Infrastructures of Censorship and Lessons from Copyright Resistance

Wendy Seltzer
Princeton Center for Information Technology Policy

Abstract

U.S. policymakers proclaim their commitment to Internet freedom while simultaneously endorsing restrictions on Internet exchange. Unfortunately, the tools – legal and technical – built to block copyright infringement, counterfeit sales, online gambling, or indecency, often find use to censor lawful expression here and abroad. In particular, the United States and its entertainment industries have prioritized online copyright enforcement such that its attack and riposte can be instructive in the Internet freedom arena.

1 Copyright as Information-Control

The United States Internet is largely free from government-mandated censorship. The 1997 *ACLU v. Reno* set an early bar, striking as unconstitutional provisions of the Communications Decency Act that would have required Internet Service Providers to block children's access to materials deemed "harmful to minors."[2] The First Amendment, the Supreme Court held, forbade these restrictions on speech. While parents in their homes (and later, libraries and schools operating with federal funds) might filter their children's Internet connections, a law mandating ISP-controlled blocking was not "narrowly tailored" to government purposes.

Copyright, however, stands as one of the rare permissible restrictions on speech. As the Court said in *Eldred v. Ashcroft*, copyright is an "engine of free expression," and therefore, "The First Amendment securely protects the freedom to make – or decline to make – one's own speech; it bears less heavily when speakers assert the right to make other people's speeches." [8]

While numerous scholars [16, 22, 21, 26] and litigants [10, 12] have criticized copyright's seeming free-pass from First Amendment scrutiny, its anomