted, copyright ‘protection has never accorded the copyright owner complete control over all possible uses of his work’”. Fair use, the first sale doctrine and terms limited to reasonable time spans are “important limitations”, and establishing a violation of copyright required legal (in the courtly sense) procedures. This school of thought contends that “the law was never designed to simply do the author’s bidding. It had public purposes as well as the author’s interest in mind” (Lessig, *Code*: version 2.0. Online version 2006, 179). While fair use can come from implied assent— when a copyright owner would be likely to authorize use, but seeking it would be too great a trouble— it is also derived from enforced consent, because the value of free critical speech also overrides an author’s objection to critique or parody the original work¹⁵.

¹⁵ The particular terms implied assent and enforced consent are Professor Latman’s (Litman 2006, 84).

Web crawlers and the asymmetric power held by reporting parties in copyright disputes are part of an architecture of control that supersedes the enforcement capacities imagined in the constitution. Copyright tradition did not aim at shutting down every unauthorized mean of distribution, nor did it seek to find and destroy every infringing copy. The current copyright legislation was informed by the desire to embed the internet with a compulsory and ubiquitous licensing system. As a means to do so an unprecedentedly low cost private monitoring and enforcement system was legalized. Authors like Niva Elkin Koren worry that this shift in power is not just a limit on our rights as consumers, or citizens for that matter, but that it also disrupts historic practices of consumption and symbolic meaning attachment to cultural artifacts like books or songs (Elkin Koren 2007, 1121).

Jennifer M. Urban and Laura Quilter's 2005 paper, *Efficient Process or Chilling Effects – Takedown Notices under Section 512 of the Digital Millennium Copyright Act* analyzed 876 notices (Urban and Quilter 2005, 641) individually systematically categorizing their effects on fair use and their deployment as anti-competition tools from the Chilling Effects database. The bulk, 743 notices, were submitted from Google since March 2002, and the rest were “non-Google” notices acquired by Chilling Effects through willing participants (Urban and Quilter 2005, 641). These researchers characterized the database's Google-centric data (84% of the dataset) as one of the two limitations of their study, although it was robust insofar as it had turned in data for a large period of time, providing “a complete picture of a major online service provider's interaction with §512” (Urban and Quilter 2005, 642). They construed a second limitation on their investigation, which was that the amount of “non-Google” notices was “relatively small and largely individually reported” (Urban and Quilter 2005, 642).

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Urban and Quilter's 2005 work¹⁶ is an interesting point of comparison to the earlier stages of the process because the low volume of takedowns made individual request analysis possible through the Chilling Effects clearinghouse project. At the moment, a similar effort to compile notices from big OSPs and individuals after over 700,000% percentage increase (Ernesto, Google Takedown Notices surge 711,887 Percent in Four Years 2014) makes the feat unrepeatable in a limited project like this one.

Daniel Seng's recent article *The State of a Discordant Union: An Empirical Analysis of DMCA Takedown Notice* provides deeper insight into the two aforementioned issues raised by Urban and Quilter. He characterizes the takedown scene after their 2005 study as "an explosion in the use of takedown notices" (Seng 2014, 3). He reached the conclusion that studying Google-centric search and indexing takedowns (§512(d)) is not a limitation, but an important endeavor, a statement I agree with. Seng's investigations suggests that copyright holders are targeting Google searches as a first-basis approach, in order to bottleneck access—whatever the underlying availability of pirated content actually is. This thesis works under Seng's proposition.

Seng's first dataset only six Facebook takedown notices are reported, although he too adds that "only a small number of service providers —Google and Twitter— consistently contribute their takedown notices to the Chilling Effects repository" (Seng 2014, 7). Seng proposes this Google-centric data may be function of a website's "maturation" as a service provider since Google only received its 1,000th takedown notice in 2009 (Seng 2014, 10). Twitter's transparency report is from 2012 onwards, although it was founded in 2006. Its reception of takedowns jumped from 307 in 2010 to 4,138 in 2011 and 6,464 in 2012 and 12,433 in 2013 (Twitter 2014). Whether this "maturation" theory is true or not, the numbers put out by Google are certainly worth scrutinizing.

¹⁶ See Jennifer M. Urban and 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 (2005), Available at: <http://scholarship.law.berkeley.edu/facpubs/501>

The methodology in Daniel Seng's article consisted of downloading the individual notices from the Chilling Effects repository and saving them to a MySQL database, creating a dataset of 539,558 notices. Almost 92% were notices submitted to Google through their web form. This again corroborates the centrality of §512 (d) as a highly deployed mechanism by copyright owners and their agents. The author addresses previous concerns over the representativeness of the takedowns from Chilling Effects because this entity relies on recipients of takedowns being willing to share their notices with them. Megaupload and Rapidshare, the "top two file hosting sites" (Seng 2014, 7) at one point had not provided any notices to the Chilling Effects repository. Yet Google's search-based takedowns can adequately paint a picture of the internet since they will contain references to the material these cyber-lockers index. It still remains difficult to gauge if copyright owners are also dealing with cyber-lockers directly or only seeking to remove search accessibility by delisting Google searches. Seng believes it is likely that a party concerned with copyright infringement would serve a takedown to search indexes (Google in particular), and the hosting site "to prevent the same content from being located" (Seng 2014, 8) by other means. As I have pointed out in chapter 1, this would be a more comprehensive approach to anti-piracy, since solely issuing §512(d) notices has the effect of bottlenecking information accessibility, but not the actual availability of it.

After reviewing Urban and Quilter's and Seng's articles, I have concluded there is merit in using Google's Transparency Reports. The reasons, extensively outlined in the introductory part of this thesis are compelling when using Lawrence Lessig's framework of "architectures of control" for copyright policy.

3 The data's evolution: trends in Google takedowns

As mentioned in the introduction, I obtained from Google's dedicated website¹⁷ two sets of the Transparency Report. The first database, referred to as (2011-2013) in the following review, contains 292,275,017 links requested by 30,981 copyright owners and 27,213 reporting agents to be taken down,

¹⁷ <https://www.google.com/transparencyreport/removals/copyright/data/>.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

from March 2011 to December 13, 2013. This is the main database I used for the analysis, and should be considered the database referred to throughout unless otherwise stated. A second database (referred to in shorthand as “M2011-May2014”) was downloaded on May 19, 2014 as a means to corroborate general longitudinal trends using 4 years’ worth of time-stamped requests.

I concatenated the GTR’s 2011-2013 domain file in order to show a summary of how many URLs had been removed, were pending action and where no action had been taken, as pertained to each website. This made over 17 million rows of URLs flagged in 912,659 notices more wieldy. The instances of URLs removed, no action taken and action pending were added into a column labeled “sum attempts”. A success index was created with the ratio of removed URLs / sum attempts. I also concatenated the GTR requests file from 2011-2013 to compare my findings on reporting agents, copyright owners and trusted partners with Seng’s data, in an attempt to determine whether the increased deployment of takedowns is a trend that has continued in a seemingly exponential manner. So far, this seems to be confirmed (see figure 7).

From the GTR 2011-2013 database there are 47 websites which have received over a million requests, and from this pool, the most requested (filestube.com) presents 10 times as many requests as the least (torrentdownloads.me), although the mean is 1,110.8712 for the 262,730 websites. With over 60 trillion addresses indexed (Google 2013, 14), such concentrated numbers do not suggest ubiquitous piracy as a share of all posted material. It does seem to point, however, to copyright owners and their representatives’ focused efforts on particular websites. For example, of the targeted websites, 3,884 had the terms listen, music, song or mp3; 14,361 terms like movie, stream, TV show, or watch. 4,155 links referred to software, play or game terms, 4,884 the terms xxx or porn, and 2,052 links for books. Yet the

most targeted websites are cyber-lockers, indexers of these and torrent websites¹⁸ which require a more sophisticated search system for petitioning parties to find their content and flag it.

Google's recent report, *How Google Fights Piracy* notes a "striking" trend, from 3 million URLs reported for all of 2011 to 4 million per week during 2012. Knowing this, with Seng's data (Seng 2014, 28, 31, 32) as a comparison, we can assume that for the top ten owners, the numbers prior to 2012 were comparatively negligible, as only Microsoft comes close to issuing more than half a million notices. Comparing and combining the two sources with the GTR 2011-2013 database the growth in the most prolific petitioners is apparent, more than doubling in all cases:

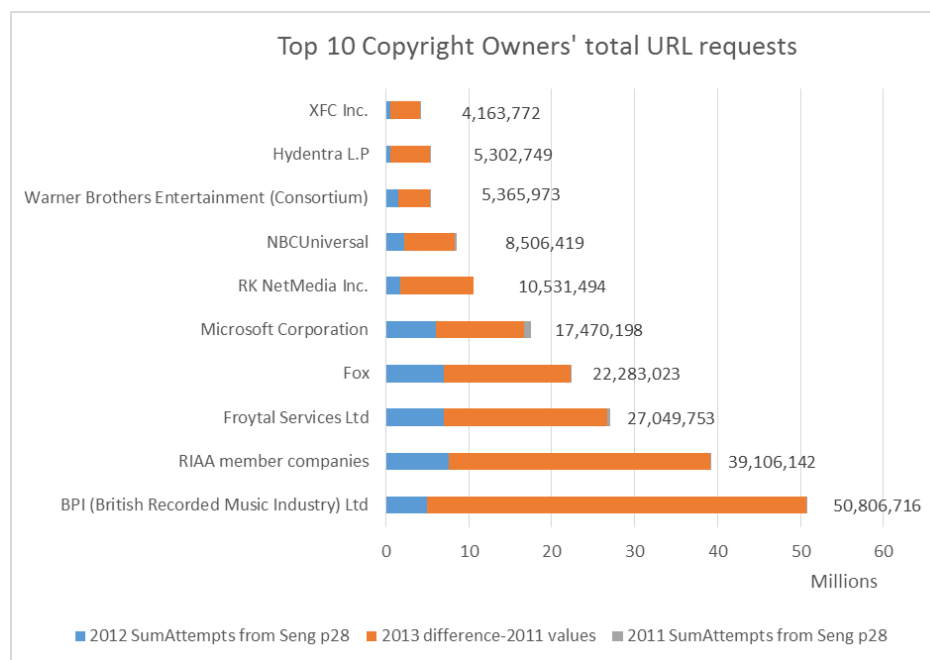


Figure 2 Top 10 copyright owners by requests (2011-2013)

Seng's longitudinal study of Chilling Effects reveals an increase in reporting agents (Seng 2014, 17) who act as intermediaries, authorized on behalf of copyright holders" (§512 (c)(3)(B)(i)), a position that was nonexistent in the takedown notice records before 2005 (Seng 2014, 20). Agents, enforcement agencies, and trade associations combined only accounted for 1.3% of § 512(d) notices prior to 2005

¹⁸ For example, the top 5 websites by requests received are filestube.com, zippyshare.com, dilandau.eu, rapidgator.net, and torrentz.eu.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

(Urban and Quilter 2005, 654). Their proliferation, as a professionalized effort to deploy search-based takedown notices is documented below. The longitudinal increase is also displayed using the M2011-May2014 database in a figure 6.

One note on the data is that duplicate entities do not mean double the notices, but rather that the entity calls itself both owner and agent when filling out the request form. In a larger dataset that concatenates the top 50 participant content owners and top 50 reporting agents, The Publishers Association shows 0% discrepancy in its filings, meaning that it has claimed both status in each petition. In BPI's case for example, 95% of the cases listing it as reporting agency also reported it as the copyright owner.

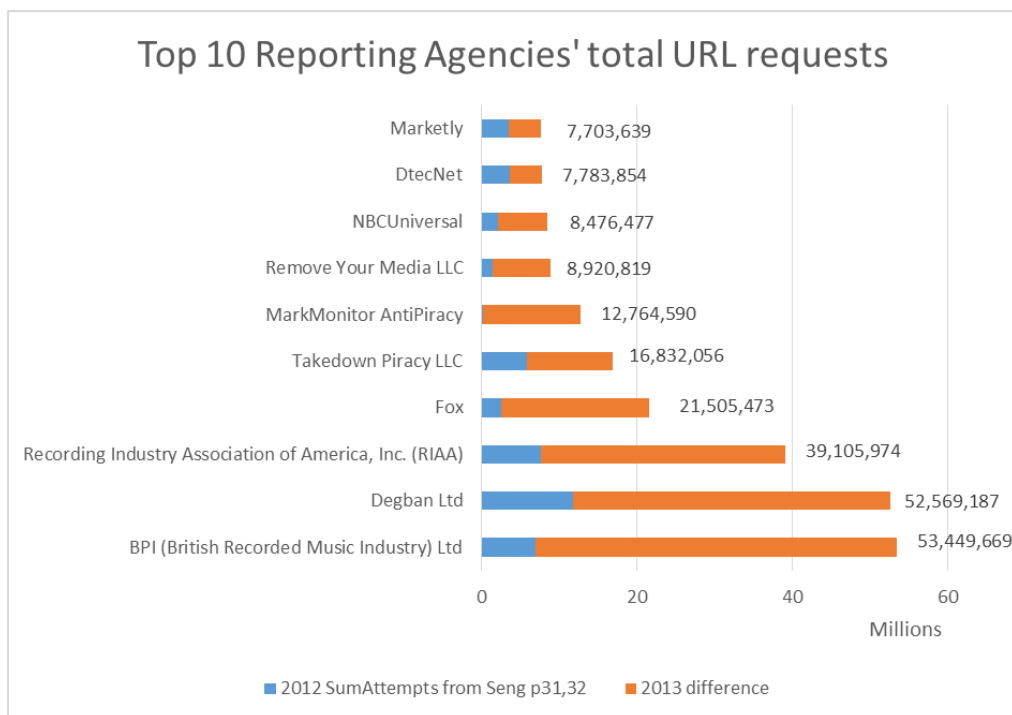


Figure 3 Top 10 reporting agencies by requests (2012-2013)

It is also worth noting that Google does not have any discretion for slight changes in agencies' names. For example, Degban, one of the most active senders of takedown requests, is listed with two different names (Degban and Degban Ltd) throughout the GTR. There is no way to link the two agencies as

the same entity automatically, as sometimes name changes are accompanied by changes in the agency's ID number. Even when the name is the same, by presenting two names in a stylistically different fashion (NBCUniversal and NBCUNIVERSAL) will result in sectioned data. This issue could be solved if Google linked their names to a single reporting ID and did not issue more than one for each entity, resulting in clearer accountability of request volumes. Perhaps suggesting previously existing similar names when a report is being submitted would avoid accidental name duplication. Aggregating GTR data by hand will be done when the particular query makes the duplication apparent, but for whole queries like averages or maximums, this duplication may impact the results. A company seeking to deliberately diffuse the perception of the amount of notices it is making could outsource its work to reporting agencies, but also make slight changes in their reporting name or take out another ID number and hide its zeal for takedown procedures.

Concern over senders' identity is a common theme in the literature, and a topic of the "architecture of control" framework seeks to understand. Urban and Quilter's study identifies repeat senders as being responsible for 62.4% of the notices" (Urban and Quilter 2005, 652) even though they made up only a quarter of all senders in 2005. In the same study, only 2% of repeat senders issued more than 10 notices, and the maximum notices a single sender issued was 54 (Urban and Quilter 2005, 653). But the numbers have continued to increase for repeat senders in the 2011-2013 data, where 8.27% of reporting agents sent more than 10 notices and 39.66% (5,998 of 15,121) are repeat senders. The following table compares the GTR 2011-2013 to Urban & Quilter's cumulative data. Seng did not provide disaggregate percentages as a result of notices, so his data cannot be contrasted. This data shows how one-time-notice senders have diminished in 8 years as a ratio of all notices sent.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Senders have sent up to:	Count	Percent	Reporters have sent up to:	Count	Percent	Owners have sent up to:	Count	Percent
1 notice	329	75%	1	9,523	62.98%	1	16,389	56.43%
5	86	20%	5	3,572	23.62%	5	6,740	23.21%
10	10	~2%	10	774	5.12%	10	1,728	5.95%
20	6	~1%	20	514	3.40%	20	1,394	4.80%
More	5	~1%	100	506	3.35%	100	1,807	6.22%
			1,000	179	1.18%	1,000	886	3.05%
			10,000	41	0.27%	10,000	91	0.31%
			100,000	10	0.07%	100,000	8	0.03%
			More	2	0.01%	More	1	0.00%
Total:	436		Total:	15,121		Total:	29,044	
	unique senders			reporters			owners	
Maximum notice number: 54			Maximum notice number: 2371,130			Maximum notice number: 142,274		
Source: Urban and Quilter			Source: Isadora Borges, 2011-2013 GTR					

Table 1 Comparison of notice frequency by type of sender (2002-2005, 2011-2013)

Studying notices as a measure of frequency, has its shortcomings, however. In particular, ranking by notice frequency can thus obfuscate the distribution of requests within each notice. For example, individuals¹⁹ (called unspecified reporting agents²⁰ in the GTR) are 5th by the frequency of notice measure on the 2011-2013 database, but submit only 4.8% of all requests, and resulted in only 0.2% of removed URLs. Through their unique IDs however, one can infer there are 9,713 individual senders. Individual senders are more likely to report instances of personal, limited piracy, instead of widespread copyright infringement. Thus, even as reporting agents' high volume of activity is easily spotted when looking at the amount of notices issued, if one looks at figure 4, one must remember that from the top 30 reporting agents who produce 85% of notices almost 5% belongs to over 9,000 individuals.

¹⁹ Marked in the GTR as "unspecified agents", Urban and Quilter and Seng infer an individual sent a notice because Chilling Effects removes identifying information for this type of sender, as opposed to copyright owning corporations and reporting agents.

²⁰ Google's full disclosure for these entries reads: Unspecified: The individual who made this removal request did not provide a name for their reporting organization. This means that they may not work for an organization or chose not to provide its name. Different individuals who did not provide an organization's name can be distinguished by other means, which is how we're able to identify and list each item.

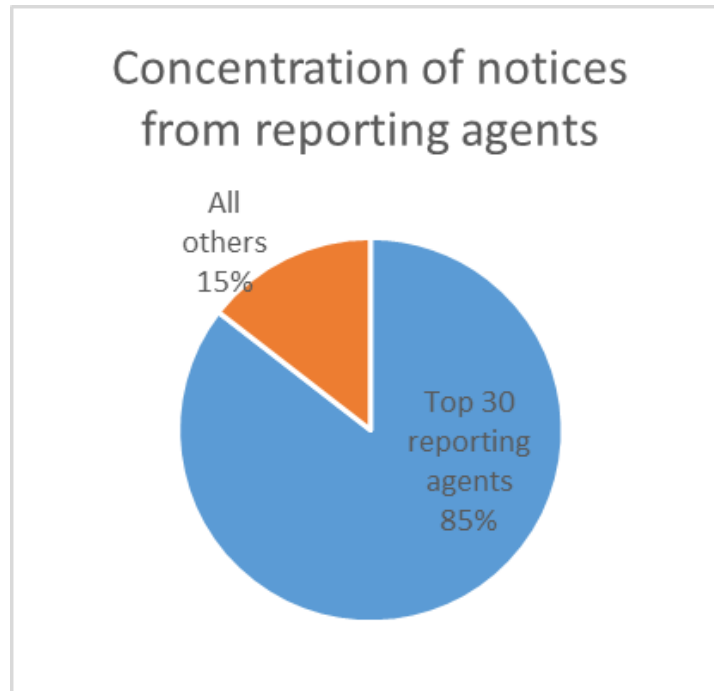


Figure 4 Reporting agents' concentration by notices (2011-2013)

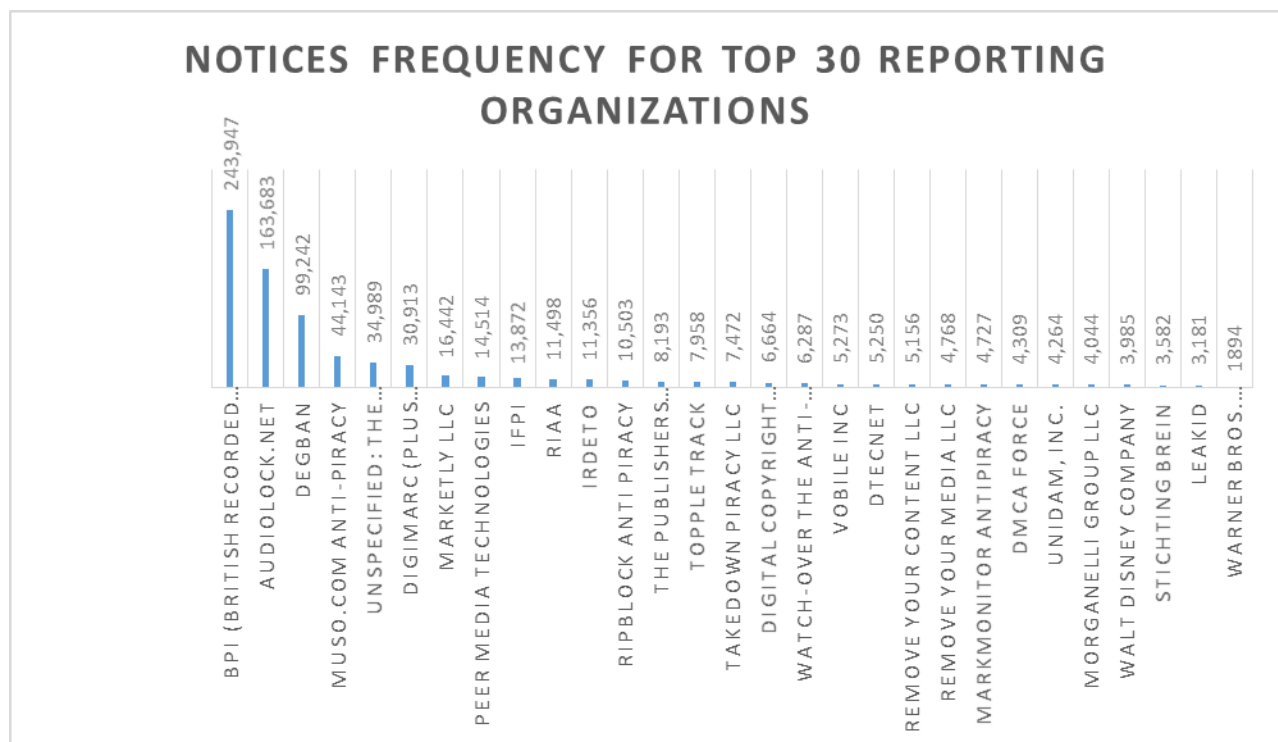


Figure 5 Top 30 reporting organizations' frequency of notices (2011-2013)

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Instead, let us review the data by requests, or individual URLs reported. Longitudinal data from the M2011-May2014 GTR shows that the year 2012 had almost 18 times as many URLs requested to be taken down than in 2011, and that in 2012 the numbers grew to almost 4 times as much as the previous year. By May 2014, the data is almost 50% of 2013's. An annualized histogram of removed URLs was created using M2011-May2014 data for both copyright owners and reporting agents. An "expected" measure was created by dividing the 2013 data by 12 and multiplying it by 5 to represent the first 5 months of 2013, and was compared to the first 5 months for 2014. In all but one measure (reporting agents who have successfully removed between 2-10 URL requests) is this year on track to surpass 2013's numbers. These instances are highlighted in red.

	<i>Upper interval limits for removed by reporting agencies</i>	<i>From March on, 2011 Frequency</i>	<i>2012 Frequency</i>	<i>2013 Frequency</i>	<i>Until May 2014 Frequency</i>	<i>(2013 data/12) *5</i>
	1	2060	5017	8897	4316	3707.083
Reporting	10	907	2369	3976	1638	1656.667
Organization	100	541	1418	2470	1032	1029.167
	1,000	178	471	832	373	346.6667
	10,000	45	116	232	150	96.66667
	100,000	8	36	91	51	37.91667
	1,000,000	7	19	26	20	10.83333
	More	0	10	23	19	9.583333

Table 2 Longitudinal histogram of URL removal (M2011-May2014) by reporting organizations

	<i>Upper interval limits for removed by reporting agencies</i>	<i>From March on, 2011 Frequency</i>	<i>2012 Frequency</i>	<i>2013 Frequency</i>	<i>Until May 2014 Frequency</i>	<i>(2013 data/12) *5</i>
	1	2065	5096	8816	4347	3673.333
Copyright	10	928	2520	4090	1816	1704.167
Owners	100	662	1928	3099	1711	1291.25
	1,000	328	1143	2077	1618	865.4167
	10,000	128	471	1012	970	421.6667
	100,000	32	149	357	300	148.75
	1,000,000	5	46	96	67	40
	More	0	9	28	26	11.66667

Table 3 Longitudinal histogram of URL removal (M2011-May2014) by copyright owners

The following tables show the annual tally of each year's top 10 reporting organizations. The top ranges double every year, as the scale grows larger. Again, 2014 is on track to show similar growth, as in the 5th month the biggest reporter has already surpassed half its request numbers for the previous year.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

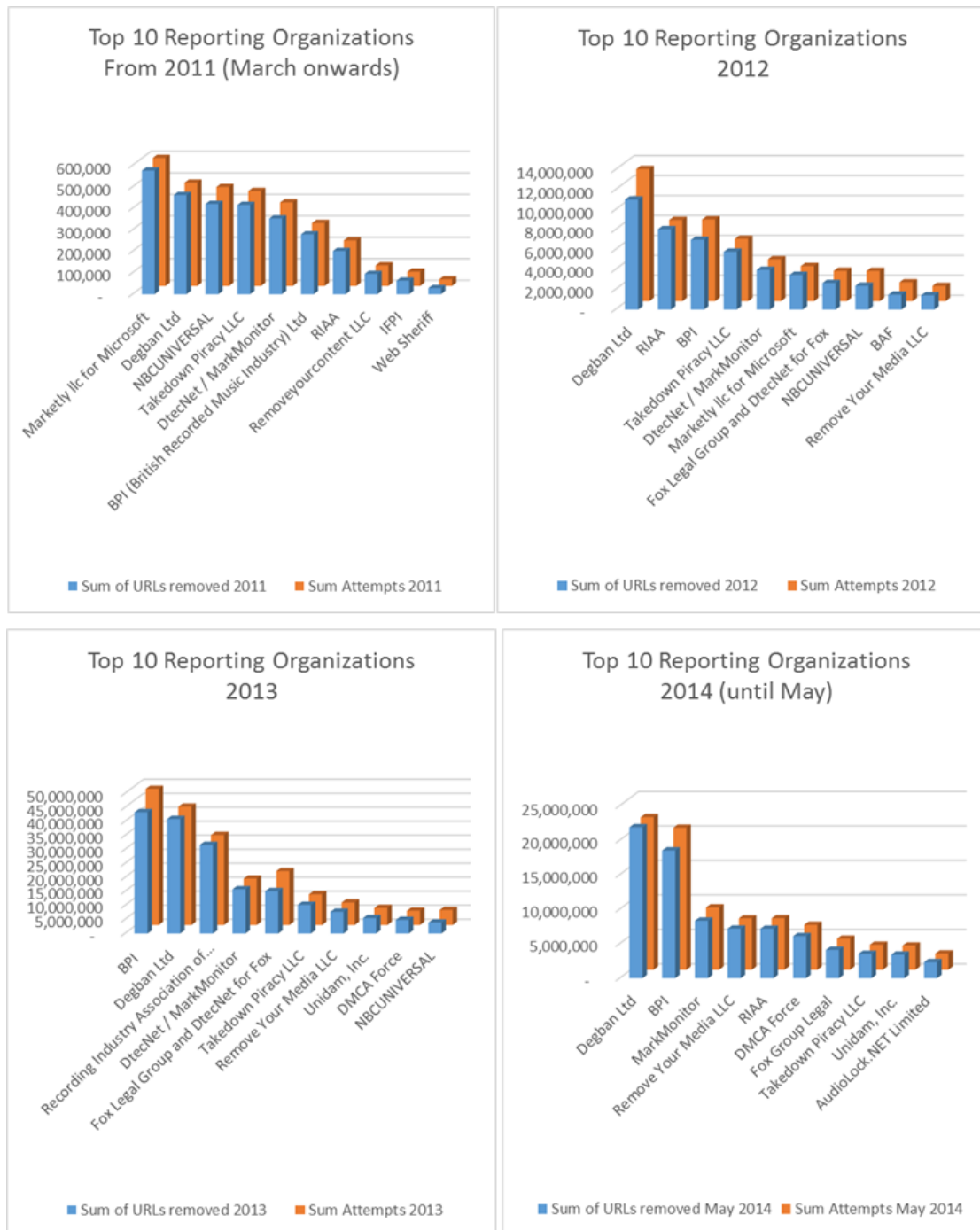


Figure 6 Four years of request by the top 10 reporting organizations (M2011-May2014)

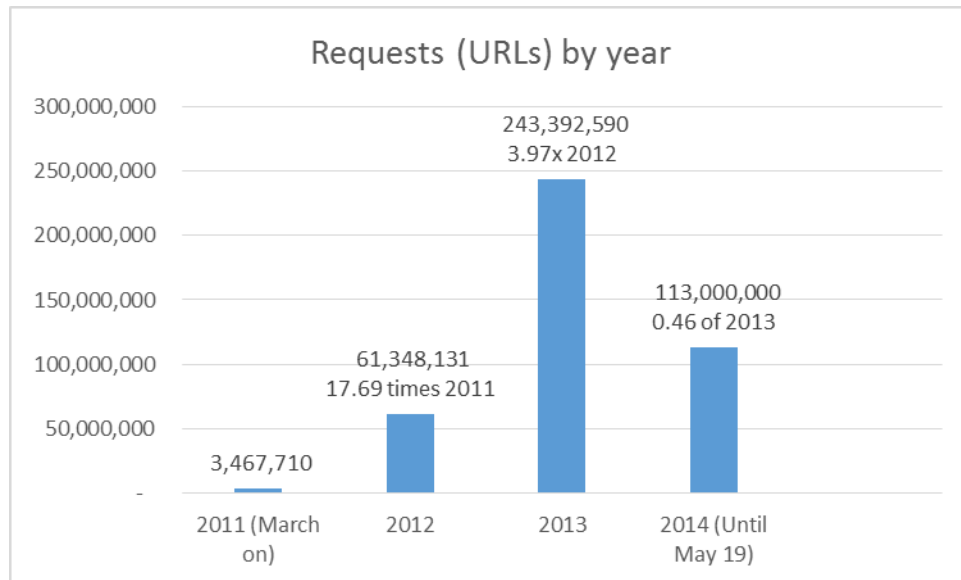


Figure 7 Total requests by year (M2011-May2014)

So what is the breakdown for requests? Of the top 100 senders, aggregating copyright owners and their agents, consolidating their name disparities and presenting the maximum data for organizations that list themselves as copyright owners and agents yields entities that have flagged at a minimum, 134,152 URL links (by DM Network Limited) and at a maximum 53,449,669 URLs (put forth by the British Phonographic Industry (BPI)). An industry breakdown ranked by their ability to take down links ranks TV and film industries almost consistently above those defending the pornography industry. Representing agents, the third parties that carry out automated searches and takedowns on behalf of a portfolio of clients represent for the four kinds of rankings slightly more than double of the most represented content industries.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Industry Categories	frequency top 100 by removed	frequency top 50 by removed	frequency top 30 by removed	by over a million removed	% over million removed	by over a million attempts	% over million attempts
Books	2	1	0	1	1.85%	1	1.82%
Photography and Design	1	0	0	0	0.00%	0	0.00%
Videogames	4	2	1	2	3.70%	2	3.64%
Music	8	5	2	6	11.11%	7	12.73%
Software	3	3	2	3	5.56%	3	5.45%
Pornography	20	6	5	8	14.81%	8	14.55%
TV and Films	18	11	6	9	16.67%	9	16.36%
Representing Agent	44	22	14	22	40.74%	22	40.00%

Table 4 Top removals by industry categories (2011-2013)

Seng at times distinguishes collective management organizations (CMO) from content owners and reporting agents (Seng 2014, 52), but the distinction seems artificial against reporting agents. CMOs are not a category in Google's reports, and they either act as reporting agents in the sense that they take care of takedown notices for a variety of interests, like the Netherlands' Protection of the Rights of the Entertainment Industry (Stitching BREIN), or easily identified to a particular industry, like the UK's Publishers Association, which is why I aggregated the data accordingly. The Publishers Association ranks 64th by successfully removed URL volume and 61st in overall requests sent to Google. It is also one of only two reporting agencies solely dedicated to the publishing industry. The other is Digimarc, previously Attributor, which ranks 46th and 47th in each category.

On the other hand CMOs that also state themselves as the copyright owners serve to make sure the takedown process takes advantage of their collective resources. For example, representing agencies make up 40% of those entities that have submitted over a million URLs to be taken down. 54 reporting organizations and owners have submitted and had removed over a million removed requests, 37.04% of them come just from reporting organizations, 40.74% if you count CMOs as serving a reporting organization purpose, placing them squarely in the majority of all kinds of participation. In contrast, all individuals aggregated have submitted 735,971 URLs, of which 527,058 were removed a 0.7161 success index, no doubt hampered by its unprofessional nature.

As an industry, the GTR data of the 50 content owners with the highest URL submissions was dominated by the music industry and pornography, with TV and film copyright owners rising as compared to Seng's data (Seng 2014, 30) to a close third, overtaking videogames and software. Books were only represented in the top 50 content owners by the UK's The Publishing Association, which issuing 898,857 requests from 2011-2013, or 0.3052% of the data subset.

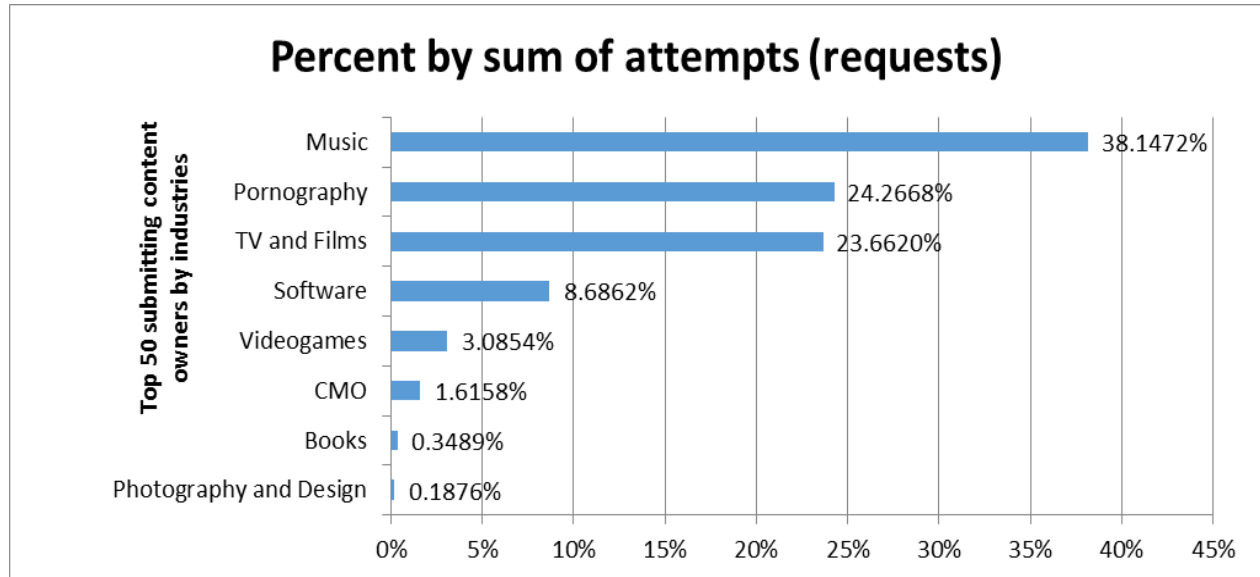


Figure 8 Top 50 sending content owners' request percentage (2011-2013)

As takedown notices have become increasingly deployed Google has begun issuing "trusted" status to some, adding an express queue to solve their requests. The result is an entrenchment of the architecture's bias towards copyright holders' claims of copyright infringement as being valid from the outset. But this kind of program makes sense from the point of view of efficient management of copyright owners' expectations because the bulk of submissions come from a few, concentrated, complaining parties. For example, Seng calculates that of all notices issued by reporting agents between 2008-2012 36% and 59.6% are top 30 reporters (Seng 2014, 20), by frequency of reporting.

Google adds providers and agents to their Trusted Copyright Removal Program in notoriously obscure ways (Seng 2014, 36). Seng surmises the program began in March 2011, but Google places a

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

more general effort to “streamline” the search takedown process in 2010, when they began receiving 15 times more URL requests (von Lohmann, Report: How Google fights piracy 2013). According to Fred von Lohmann, Google’s senior copyright counsel, the company began processing approximately 4 million requests weekly, averaging less than 6 hours per request (Google 2013, 3).

At first, the program had a tiered system where some copyright owners and reporting agencies had a 1,000 request cap (including Warner Brothers and Attributor, now Digimarc), while others had 10,000 requests (including Fox’s Legal Group), and a tier that allowed for 10,000 requests in *each* notice (Seng 2014, 38)—implying complaining parties like the RIAA, NBCUniversal and British Phonographic Industry (BPI) could send more than one notice with the maximum for the second group each time they did. The levels may have been awarded by prior volume, as an appeasement move from Google, which is often criticized of not doing enough (Jones 2012). The caps were controversial with many content owners, who thought the limits were artificial (Buckles 2012), and not, as Google claimed, a safeguard against crashing their website with distributed denial of services—as the volume of automated requests would act.

Overall limits were finally removed in April 2013, and although limits per submission were sustained, the amount of times submissions could be made became unlimited. The effect was eliminating the tiered system for those TCRP partners, “increase[ing] these limits for partners who have demonstrated a consistent track record of submission quality and volume” (Ernesto, Google Relaxes DMCA Takedown Restrictions, Eyes Abuse 2013). The program has also changed to allow .txt formatted files with a 1 MB limit (equivalent in disk space to the size of a pdf book) of URLs (von Lohmann, Google’s DMCA notice and takedown tools 2014, 7), among other ease-of-use measures to make reporting less cumbersome for automated batch-processing reporters.

As the advantages of the TCRP have grown, its membership has increased. Seng identified 46 trusted users (Seng 2014, 36-38) between 2011 and 2012. This corroborates Google’s calculation that in 2012 there were approximately 50 participants (Google 2013, 15). A look at the GTR data that includes

2013 suggests that there are *at least* 90 TCRP participants now, who have submitted 292,275,017 of 422,198,832 requests, 69.23% of the URL requests submitted for 2011-2013. Thus, only 10% of the top URL link submitters do not participate in the Trusted Copyright Removal Program. Access to streamlined takedowns could be an asset for companies that do piracy monitoring on behalf of content owners. To the extent that limits were expanded to members of this program, and then dropped because of their petitions to Google, they are able to send larger amounts of requests and relay to their clients impressive figures. This may help reporting agents make the case for being genuinely needed—firstly because they have access to the TCRP's benefits, and secondly because they are "taking down" thousands of links.

Copyright owners successfully got an automatic content sweeper for YouTube in 2007 called Content ID, where right holders can upload a database of reference files and metadata, and direct Google's video service to (1) make money from the upload (the most used option), (2) leave it and track the viewing statistics, or (3) take it down from YouTube (Google 2013, 10), effectively putting the NII's dream of an omniscient licensing system into effect. From the voiced complaints, it seems like copyright holders would like all internet content to be monitored thusly. However, this approach seems more difficult for searches. The volume of addresses (more than 60 trillion according to Google's September 2013 report (Google 2013, 14)) alone could make the feat a daunting one, even if considerations of fair use were ignored. The issue scholars like Urban and Quilter raised in 2005 is that Google's reliance on copyright owners to distinguish fair use places too much trust in their judgment. Nonetheless, Google insists on copyright holders' and their agents' indispensable quality as reporters of infringement (Google 2013, 14) and has ended up enabling the automated discovery mechanisms these parties deploy by removing the cap for any authorized entity to report. However, the change in policy has come with added caution to request abuses. The touted 97.5% request removal rate from December 2011 to November 2012 (Google 2013, 15) dropped to 90.5% in the GTR 2011-2013 database, although the M2011-May2014 data is presenting a slight uptick so far.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Year	Requests (URLs) removed	No action URLs	URLs pending review	Sum Attempts	Success Index	No counter-notice	Counte-rnotice	Total Notices
2011 (March on)	3,198,179	269,531	-	3,467,710	0.9223	35,416	112	35,528
2012	54,354,143	6,993,988	-	61,348,131	0.8860	427,521	636	428,157
2013	221,958,501	21,434,037	52	243,392,590	0.9119	479,834	363	480,197
2014 (Until May 19)	105,250,304	8,135,695	61,223	113,000,000	0.9314	250,481	29	250,510
Total	384,761,127	36,833,251	61,275	421,208,431	0.9135	1,193,252	1,140	1,194,392

Table 5 GTR longitudinal data (M2011-May2014)

Seng sees a similar drop in takedown rates, as compared to the vaunted number, as cause for worry (Seng 2014, 39) for participants of the TCRP, as their rates are lower than the overall average in his data. Yet the 2011-2013 GTR database presents a 90.495% overall success index for entities of all kinds, and an 89.10% success index for the 90 TCRP participants in the top 100 index, a very slight underperformance.

One could even make the counterargument that if speed is prioritized, a degree of extra scrutiny should follow. There is also a possibility that the reason why links are being marked as not taken down is because the speed at which petitioners' automatic systems compile links outpaces Google's responses, so that links are repeatedly being asked to be taken down before Google has a chance to act on the first request. On subsequent petitions, the search engine will mark it as not having been taken down, not because it did not try to remove a URL, but because Google could no longer find a result to remove (Google 2011, Why don't you list all of the URLs that you didn't take action on?). Recent data from the M2011-May2014 database shows this is especially the case for the two biggest agencies which are dedicated to book, journal and document piracy, The Publishers Association, and Digimarc (Attributor before). They remain the most active²¹, but their success index has dipped below their previous rates.

²¹ Other reporting agencies that serve the publishing community were found through a "publish*" text filter in the agency name. Because of the volume of reporting agents, the percent displayed in the table is an approximate and will not contemplate other agencies that do not refer to publishing in their name. For example, the Dutch Stitching BREIN may report on authors' works, but it is not obvious if they do from the GTR.

Yearly requests:	Sum of URLs removed	Sum of Sum Attempts	Success Index
2011	2,081	2,514	0.8278
The Publishers Association	1,860	2,141	0.8687
Percent of 2011 publishing takedowns	89.38%	85.16%	
2012	292,537	322,499	0.9071
The Publishers Association	226,917	255,154	0.8893
Contributor	62,967	64,545	0.9756
Percent of 2012 publishing takedowns	99.09%	99.13%	
2013	2,028,484	2,326,161	0.87203
Contributor	1,072,821	1,196,348	0.8967
The Publishers Association	514,633	677,769	0.7593
Digimarc	434,378	442,261	0.9822
Percent of 2013 publishing takedowns	99.67%	99.58%	
2014	18,844	1,356,681	0.01389
Digimarc	16,808	921,059	0.0182
The Publishers Association	1,754	427,078	0.0041
Contributor	4	49	0.0816
Percent of 2014 publishing takedowns	98.52%	99.37%	

Table 6 Publishing industry reporting agents' longitudinal request rates and success indexes

Another complaint from copyright owners Google has adapted to is that websites that have repeatedly infringing users (dubbed as “pirate websites”) continue to be online. As measures to change the search results Google has, for example, removed from its instant and autocomplete features piratebay.org, Pirate Bay, BitTorrent, torrent, utorrent, RapidShare (Ernesto, Google Starts Censoring BitTorrent, RapidShare and more 2011). It has also made such repeated claims affect their overall rankings (Google 2013, 15). It has not, however, caved to pressures to refuse to show them altogether, unlike the UK government, which pressured ISPs into filtering and refusing access to sites like The Pirate Bay (Clark 2013). Furthermore, despite the TCRP program getting more press coverage, it is not advertised in its general takedown notice information, and Google maintains a veil of secrecy over how trusted status is obtained and who all of its members are, as well as what constitutes streamlined treatment.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Counter-notices remain a blip on the radar. Even though Google states it notifies webmasters of removals, it can only do so if the person uses the company's Webmaster Tools (Google 2013, 17). Legally, Google has no reason to go out of its way to find and inform anyone of a §512(d) takedown, thus, removing the 6 counter-notices Seng found for 2011 (Seng 2014, 46), and the possible 7 for 2012 (if we allocate the 5 marked as "unknown" in his table 12), we are left with 1,095 counter-notices filed for 912,659 unique notices, or 0.12%. Urban and Quilter speculate that few bother with filing a counter-notice because the 10-14 day delay that is required by law to give complainants a chance to sue is too cumbersome as compared to just moving the material to another hosting service (Urban and Quilter 2005, 60). But not only is an informed decision impossible because counter-notices are impossible to solicit unless there is a web hosting relationship with Google, there is no legal requirement for Google to seek it out. In this sense, the DMCA is entirely upside down for uploaders accused of copyright infringement, as they are protected neither by a technological fix nor by the scope of the law.

One could assume that the small proportion of counter-notices filed, means that the system is working. This is the position of Ben Sheffner, vice president of legal affairs for the Motion Picture Association of America, who cites a study made with only data his organization provided to the authors. His opinion is that because the problem of abusive notices appears "anecdotal" and "tiny" compared to the overwhelming amount of notices (25 million takedowns are cited by the piece), their completeness, legitimacy or truthfulness need not be explored. This view, as the article reflects, ignores that §512(d) takedowns do not require notifying accused infringers that their content is being un-indexed (and hence that they are being accused of copyright infringement). Furthermore, starting a counter-notice procedure also gives the copyright holder the option to sue the alleged infringer—a great dissuader.

Finally, while Google's broadband internet development might pose an increased pressure to comply faster with takedowns in the future, it could also incentivize more counter-notices being filed if OSP uploaders can be identified as ISP users and contacted in an efficient manner. This consolidation of

functions would not pose a conflict with the legislation, since Google enjoys blanket safe harbor under subsection (a), but liability may still arise from its other functions it carries out under each subsection's particular conditions.

4 Academic publishing's use of takedown notices

As we can see from the data, the majority of requests made and fulfilled come from the efforts of a few entities. The requests are also heavily segmented by industry, of which popular media (music, pornography, TV and films) are the most prevalent sector. But books are represented by dedicated agents on a smaller level in the subset of top 100 requesting parties (.565% of requests) by Digimarc (called Attributor before), and the British Publishing Association. Of these two companies, Digimarc acts on behalf of 8 non-profit university presses in the US and the UK²², which are part of the American Association of University Presses. It also acts on behalf of for-profit journal publishers like Elsevier, non-profit organizations like the Institute of Electrical and Electronics Engineers publications, and commercial trade book publishers²³.

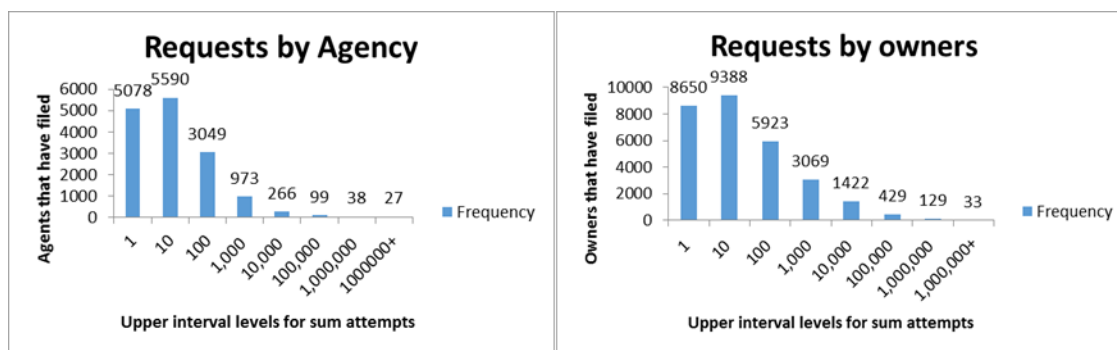


Figure 9 Histogram of requests by agencies and owners (2011-2013)

²² Cambridge University Press, Columbia University Press, Duke University Press, Harvard University Press, Northwestern University Press, Oxford University Press, Princeton University Press, University of Chicago Press.

²³ Random House, Harper Collins, Pearson Education, Simon & Schuster and Shueisha (Japanese publishing), Grupo Planeta (Spanish and French publishing), Quercus (fiction publisher) among others. See <http://www.digimarc.com/guardian/customers> and <https://www.google.com/transparencyreport/removals/copyright/reporters/28460/Digimarc/>.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Top 100 Submissions Industry Categories	Sum Attempts by Industry	Percents
Photography and Design	483311	0.113%
Books	2408607	0.565%
Videogames	7947857	1.864%
Software	22375518	5.249%
TV and Films	60952815	14.298%
Pornography	62510926	14.664%
Music	100909551	23.671%
Representing Agent	168709542	39.575%

Table 7 Numeric percentage of sum of requests by top 100 industry categories

The Association of American University Presses (which also has other international members) (AAUP 2013), is one of the organizations which represents scholarly publishing. It distinguishes itself from other academic publishers like Elsevier by being an umbrella organization to university presses with non-commercial ethos (AAUP 2013), looking for intellectual impact over sales performance when choosing contemporary pieces. This is not to say that university presses are charity organizations, but they can publish niche items cross-subsidized by classical, well sure-selling titles without the pressure of turning a profit on every title they choose. The AAUP however agrees with the Authors Guild that the Google scanning and book indexing program infringes authors' copyrights (AAUP 2013), and prioritizes embargo periods for open access to other mechanisms for the recovery of production costs for journals, despite acknowledging the short life span science research has (AAUP 2010, 2). And on the point of copyright infringement, its website currently cites intellectual property defense industries' estimates of how damaging (in scope and monetarily) e-book piracy is.

The Publishers Association does its own copyright surveillance, claiming itself as the owner and reporting agent on all takedown notices in the GTR. Like Digimarc, it has both for profit and non-profit publishers in its portfolio. Two British universities are members of the Publishers Association but choose to employ the American company Digimarc to do their search-based takedown notices, as the Google's Transparency Report reveals. The Publishers Association's Chief Executive Richard Mollet has described

as "very timely" a report by Mike Weatherly MP, intellectual property advisor to David Cameron, arguing that Google is not doing enough to curb online piracy (The Publishers Association 2014). What these examples of academic publishers' stance on digital technologies and the prospect of copyright infringement show is that they apply the same logic to academic publishing than trade book publishers do to their products. They therefore align themselves within umbrella organizations that pursue a homogenous tactic against piracy, as if the incentives to produce and the effects of sharing their works impacted both kinds of authors in exactly the same fashion.

American copyright law, the DMCA included, has a firm tradition of a utilitarian perspective that draws a causal relationship from market-based, economic incentives to individual creative production, resulting in technological and scientific progress. Yet, those who study digital semiotic technologies and/or digital copyright encounter a required assumption within anti-piracy rhetoric: equating tangible property to intellectual property can be a misleading premise. First, intangible works are non-rival and non-exclusive goods. At most, scarcity is a function of time and funding, ideas do not run dry like water wells do. Lessig reminds us that the production of property (physical or intellectual) is boosted by laws that create an environment of "sufficient incentive" to produce as a means for creating marketplace—which means an environment of "less than perfect control" (Lessig, Code: version 2.0. Online version 2006, 183). Gillespie warns us, however, that it is the "logic of a market economy [which] suggests that authors need an assurance that they'll be able to sell their work" (Gillespie 2007, 23). In fact, the "most logical strategy in purely economic terms is in fact piracy, except for the fact that copyright makes piracy illegal" (Gillespie 2007, 25). Thus, the more rights a property regime bestows, the more its underlying architecture of control seeks to fix the "market failure" that will arise if there is no legal incentive (i.e. the promise of remuneration) to produce property.

Because a physical property is rival and exclusive in terms of another physical property, the state also needs to act as a watchman and protect the integrity of an object's possession. But for intellectual

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

property, particularly the non-exclusive, non-rival digital kind, “the law need only create the incentive to produce” (Lessig, Code: version 2.0. Online version 2006, 183). Because “sufficient incentive” does not equate perfect control (Lessig, Code: version 2.0. Online version 2006, 198) pre-digital copyright law did not create an architecture of control that sought out this goal. This is why US copyright law, regardless of what the DMCA omitted, still upholds the value of fair use, time-limits to copyright (however extended these are) and the first sale doctrine. While the DMCA framed intellectual property from a wholly utilitarian perspective, the privileges it grants are still part of a constitutional tradeoff to make public valuable creations and information²⁴.

As discussed in the introduction, the structure of universities and other research institutions provides a sufficient incentive to produce (through a salary, tenure, fellowships, the promise of sabbaticals, etcetera) for academic authors that is different from commercial authors, artists and even industry scientists. The first group has historically not been expected to produce widely consumed best-sellers. When they do, they are remunerated beyond their sales through the research system they are a part of, but this reward mechanism is not exclusive to such exceptional cases.

But where an incentive to produce exists regardless of the law, its application can become a hindrance. Here, I will work out Lessig’s proposition that “the ideal protections of intellectual property law are something less than the ideal protections for ordinary or real property” (Lessig, Code: version 2.0. Online version 2006, 198) in digital academic works. We should start out by distinguishing that not all publishing is created equal, as the open access UK legislation and the US directives for journals only have recently reminded us. While this thesis focuses on academic e-books, the drive for sustainably funded open access speaks to a similar concern of making high academic work available.

When it comes to academic work, there is an added issue in equating tangible goods to intangible goods’ property rights, which resides in defining what motivates scholastic production nowadays. While

²⁴ As a reminder, copyright is granted over the style or form of a statement, not the factual information it conveys.

there are costs involved in the production of knowledge, it is a rare event when an academic book becomes a best-seller. Those involved in academic publishing are most often rewarded for their hard work on the basis of the exposure it gets, as translated into numerous impact metrics which feed into perceptions of status within a field, in turn informing grants, research and teaching positions, unpaid interviews and conferences, etc. Sometimes publishers' advances are an attractive incentive. But publishers are not record labels. They are hardly wont to approach an unknown scholar to "make them a star" with a big signing bonus. They are cognizant of interesting projects from researchers who have made thought-provoking article publications, turning an enticing idea into a book deal that will circulate 350-500 copies²⁵ at most.

Lead researchers in scholarly copyright also contend that the particular market of scholarly works at the moment makes each individual both consumer and producer (McGuigan 2014, 252), with little readership outside academic circles—though proponents of open access seek to change this. Hence, while most copyright discourse overlooks the role of the objects' consumer (Elkin Koren 2007, 1125) scholarly authors' interests are more aligned to consumers' interests than other forms of media. Of course, if controls on acquisition were less restrictive, consumers could expand to include researchers from smaller schools, foreign students of universities whose libraries have lesser budgets ("the global middle class" (Kelty 2012), journalists, ambitious high school students, and members of the general population interested in higher research. In Toffler's tradition of consumers as participant producers (Toffler, *The Third Wave* 1980) (Toffler and Toffler, *Revolutionary Wealth* 2006), or Professor Elkin Koren's concept of consumers-as-authors (Elkin Koren 2007, 1139), consumption of informational works is itself productive to social dialogue, nurtures production of human capital and new content, and informs citizen participation (Elkin Koren 2007, 1137, 1142), which can itself have a positive impact on researchers' work.

²⁵ This figure does not include on demand printing or authorized distributions of electronic versions, but private communications with sources familiar to academic publishing have corroborated that increasingly print on demand and e-books have not expanded sales figures but replaced the traditional book run market.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

Professor Cristopher Kelty goes so far as to compare the academic publishing industry's takedown tactics to the pharmaceuticals' patent frenzy in the face of the AIDS epidemic in an op-ed about the closing of the website library.nu's closing. Because academic publishing's model has centered on selling few books at high prices to North American and European libraries they have cut off foreign markets from more affordable access. Kelty addresses his peers by confronting them with his perception of their complicity in the shutdown of a website dedicated to servicing the "global middle class" of the world: students from Latin America, China and Eastern Europeans who are not the "truly poor [...]" but nonetheless they do not have much money" (Kelty 2012). In his opinion European and North American scholars should stop assenting, if not actively supporting, publishing companies' decision to bring down websites that make orphan works, rare volumes or other kinds of academic works to these marginalized markets when the uploaders are acting in a non-commercial model. To him, "books and their scholars are the losers in this trade-off" of higher profits against more accessibility to the global academic middle class, which even shuts out European and North American cash strapped libraries, because their work has become increasingly unavailable or so enshrouded in DRM "as to make using them supremely frustrating". The effect is that the "best scholarship is being shot into some publisher's black hole where it will never escape" until non-commercial sharing websites make them available.

It seems that current copyright law has forgotten the adage of intellectual creation from "standing on the shoulders of giants", as if artists and researchers create[d] from nothingness, creation is an incremental process. But the progress constitutionally ascribed to the right to copy requires that the public be able to gain access to these works and to extract their value as part of the quid pro quo a temporary production monopoly bestows. However, the liberal democracy tradition of the US leaves this access up to the market, which is a policy choice just as much as the DMCA's expansion of access control is a policy choice. Using copyright law to stress economic incentives for present creators and their publishers, record labels, etc. by using access controls excludes a second wave of creative use on media works. By

bottlenecking the promotion of future creativity and innovation, copyright policy weakens the ability of consumers to become creators of transformative works. Academic authors in particular have little to gain from such an environment, as their work impinges on their ability to access relevant and new research, and critically engage with it. As Elkin Koren bluntly states, “[p]romoting creativity requires expanding access to creative works. Copyright law must therefore expand the balance it strikes between authors and users to cover not only subsequent authors but also simply consumers of cultural goods who might become authors in the future” (Elkin Koren 2007, 1141). Because the spread and use of knowledge, particularly in citations, reviews, etc. is the driver in the non-commercial incentive structure of academia, academic *writing* is driven by reputation, not profits. This doesn’t mean academic authors have taken, or should take, vows of poverty. Rather, the incentive to produce for this group does not lay mainly in the immediate success of a book’s profits, as this figure is usually not their main source of income. Instead, the works’ ideas are subjected to a set of quantitative and qualitative metrics (such as citation impact, conference participation, uncovering an unknown fact of revolutionizing a field, etcetera) which are rewarded through the researcher’s employment system (better salaries, tenure, smaller classes, sabbaticals, fellowships, etcetera). The same is not an inherent system for all the publishers that reproduce said material. Those publishers may not share their authors’ interest in more accessible circulation. Like commercial authors, most publishers will see their gains (economic and in status) closely tied to the publishers’ profit margins.

By subjecting academic books to the same degree of automated surveillance of distribution to trade books that aim to be best sellers, publishers are construing all types author’s incentives as one and the same in wholly different settings, judging information works by their economic utility and price (Elkin Koren 2007, 1132) as other commercial media and obfuscating other merits to their circulation. Restriction on use does not just have the effect of equating the rights of intellectual work to tangible goods, with the incorrect notions of rivalry and exclusivity doing so entails; it also makes us stop thinking of these as a self-

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

expression works (the style of which is copyrighted, not the facts it describes), making us view discoveries and research as commodities owned by someone who exercises sweeping control over their use.

Academically rigorous studies have found there is no correlation between piracy and dwindling sales²⁶, and even with all the “rampant” commercial piracy, book sales are going strong (Publishers Weekly 2014), showing an increase in e-book sales and university presses (Publishers Weekly 2014). Furthermore, academic book piracy is a small subsection of digital piracy. Usually tended to by specialized indexing websites instead of the larger websites which host bestsellers, they are highly vulnerable to domain seizures and do not count with a robust “backup” system torrenting or mirrored websites²⁷ have. When websites specialized in I have also tried to argue that authors have more to gain from a system of imperfect control of their production where their findings can be discussed and shared throughout larger sectors of the population. Digital piracy has not been the reason why most academic authors do not make their livelihood from selling 350-500 copies of a monograph. Most academic books do not crossover to popular science genres and become the next *Freakonomics*. Other activities like journal article authorship and peer review as part of academic activities and outside a for-profit logic of production are incompatible with the aspiration of perfect control of distribution in other cultural industries like TV or music. I would like to propose that academic authors should have the option of opting out from the automated takedown system their respective publishers (often academic, university non-profit presses) are presently enrolling them in.

²⁶ “[M]usic sales in the United States are less than half of what they were in 1999, when the file-sharing site Napster emerged”, said Cary H. Sherman, RIAA lobbyist, during the SOPA and PIPA aftermath **Invalid source specified.**, contradicting empirical and academically rigorous findings. For example Oberholzer-Gee and Strumpf concluded that “[d]ownloads have an effect on sales that is statistically indistinguishable from zero. Our estimates are inconsistent with claims that file sharing is the primary reason for the decline in music sales during our study period” **Invalid source specified.** Similarly, for TV and movies, Smith and Telang found that “the post-broadcast increase in DVD sales suggests that giving away content in one channel can stimulate sales in a paid channel if the free content is sufficiently differentiated from its paid counterpart. Likewise, our finding that the presence of pirated content does not cannibalize sales for the movies in our sample suggests that if free and paid products appeal to separate customer segments, the presence of free products need not harm paid sales” **Invalid source specified.** These studies clearly contradict the intuitive, if unquantifiable assertions that free content versions are the cause of a decline in media products’ sales.

²⁷ Mirroring refers to how bigger piracy repositories have alternate domains they can bring online if their main domain URL is seized. Torrenting websites can also have mirrored versions, but the content itself is less prone to being taken offline because it is not centrally hosted by a single server. Rather, users connect through a software and share the versions they have in their hard drives while they are online.

The DMCA had in mind the creation of a mostly commercial network that would monitor, track and license every piece of information into a compulsory license system, deliberately avoided legislating on the subject of fair use and instead provided copyright owners with the legal ability to technologically enforce absolute control on access to media and information. “Progress of Science and useful Arts” is the constitutional goal, and the property rights over their works is the means by which they are as a whole meant to be incentivized into the marketplace. The goal is not, as the legislative proceedings might suggest, a jobs program. But copyright law has built incentives to the whole of intellectual and creative endeavors as if progress was defined the same for all communities, and as if all recipients were one and the same class of consumer. But academia benefits from the innovation of “a wide participation of individuals and collaborative communities” (Elkin Koren 2007, 1152) produces in ways consuming TV or trade books does not. Academic authors themselves benefit through their production incentives when their work is accessed and used by their peers in ways actors or commercial authors do not because of the non-rival, non-exclusive nature of intellectual work within modern research institutions. I join the ranks of Lessig, Litman and Elkin Koren in asking, does bottlenecking access to informational searches or sweeping the internet with automatic takedowns make sense in the world of academic digital media, from a public policy perspective? Just because technology now grants us the power to enforce what was hard to enforce before, should we? The principles behind copyright’s limits should not be modified merely because of new technological affordances. Copyright law should be based on an architecture of control that prioritizes the public values the Constitution aspires to, like enforced consent fair use. Of course, I do not seek to re-litigate the merits of a historically institutionalized system (American market-oriented copyright). Instead, I point out that when publishers make use of automated takedown notices, search-based and otherwise, they are reflecting a set of priorities embedded in the DMCA’s architecture of control that academic authors may not benefit from or might disagree with its consequences. This work seeks to champion their right to take an informed personal stance in the matter and make sure their publishers and reporting agents truly

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

represent their general position when it comes to the DMCA's takedown system. If academic authors had the option of opting out from the automated takedown system their publishers are enrolling them in (not as a case by case basis, but as a part of the general contractual offer), this autonomy would be returned to them so each can decide if an architecture of perfect access control is the most appropriate copyright protection for their work.

In fact, if a comparison of the brief history of copyright against the history of human creativity is made, the result is that authors have been willing to produce their works *despite* little to no copyright protection, and despite the difficulties in its enforcement, let alone the recent aspiration to perfect control. Copyright protection as an incentive to invest and distribute books and papers has been clamored for by publishers (Litman 2006, 104). The DMCA's takedown system gives publishers and the reporting agents they hire on authors' behalf nearly perfect control of distribution through legally supported technological advances and automated monitoring. Unlike encryption and digital rights management, fair use of information traditionally comes as a right once access has been obtained through purchase. However, as the open access debate questions, we should also be considering the matter of increasing accessibility alongside other legal values in copyright like fair use, the role of public funding in scholarly diffusion, etcetera.

Furthermore, e-book sales have kept growing (Milliot 2013) *despite* the "devastating" effects of digital piracy—which in any case is much less insidious than in music, pornography or TV and films, if we use takedown rates to measure the perception of infringement. In the particular case of books, the Association of American Publishers' own data pointed to a "resurgence [...] across all markets — trade, academic²⁸, professional", showing a 5.6% in net revenues (of \$27.9 billion) from 2008 to 2010 (Bosman 2011). This despite one of the major book reporting agents claiming 1.5 million to 3 million daily Google

²⁸ The academic section included professional and scholarly journals and databases, multimedia teaching materials and mobile apps.

queries for pirated e-books, 50% more than in 2009 (Silicon Republic 2010). Substantiation of these sets of claims are hard to obtain since the report has been pulled from the website, but organizations like AAUP trust in the accuracy of the agents their members hire (AAUP 2013), even though they have an incentive to assess their proprietary data in a dire fashion. Google for example, presents another way of viewing such search term growth by contrasting it with non-piratical searches of the same material. For example, they advance that a simple search for “carly rae jepsen call me maybe” that would place authorized websites at the top of the searches occurred 16 times more often than a search that included “mp3” in the term (Google 2013, 20). This would explain why e-books continue to gain market share, as authorized distribution becomes widely available regardless of bestsellers leaking online.

Then again, the impact of free copies on scholarship is in principle not the same as trade books or commercial entertainment, because the goal of authors is to increase readership in the hopes of readership, citations, higher impact rankings and benefits that come with being considered a prominent expert in the field²⁹. In this regard, academic publishing of e-books should be included in policy discussions of open access, and decoupled from the goals of access control and manipulation that the DMCA’s Title II embedded as a first line of defense against digital rights management. But by placing control of their intellectual property in their publishers’ control (non or for profit, university or commercial publishing), authors risk being subsumed into a system created with gatekeeping for trade only in mind, bottlenecking the availability of the work they would benefit from sharing through a takedown mechanism that currently targets the largest search engine as a first stop at ever-increasing rates. As a whole, e-book (trade or academic) sales have continued to grow despite the possibility of un-degraded sharing, even when anti-piracy efforts for this industry are nowhere near comparable to the top three industries. But their

²⁹ The open access movement has centered on public funding for research, arguing that if the public has already paid for the work to be done, they should have access to it. While a layperson’s readership would not contribute to the academic incentives to produce because it does not increase citation, one can argue that the general public also includes the scholastic “global middle class”. As Kelly contends, these expanding research and development markets can also contribute to American and European researchers’ incentives to produce through citations, conference invitations, and etcetera.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

homogenized efforts seek to keep all printed works out of an unauthorized public, undistinguishing of the public's right to fair use for educational purposes, and even their authors' discrete priorities as opposed to other artists who create entertainment commodities.

5 The Making of the Digital Millennium Copyright Act

As part of the architecture of control framework this thesis has used, understanding the DMCA requires examining the motivations, concerns and negotiations between stakeholders, lobbyists and legislators during the process of the law's construction. The following chapter documents in an abridged manner salient moments that shaped the parts of the architecture of control for digital copyright that have been the focus of this thesis. In the High Performance Computing Act (HPCA) of 1991, the US government was eager to shift ongoing "anarchic" internet behavior from "a dirt road" (U.S. Congress 1995, 2) to an "information superhighway". The thinking was that the internet was not yet as content rich as it could be because it lacked the assurances copyright owners needed to go online. In the HPCA Congress coined the term "National Information Infrastructure" (NII), an imagined network used to allude to a future network with robust copyright enforcement capabilities. The goal of creating the NII seemed attainable at first, as the internet had just recently been weaned from military-cum-academic uses to a commercial audience and appeared to be dominated by the amateur realm. But concern over copyright infringement was already present when the HPCA suggested creating an "accounting mechanism" to charge for use of copyrighted material as ubiquitous licensing seemed technically feasible. Micropayment plans were put forward by technologists like Theodor Nelson (Nelson 2004). The HPCA also called for reports to be submitted to Congress, creating an Information Infrastructure Task Force (IITF) to this effect. Congress had often benefited from the conference system in copyright negotiations because it allowed the lawmakers to hear from the Copyright Office after more radical proposals had been floated. These were published in the

Working Group's draft recommendation (known as the Green Paper), and then massaged into the final report (informally called the White Paper) (Litman 2006, 94).

The Task Force's Working Group on Intellectual Property Rights was headed by Bruce A. Lehman, U.S. Assistant Secretary of Commerce and Commissioner of Patents and Trademarks, who spearheaded all DMCA related efforts at home and abroad throughout the years. Lehman's report and work with Congress laid the foundation for the DMCA. In 1995, after carrying out two years of hearings and collecting Requests for Comments, the Task Force issued the Report of the Working Group on Intellectual Property Rights (the White Paper). The report stated that the National Information Infrastructure was to be understood as future digital services not yet developed. Meant to replace the internet, it encompassed "digital, interactive services now available, such as the Internet, as well as those contemplated for the future" (Lehman and U.S. Information Infrastructure Task Force September 1995, 2). This "fantasy digital network" (Litman 2006, 90) was first talked about like a TV, with some interactive shopping, and thereafter imagined as a network where computers do what we can now do with a smartphone: "a telephone, a TV, a camcorder, and a personal computer [...], you could browse the latest additions to the library, or you could find the best prices in town on groceries [...]", etcetera (Commerce 1993).

Copyright policy going forward would seek to address concerns content producers raised after the Home Recording Act of 1992 regarding hardware manufacturers' role in helping enforce their copyrights. This is why a significant part of the negotiations for the DMCA centered on implementing anti-circumvention protections in hardware, and on running software encryption on transmitted data (Title I of the DMCA). Digital rights management (DRM) was to become the cornerstone of the DMCA's architecture of control.

The Task Force was envisioning a client-server system of information retrieval where access controls implemented by ISPs on behalf of copyright owners were the first level of protection. This clearly shifted the burden of policing, so a second level of protection was proposed whereby an electronic file used

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

embedded encryption features³⁰ to authorize or deny any subsequent manipulation of the work by embedding appropriate authority over the content. The model could not account for peer-to-peer downloading because it was not around yet. In this way, the groundwork was laid out for DRM to execute an architecture of control which afforded compulsory licensing.

One of the conclusions the Working Group reached was the idea that the 1976 Copyright Act's definition of transmission had to be updated. As they interpreted the law, transmission required the performance or display of a work. But because a digital file can be transmitted without playing it at any point of the upload or download process, transmission had to be re-defined beyond reproduction if the law was to effectively discourage unauthorized copies of content from digitalizing. Since a performance requires reproducing the content, the Task Force advanced, and a computer "logically and legally" does so within its memory, transmission is a means of distribution, and the new law should reflect it (Lehman and U.S. Information Infrastructure Task Force September 1995, 214).

Transmission liability of copyrighted works was a concern of ISPs, but in 1993 (when the Task Force's mandate began) the real internet was far from the imagined NII TV-like system. The proposed efforts on transmission liability was preemptive, and the service providers that would be required to comply with copyright holders' demands at the time were not at the negotiating table. However, as the internet became commercially available to the public at large, two traditions in communication law came into conflict with one another. On the one hand, publishers of analog content were liable for printing and distributing material that infringed on another's copyright. On the other hand, previous information exchange mediums like the Post Office and telegrams carried a "don't shoot the messenger" principle which embedded privacy rights to their distribution of data. This relieved them of liability for transmitting a package that may contain harmful information, stolen goods, etc. As internet use grew, OSPs like forums and Yahoo, and ISPs like America Online sought relief from copyright liability by alluding to the second tradition, trying to liken

³⁰ These would include rendering software, scrambling encryption, digital watermarking, unique file formats.

themselves to the Post Office against copyright owners' accusation of their complicity in copyright infringement. Meanwhile, copyright owners wanted to treat ISPs and OSPs with the rules publishers abided by which would have required compulsory licensing. Congressional definition of digital service providers' role as transmitters became a pressing issue.

Thus, ISPs and OSPs had to be implicated in any negotiation of responsibility and liability for information management on the network, as they were the conduit to transmission and distribution. As the bill progressed Congress agreed with Commissioner Lehman's advocacy that transmission be equated to the distribution of a work. This way, the Copyright Act of 1976's definition of transmission as requiring the performance or display of work (U.S. Congress 1995, 46) could be altered to transmission without reproduction in order to trigger a copyright violation.

The House and Senate subcommittees on Courts and Intellectual Property held a joint hearing November 15, 1995 to bring contextualizing testimony and expertise on the NII Working Group's final report because proposed legislation from both chambers was based on it. Commissioner Lehman, who authored the White Paper, was the main witness. The White Paper, a product of conference procedures, not only became the basis for legislation on both House and Senate, but the representatives actively sought the Commissioner's stamp of approval and interpretation on the proposals "they" had written, during the hearings. The bills, H.R. 2441 and S. 1284, were only a few pages long, and Commissioner Lehman often repeated that the adjustment to existing copyright legislation should be purposefully minimal, to attend to the most urgently needed tweaks. But the conference leading to the White Paper did not fully work out all the conflicts between emerging and established parties. Protestations persisted over technical locks and control access, and librarians raised the issue of digital fair use throughout, so both chambers of Congress had to deal with the mediation processes openly and directly.

Despite service providers' concerns over their liability for third party copyright infringement, a sense of urgency was lacking in 1995 when Lehman talked to the joint session about digital piracy. As he saw it,

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

the content that abounded on bulletin boards was not content from business that could be monetized (U.S. Congress 1995, 31). Because of the Commissioner's comments legislators assumed a low incidence of infringement online and did not initially see ISPs and OSPs as key players in the negotiations over their liability. This situation would surely change once the NII was developed and commercially viable content was online, he remarked. But the matter of future distribution of copyrighted works had to be settled swiftly if content producers were to be persuaded to add their products to cyberspace—an *a priori* issue in Lehman's mind³¹. A recent ruling over the LaMaccia case³² had concluded that copyright infringement for non-commercial motives could not be prosecuted under the cybercrime and copyright law of the time, as protected by the Audio Home Recording Act (Gillespie 2007, 321)³³. According to Commissioner Lehman the prospect that this could be the lay of the land would discourage content producers from participating in cyberspace, and the digital realm would lose any attractiveness as a marketplace for intellectual property—crushing the prospect for the NII's license scheme. The possibility that the success of the internet as a communication and commerce medium could be at risk was a real concern of lawmakers and the Working Group.

The matter of ISP liability was raised various times during the joint hearing by Congressman Boucher (D-VA), who found it hard to get a clear answer on how the legislation going forward should treat ISP and OSPs according to the Patent and Trademark Office and the Register of Copyrights. The Congressman held throughout the questioning that ISPs and OSPs should not be subject to contributory or

³¹ The Commissioner was concerned, for example, that Microsoft was starting to sell software online without the domestic or international rules being set, which could produce a "very chaotic situation" that would hamper the possibility of "explosive economic growth" (U.S. Congress 1995, 64).

³² In 1994 the David LaMacchia used MIT computers to create a bulletin board system where commercial software could be uploaded and downloaded to. Because he had not acted to obtain private financial gain or commercial advantage, he could not be charged with criminal copyright infringement, which requires establishing such motive. In *United States v. LaMacchia* David LaMacchia was charged by the federal government with violation of the federal wire fraud statute. The statute "prohibits the use of interstate or foreign wires to execute any 'scheme or artifice to defraud'" (Hoag 1995, 509). LaMacchia argued that his was a case of copyright infringement, and applying the wire fraud statute would create such a broad interpretation of the statute as to render it "unconstitutionally vague" (Hoag 1995, 510). Judge Stearns agreed and moved to dismiss the case.

³³ The NET Act, passed in 1997 remediated the loophole pointed by the LaMacchia ruling when it created the right to prosecute against individuals accused of copyright infringement even when no commercial benefit occurred.

vicarious infringement liability but treated as telephone companies³⁴ (U.S. Congress 1995, 27, 28). He held that since monitoring every upload would be economically unfeasible, in the absence of actual knowledge of specific infringement, they should be exempt from liability. This position would be triumphant in the DMCA, but Commissioner Lehman and content owners demurred in showing open support for it because they sensed their most of their proposals were not being as well received by the House as in the Senate.

After the joint session, the House remained dissatisfied with the clarity of the language on service provider liability. Copyright owners were also unsatisfied because it seemed like support for secondary liability against ISPs was waning among the 104th Congress' lower chamber (U.S. Senate 1998, 4) as the proposed bills for the NII Copyright Protection Act stalled. Copyright owners contended that measures to incentivize internet traffic monitoring should not be discarded because ISPs were in a better position to track transmission of piracy than they were, but would be unwilling to act without a provision that made them liable. Yet, in 1997 internet adoption rates were proving that the internet would not wait for an imaginary, NII-idealized structure to bring "copyright worthy content" online. During this year the House also found itself under the looming threat of being squeezed out of the copyright legislation procedure altogether by the Senate. This chamber of Congress was working on the ratification of the World Intellectual Property Organization (WIPO)'s Copyright Treaty and the Performances and Phonograms Treaty. The accord dealt with digital rights management at the international level (which would turn into the DMCA's Title I). The House thus decided to resume hearings on service provider liability in order to introduce the On-Line Copyright Infringement Liability Limitation Act (OCILLA, which would become the DMCA's Title II) and contribute to the national process to update copyright legislation.

During the first version of OCILLA negotiations (H.R. 3209) California's Congresswoman Lofgren (D-CA) revived concerns that the language did not protect OSPs such as the Yahoo! search engine from

³⁴ Telephone companies enjoy a "common carrier" status that, while forcing them to connect subscribers of other companies to their customers, conveys the same liability exclusions the Post Office has when transmitting packages (the "don't shoot the messenger" principal, as this paper has called it).

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

unknowing liability. Content owners lobbied the House again, arguing that service providers actually had more control over their networks than they claimed, and thus any exemption would be too generous (Congressional Quarterly 1999, House Panel Moves Copyright Bills; Debate Escalates, p. 3). As the issue of ISP and OSP liability began commanding more attention from the House, these participants' voices grew in their ability to reach the negotiating table. Furthermore, local ISPs began consolidating into larger companies or merged with telephone companies, which made Congress grant them an opening in the negotiating table as the House sought out their approval of the bill.

The deal the House struck included a provision where an ISP could claim it, like telephone companies, was merely a conduit for unlawfully obtained copyrighted material. The amendment would turn into subsection (a) of the DMCA's Title II. ISPs and OSPs also negotiated protections for themselves if they took down a file that was later found not to be infringing copyright. Thus service providers which removed content or delisted it from searches at the behest of a third party could not be sued by the uploader down the road. This language is included for in subsections (c) and (d) in the final Act. Service providers which were unwilling to cooperate with copyright holders' demands would lose their exemption from liability over an uploader's actions. OSPs contended that monitoring all posts would be economically prejudicial and asked copyright owners to take over surveillance of their own content since they had the most to lose from an OSP keeping tabs inefficiently. In return, OSPs assured rapid rates of responsiveness, and to not question copyright owners' individual claims if their forms met the basic legal requirements for an acceptable submission. This way, the access controls service providers implemented would not have to consider fair use or other complex legal exceptions to copyright law. The architecture of control laid out retained a mostly private character, as an uncontested takedown would not have to go through a court review to be put into effect. This made the mechanism attractive and cheap for content owners to deploy in the future *en masse*. Although service providers presently have better lobbying power and can dictate some of the technical procedures left ambiguous by the legislation, when carrying out takedowns service

providers still avoid eliciting as much discontent as possible from content providers without actively monitoring and filtering all content posts, thus maintaining the bargain struck between content owners and service providers during DMCA negotiations. This can be corroborated by Google's average success rate of over 90%, and their efforts to maintain a responsive rate of under an 8 hour workday shift for the thousands of requests it receives.

After the House's negotiations, by the start of 1998 Senate Judiciary Committee Chairman Orrin Hatch (R-UT) held hearings and conferences with chairmen and CEO's of an extensive list of vested interests, as well as with some groups that had opposed the past years' proposals³⁵. The Senate incorporated language from other House bills related to technology, such as the computer maintenance authorization provision (DMCA's Title III) and new copyright protections for nautical ship hull design (Title V) into the Senate's proposed bill (S. 1121). After hearing from America Online and smaller ISP's, Senator Ashcroft proposed and passed an amendment akin to H.R. 3203 (OCILLA) which dealt with limited liability for service providers. During that time the content industry voiced discontent over university and library fair use provisions, alleging academics' demands would hurt their economic interests (Congressional Quarterly 1999, 6, 7). To the dismay of content owners, the House Commerce Committee proposed an amendment to oversee the content industry's allegations. The Senate intervened by asking all parties desist from bringing up fair use in the legislation (Litman 2006, 137-141), which is why the DMCA makes no mention of the matter.

³⁵ These meetings included the outspoken Motion Picture Association of America (MPAA) chairman and CEO Jack Valenti, music producer BMI, the National Music Publishers Association (inaugurators of the conference system in 1909), the Business Software Alliance, the Computer & Communications Industry Association, McGraw Hill, Home Recording Rights Coalition (a consumer protection lobby), Consumer Electronics Manufacturers Association, Activities Board Institute for Electrical and Electronics Engineers, U.S. National Commission on Libraries and Information Science, American Printing House for the Blind, Scholastic Corp. for the Association of American Publishers; Association of American Universities, CompuServe, HLC-Internet, William, Brinks, Hofer, Gilson & Lione lawyer firm, International Trademark Association, the American Society of Composers, Authors, and Publishers (ASCAP, another legacy participant), America Online, Inc., the Digital Future Coalition (a new coalition of business and nonprofits unhappy with the White Paper); and Daniel Government Relations, Novell, Inc. (U.S. Senate 1998, 3, 4).

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

The Senate's final proposal included the aforementioned House resolutions but was able to strike a compromise. It offered content owners language that did not involve a House Commerce Committee hearing. In return for not adopting the hearing proposal, content owners finally agreed to the liability exceptions for ISP and OSPs. The exceptions would still require service providers to be unaware of a particular user's copyright violation, and upon being notified of infringement, they agreed to shut down violators, remove the content, and turn over identifying information where the law required it. Importantly, content owners would not need to substantiate a copyright infringement claim in court in order to trigger service providers' cooperation. Service providers also received immunity from being sued by subscribers if the material had been taken down or if their service had been terminated due to an erroneous complaint.

As the bill went into committee and the election recess loomed, both parties were in a hurry to pass a copyright bill. The odd combination of the DMCA titles, ranging from DRM technology to digital copyright, to the provision ensuring that fixing someone's computer will not get a repairperson in jail, to ship design copyright, these measures were cobbled together into a "hodgepodge" legislation (Litman 2006, 143) that combined the side deals the House and Senate made. Yet the most notable effects of the DMCA were Title I's anti-circumvention requirement³⁶, and Title II's liability exceptions for internet and online service providers through takedown notices.

The fact that the DMCA contains Title II's liability exclusions is an effect of the time between the 1995 joint hearing and the 1998 passage of the bill. During this period ISPs in particular gained negotiating power. By 1997 18% of US households had internet access, and the figure more than doubled by the year 2000³⁷ (U.S. Census Bureau 2012, Table 4. Households With a Computer and Internet Use: 1984 to 2012). During the 1995 hearings the matter had been marginal compared to the effort spent laying the legal

³⁶ Title I makes participating in any way in the distribution of a service or device meant to frustrate a copyright owner's locks on access to their product illegal.

³⁷ In 1993 the U.S. Census Bureau, Current Population Survey, asked about home computers, but not internet. In 1997 half of computerized households had internet access.

groundwork to justify anti-circumvention measures and equating transmission to distribution. The rise of an industry—internet and online service providers— made an outlier issue which was downplayed in the first draft version by Commissioner Lehman second only in importance to the anti-circumvention crux of the final copyright bill. The expedient nature of takedown notices, which are executed without individual court reviews reflect the lobbying power of the media industry on Congress. Through this mechanism, the legislature vested content owners with vast *de facto* decision-making power in copyright enforcement. Because the architecture of control was legislated as a means of preparing internet users for a compulsory license system (the NII), the takedown mechanism trusts copyright holders to make judicious claims of copyright infringement, instead of having a court review them. Furthermore, the legislation also requires their demands be attended to expeditiously, even though content owners have the most to gain from the narrowest applications of fair use, the first sale doctrine, and nothing to gain from less stringent access control on academic works. This ensures that copyright's current architecture of control stay skewed towards companies that can issue high volumes of takedown notices, thus commanding the attention of the receiving service providers.

Conclusion: Sixteen years of the DMCA's architecture of control

A historical view of the Digital Millennium Copyright Act yields the interesting observation that the legislation was crafted for the goal of establishing a compulsory license network called the National Information Infrastructure (NII) to replace the internet. Its purpose would be to create a content producer-friendly network that would track and enforce copyright, effectively ending the first sale doctrine and fair use. Commissioner of Patents and Trademarks Bruce Lehman assured internet and online service providers that they were already covered against copyright infringement liability by existing law while at the same time assuaged media makers that his first priority was to create an idealized version of the internet that would afford them the highest level of access control over owners' property. But as internet and online

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

service providers grew as an industry, they gained more weight in negotiations and lobbied Congress for explicit liability exceptions. The bargain struck between service providers and copyright holders became an architecture of control that affords low-scrutiny and is cost-effective to content owners' and their representatives' accusations, while providing little tools for those accused to receive notice and ask service providers to put back the content. Service providers were relieved of having to actively monitor all internet postings, but are under implicit pressure to accelerate the speed and success rates of the thousands of automated takedown notices the largest complaining parties send.

The DMCA also inaugurated an era of legalized technical measures to track and censor internet unauthorized postings and make illegal any service or device that would remove digital rights management, regardless if the aim was grounded on fair use or piracy. As Google's role in managing internet content access has grown, so has its power to make decisions one way or another in the effectiveness of takedown notices. By favoring web based requests and even .txt files that can be created and implemented without human assistance, Google has afforded bulk reporting in growing numbers. The legislation also allows for reporting agents to act on behalf of content owners, leading to the creation of a new industry chartered with defending the copyrights of all their clients in a homogenous fashion.

The measure of the volume of piracy post-DMCA is still hard to grasp, partially because its definition is still not internationally consistent. This is because the standard by which we measure infringement is defined by nations—as chapter 5 makes explicit—although, as I've argued, efforts to internationalize standards are always at work. Examples of such standards are temporal terms of protection, whether commercial gain is a requisite motive in defining piracy, the limits of fair use, the applicability of the first sale doctrine in digital goods, etcetera. In this sense, the DMCA has been an effective tool in internationalizing an American definition of what is and isn't piracy, which is a precondition to measuring it.

Google, the DMCA and the most avid issuers of takedowns are mostly American. A few of the largest search-based takedown issuers are not: BPI and the Publishers Association are British, Stitching BREIN is Dutch and Bundesverband Musikindustrie is German; but they are all bound by the European Union's E-Commerce Directive which imitates the US' safe harbor provisions for OSPs if they take down content accused of infringement (Google 2013, 14). However, downloaders may be Brazilians—a country where piracy is still defined in terms of commercial profit. In that case, if a downloader obtains an unauthorized (as defined by American law) version of a movie, watches it privately and deletes it, under Brazilian law he is not at fault. Instead, consider defining the matter of piracy from the point of view of the content industry, as established above often under American jurisdiction. Then, the DMCA copyright control has been able to stem piracy because the internet's largest service providers and content owners have internationalized American standards of defining piracy. When the American government seizes the domain and servers of a website, or the content industry takes to American court an uploader, or reporting agents make a website's pages devoid of content and unreachable through §512(c) and (d) takedowns, the nationality (and domestic definition of piracy) of the website, the uploader and the public stop mattering. From this perspective, one can also conclude that insofar as Google's delisting of search results are concerned, the content industry has been successful in 90% of the cases. One should not forget, however, that these actions bottleneck casual downloaders' access to allegedly infringing content. This particular takedown mechanism does not dissuade uploading, nor does it remove media from the internet.

In looking at how takedown notices have been deployed, the overarching trend has been one of concentration: less than 50 websites in 60 trillion addresses have received over a million requests to take down individual URLs, and the 29 largest reporting agencies have sent 80% of all notices, though in 2005 they accounted for only 1.3% of § 512(d) notices. Another longitudinal trend is that both for notices (which measures the frequency of complaints) and requests (individual flagged URLs), the figures show dramatic increases in their deployment. In 2012, requests grew 20 times the previous years' sum, and from 2012 to

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

2013 six-fold. Methodologically, I have advanced that requests are a better metric than notices in the study of takedowns since the latter can hide thousands of instances or can inflate participation rates. For example, individuals, marked by Google as unspecified reporting agents, are 5th by notice frequency, but submit only 4.8% of all URLs, and their efforts result in only 0.2% of the removals. This last number shows how professional effective removals are in Google's system. Because of the rise in the use of takedown notices, Google has begun issuing "trusted" status in their Trusted Copyright Removal Program, adding an express queue to solve the most active complaining parties' requests. Membership to the program has also expanded. But copyright owners have continued to voice their discontent at the takedown process, hoping they might be awarded more discretionary powers to take down whole websites, and avoid having to continually monitor websites they believe are solely online for the purposes of infringing their content. Finally, because there is no legal requirement for ISPs and OSPs to engage in notifying potential counter-notice senders that their content has been delisted from search engines or removed from their servers, these instruments are currently issued in less than 1% of §512(d) takedowns.

The data also presents publishing as poorly represented in the fight against digital piracy as a proportion of the most active reporting agents (see table 4). Yet regardless of claims of the persistency of piracy, sales in trade and strong figures in academic publishing remain. Currently, the two largest reporting agents seemed to have so effectively removed URLs from Google search that their automated systems are reporting delisted links that continue online, but unlisted in Google's indexes (see table 6).

Finally, this thesis questions whether academic publishers benefit from a copyright architecture that affords perfect control in an attempt to foment a compulsory licensing online environment. Academic publishers have deployed the same takedown services used to resolve unauthorized distribution of their authors' content that other media industries use. I argue, however, that academic authors are part of a different production incentive structure than other content creators. The incentive structure does not disregard personal concerns for a decent quality of life and an economically secure future. Rather,

economic security results from activities where production is hardly ever measured in terms of units of research sold. Instead, scholars are assessed by their impact in a field through reviews, publications, citations, conference participation, grants, awards, and the like, result in a hard to quantify metric of prestige. In turn, prestige informs fellowships and hiring decisions, salaries, tenure, productivity stipends. Salaried scholars are expected to engage with their peers' production in a way artists are not, by collectively filtering good research from bad research and advancing the field. In the movie industry, no one but those directly tied to a movie are positively impacted by a block-buster end product. A director may benefit from another director's development of a new technique, but that benefit does not motivate the first director to produce it.

Thus, the disparate nature of production incentives for most scholars suggests automated search-based takedowns are an issue that needs to be taken up by the stakeholders themselves. Furthermore, it is time researchers are given the choice to act on their concerns, and even self-interested motivations, in affording the scholastic "global middle class" access to their work if they find this inclusiveness is worth reorienting the DMCA's architecture of control.

Works Cited

- AAUP. 2013. *AAUP Member Presses*. May 8. Accessed July 5, 2014. <http://www.aaupnet.org/aaup-members/membership-list#u>.
- . 2010. *AAUP Response to OSTP Request for Public Comment*. January 2010. Accessed July 6, 2014. <http://www.aaupnet.org/images/stories/documents/OSTPresponse12110.pdf>.
- . 2013. *About University Presses*. Accessed June 17, 2014. <http://www.aaupnet.org/about-aaup/about-university-presses>.
- . 2013. *Policy Areas' Copyright & Access: Litigation & Legislative Action*. May 08. Accessed July 06, 2014. <http://www.aaupnet.org/policy-areas/copyright-a-access/copyright-a-permissions/current-litigation-a-legislative-action/google>.
- . 2013. *Policy Areas' Copyright & Access: Online Piracy*. May 8. Accessed July 7, 2014. <http://www.aaupnet.org/policy-areas/copyright-a-access/copyright-a-permissions/online-piracy#3>.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

- Anderson, Nate. 2008. *IFPI: "Three Strikes" Efforts Hit Worldwide Home Run*. August 19. Accessed April 1, 2014. <http://arstechnica.com/tech-policy/2008/08/ifpi-three-strikes-efforts-hit-worldwide-home-run/>.
- Andy. 2014. *Google Processes Millions of Useless DMCA Notices*. July 15. Accessed July 23, 2014.
- Bosman, Julie. 2011. "Publishing Gives Hints of Revival, Data Show." *New York Times*, August 9: C1. Accessed July 7, 2014. <http://www.nytimes.com/2011/08/09/books/survey-shows-publishing-expanded-since-2008.html>.
- Brown, Bob. 2010. *Sony BMG Rootkit Scandal: 5 Years Later*. November 1. Accessed July 26, 2014. <http://www.networkworld.com/article/2194292/network-security/sony-bmg-rootkit-scandal--5-years-later.html>.
- Buckles, Brad. 2012. *Some Clear Facts About Google's "Transparency" Report*. Edited by Recording Industry Association of America (RIAA) Anti-Piracy Executive Vice President. May 30. Accessed June 28, 2014. http://www.riaa.com/blog.php?content_selector=riaa-news-blog&blog_selector=Clear-Facts-&blog_type=&news_month_filter=5&news_year_filter=2012.
- Chilling Effects Clearinghouse. 2014. *Chilling Effects: Weather Reports - In addition, we want your help*. April 1. Accessed April 08, 2014. <http://chillingeffects.org/>.
- . 2011. *Participants and Contacts: About Us*. April 11. Accessed April 8, 2014. <http://chillingeffects.org/about>.
- Clark, Liat. 2013. *Pirate Cull: UK Court Orders ISPs to Block 21 File-Sharing Sites*. October 29. Accessed June 30, 2014. <http://www.wired.co.uk/news/archive/2013-10/29/21-pirate-sites-block>.
- Congressional Quarterly. 1999. "Lawmakers Update Nations' Copyright Law For the Digital Age." *CQ Almanac 1998*, 54 ed.: 11. Accessed April 23, 2014. <http://library.cqpress.com/cqalmanac/cqal98-0000191050>.
- DePillis, Lydia. 2013. *Everything You Need to Know about the Trans Pacific Partnership*. December 11. Accessed July 31, 2014. <http://www.washingtonpost.com/blogs/wonkblog/wp/2013/12/11/everything-you-need-to-know-about-the-trans-pacific-partnership/>.
- Elkin Koren, Niva. 2007. "Making Room for Consumers Under the DMCA." *Berkley Technology Law Journal* 22 (3): 38. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1024566.
- Ernesto. 2011. *Google Now Censors the Pirate Bay, IsoHunt, 4Shared and More*. November 23. Accessed March 30, 2014. <http://torrentfreak.com/google-now-censors-the-pirate-bay-isohunt-4shared-and-more-111123/>.
- . 2013. *Google Relaxes DMCA Takedown Restrictions, Eyes Abuse*. April 11. Accessed June 28, 2014. <http://torrentfreak.com/google-relaxes-dmca-takedown-restrictions-eyes-abuse-130411/>.
- . 2011. *Google Starts Censoring BitTorrent, RapidShare and more*. 26 January. Accessed June 29, 2014. <http://torrentfreak.com/google-starts-censoring-bittorrent-rapidshare-and-more-110126/>.

- . 2014. *Google Takedown Notices surge 711,887 Percent in Four Years*. March 25. Accessed March 29, 2014. <http://torrentfreak.com/google-takedown-notices-surge-140325/>.
 - . 2013. *Microsoft Ditches Anti-Piracy Partner after Embarrassing DMCA Takedowns*. September 27. Accessed April 16, 2014. <http://torrentfreak.com/microsoft-ditches-anti-piracy-partner-after-embarrassing-dmca-takedowns-130927/>.
 - . 2014. *This is How the UK Piracy Warnings Will Work*. May 17. Accessed July 27, 2014. <http://torrentfreak.com/how-uk-piracy-warnings-work-140517/>.
 - . 2014. *UK ISPs and copyright holders praise new piracy warning system*. July 19. Accessed July 27, 2014. <http://torrentfreak.com/uk-isps-and-copyright-holders-praise-new-piracy-warning-system-140719/>.
- Gillespie, Tarleton. 2007. *Wired Shut: Copyright and the Shape of Digital Culture*. Cambridge, MA: The Massachusetts Institute of Technology Press.
- Google. July 2011-Dec 2013. "Google Transparency Report." Downloadable database, Mountain View.
- . 2014. *Googlebot*. Accessed April 04, 2014. <https://support.google.com/webmasters/answer/182072>.
 - . 2013. "How Google Fights Piracy." *Google Public Policy Blog*. September 10. Accessed June 28, 2014. <https://docs.google.com/a/google.com/file/d/0BwxyRPFduTN2dVFqYml5UENUeUE/edit?pli=1>.
 - . 2011. *Transparency Report: FAQ*. July. Accessed April 5, 2014. http://www.google.com/transparencyreport/removals/copyright/faq/#is_the_data_comprehensive.
 - . 2014. *Transparency Report: Overview*. April. Accessed April 11, 2014. <https://www.google.com/transparencyreport/removals/copyright/?hl=en>.
 - . 2013. *Transparency Report: Visible Changes*. October. Accessed April 11, 2014. <https://www.google.com/transparencyreport/removals/copyright/changes/>.
 - . 2000. *Yahoo! Selects Google as its Default Search Engine Provider*. June 26. Accessed July 22, 2014. http://www.google.com/googlefriends/alert2_2000.html.
- Gross, Grant. 2014. *What's the Problem with DMCA Takedown Notices?* IDG News Service. March 20. Accessed March 30, 2014. <http://www.itworld.com/it-management/410814/whats-problem-dmca-takedown-notices>.
- Hoag, Aaron D. 1995. "Defrauding the Wire Fraud Statute: United States v. LaMacchia." *Harvard Journal of Law & Technology* 8 (2): 12. Accessed July 27, 2014. <http://jolt.law.harvard.edu/articles/pdf/v08/08HarvJLTech509.pdf>.
- Johns, Adrian. 2010. *Piracy: The Intellectual Property Wars from Gutenberg to Gates*. Electronic, ISBN 9780226401201. Chicago: University of Chicago Press.
- Jones, Ben. 2012. *RIAA Demands Unlimited DMCA Power from Google*. June 2. Accessed June 25, 2014. <http://torrentfreak.com/riaa-demands-unlimited-dmca-power-from-google-120502/>.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

- Karaganis, Joe, et al. 2011. *Media Piracy in Emerging Economies*. New York: Social Science Research Council, 424.
- Kaufman, Joshua J. 2014. *United States: DMCA Takedown? Not Without A Registration*. Venable LLP. March 26. Accessed March 30, 2014. <http://www.mondaq.com/unitedstates/x/302498/Copyright/DMCA+Takedown+Not+Without+a+Registration>.
- Kelty, Christopher M. 2012. *The Disappearing Virtual Library*. March 01. Accessed July 26, 2014. <http://www.aljazeera.com/indepth/opinion/2012/02/2012227143813304790.html>.
- Latour, Bruno. 2009. *The Making of Law: An Ethnography of the Conseil d'Etat*. Translated by Marina Brilman and Alain Pottage. Cornwall: Polity.
- Lehman, Bruce A., and U.S. Information Infrastructure Task Force. September 1995. *The Report of the Working Group on Intellectual Property Rights*. Governmental, Washington D.C.: U.S. Government. Accessed April 18, 2014. <http://www.uspto.gov/web/offices/com/doc/ipnii/ipnii.pdf>.
- Lessig, Lawrence. 2006. *Code: version 2.0. Online version*. N/A: BasicBooks. Accessed January 7, 2014. <http://codev2.cc/download+remix/Lessig-Codev2.pdf>.
- Lessig, Lawrence. 2001. "The Architecture of Innovation." *Duke Law Journal* 51 (6): 1783-1801. Accessed July 31, 2014. <http://scholarship.law.duke.edu/dlj/vol51/iss6/2>.
- . 1988. *The Laws of Cyberspace*. Vers. Draft 3. March. Accessed July 31, 2014. https://cyber.law.harvard.edu/works/lessig/laws_cyberspace.pdf.
- Litman, Jessica. 2006. *Digital Copyright*. Amherst: Prometheus Books.
- MarkMonitor AntiPiracy. 2014. *Shut down the online distribution and promotion of pirated digital content*. Accessed June 12, 2014. <https://www.markmonitor.com/services/antipiracy.php>.
- Masnick, Mike. 2012. *Funny How Copyright Holders Only Ramped Up Google DMCA Takedowns After SOPA Failed*. December 13. Accessed November 20, 2013. <https://www.techdirt.com/articles/20121212/22445321369/funny-how-copyright-holders-only-ramped-up-google-dmca-takedowns-after-sopa-failed.shtml>.
- McGuigan, Glenn S. 2014. "Hateful metrics and the bitterest pill of scholarly publishing." *Prometheus: Critical Studies in Innovation*, May 21: 249-256. Accessed June 13, 2014. doi:10.1080/08109028.2014.891711.
- Meyer, David. 2013. *Microsoft down to fifth place in comScore's global search stats, thanks to Yandex*. February 06. Accessed March 14, 2014. <http://gigaom.com/2013/02/06/microsoft-down-to-fifth-place-in-comscores-global-search-stats-thanks-to-yandex/>.
- Milliot, Jim. 2013. *Trade Books a Winner in 2012*. May 20. Accessed July 05, 2014. <http://www.publishersweekly.com/pw/by-topic/industry-news/bookselling/article/57313-trade-books-a-winner-in-2012.html>.

- Neave, Guy R. 2012. *The Evaluative State, Institutional Autonomy and Re-engineering Higher Education in Western Europe: the Prince and his Pleasure*. New York: Palgrave Macmillan.
<http://books.google.com/books?id=Jcvy8snsqv0C>.
- Nelson, Theodor Holm. 2004. *A Thought for your Pennies: Micropayment and the Liberation of Content*. May 28. Accessed July 27, 2014. <http://transcopyright.org/hcoinRemarks-D28.html>.
- Publishers Weekly. 2014. *Monthly StatShot, March 2014*. July 18. Accessed July 31, 2014.
<http://www.publishersweekly.com/pw/by-topic/industry-news/financial-reporting/article/63373-monthly-statshot-march-2014.html>.
- . 2014. *The Weekly Scorecard: Tracking Unit Print Sales for Week Ending July 20, 2014*. July 25. Accessed July 31, 2014. <http://www.publishersweekly.com/pw/by-topic/industry-news/bookselling/article/63472-the-weekly-scorecard-tracking-unit-print-sales-for-week-ending-july-20-2014.html>.
- Remove Your Media Clients. 2013. *Solutions*. December 14. Accessed June 12, 2014.
<http://www.removeyourmedia.com/solutions/>.
- Rojek, Chris. 2005. "P2P Leisure Exchange: Net Banditry and the Policing of Intellectual Property." *Leisure Studies* (Routledge Taylor & Francis) 24 (4): 357-369. Accessed December 16, 2013.
doi:10.1080/0261436052000330438.
- Seng, Daniel. 2014. "The State of the Discordant Union: An Empirical Analysis of DMCA Takedown Notices." *Virginia Journal of Law and Technology* 56. doi:<http://dx.doi.org/10.2139/ssrn.2411915>.
- Seville, Catherine. 2006. *The Internationalisation of Copyright Law: Books, Buccaneers and the Black Flag in the Nineteenth Century*. Cambridge: Cambridge University Press.
- Silicon Republic. 2010. *Research Reveals Increase in E-book Piracy Demand*. July 10. Accessed July 7, 2014. <http://www.siliconrepublic.com/new-media/item/18192-research-reveals-increase-i>.
- Sullivan, Danny. February. *Google Still World's Most Popular Search Engine By Far, But Share Of Unique Searchers Dips*. 2013 11. Accessed March 07, 2014. 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), <http://searchengineland.com/google-worlds-most-popular-searchengine-148089>.
- The Publishers Association. 2014. *Publishers Association Welcomes Pressure on Google over Piracy*. May 30. Accessed July 6, 2014.
http://www.publishers.org.uk/index.php?option=com_content&view=article&id=2731:pa-welcomes-pressure-on-google-over-piracy&catid=79:pa-in-the-media&Itemid=1403.
- Toffler, Alvin. 1980. *The Third Wave*. New York: Bantam Books.
- Toffler, Alvin, and Heidi Toffler. 2006. *Revolutionary Wealth*. New York: Knopf.
- Twitter. 2014. *Copyright Notices*. January 1. Accessed April 10, 2014.
<https://transparency.twitter.com/copyright-notices/2013/jul-dec>.

An Architecture of Control: Automatic Takedown Notices in Academic Digital Publishing

- U.S. Census Bureau. 2012. *Computer and Internet Access in the United States: 2012*. Governmental statistics, Washington D.C.: U.S. Census Bureau, 5.
<https://www.census.gov/hhes/computer/publications/2012.html>.
- U.S. Congress. 1995. *NII Copyright Protection Act of 1995 Joint Hearing before the Subcommittee on Courts and Intellectual Property*. Government Proceedings, House and Senate Committee on the Judiciary's Subcommittees on Courts and Intellectual Property, Washington D.C.: U.S. Government Printing Office, 156. Accessed April 18, 2014.
<https://archive.org/stream/niicopyrightprot01unit#page/n1/mode/2up>.
- U.S. Senate. 1998. *Committee Report on The Digital Millennium Copyright Act of 1998*. U.S. Senate Report, Washington D.C.: U.S. Government Printing Office. Accessed December 19, 2013.
<http://www.gpo.gov/fdsys/pkg/CRPT-105srpt190/pdf/CRPT-105srpt190.pdf>.
- Urban, Jennifer M., and Laura Quilter. 2005. "Efficient Process or Chilling Effects - Takedown Notices under Section 512 of the Digital Millennium Copyright Act." *Santa Clara Computer & High Tech. Law Journal* (HeinOnline) 22: 621-694. <http://scholarship.law.berkeley.edu/facpubs/501>.
- US Library of Congress. 1998. *The Digital Millennium Copyright Act of 1998: U.S. Copyright Office summary*. Governmental Report, Washington D.C.: Copyright Office, Library of Congress, 18.
<http://www.copyright.gov/legislation/dmca.pdf>.
- von Lohmann, Fred. 2014. "Google's DMCA notice and takedown tools." Vers. Second Public Multisakeholder Forum Meeting on the DMCA Notice and Takedown System. *US Patent and Trademark Office's Policy and International Affairs (OPIA) team*. May 8. Accessed June 25, 2014.
<http://www.uspto.gov/ip/global/copyrights/Google.pdf>.
- . 2013. "Report: How Google fights piracy." *Google Public Policy Blog*. September 10. Accessed June 28, 2014. Introduction post at <http://googlepublicpolicy.blogspot.com/2013/09/report-how-google-fights-piracy.html>.
- Westley, Brian R. 2014. *United States: Google, Viacom Settle Copyright Infringement Dispute That Centered On DMCA 'Safe Harbor' Provision*. Finnegan, Henderson, Farabow, Garrett & Dunner, LLP. March 25. Accessed March 30, 2014.
<http://www.mondaq.com/unitedstates/x/302074/Copyright/Google+Viacom+Settle+Copyright+Infringement+Dispute+That+Centered+On+DMCA+Safe+Harbor+Provision>.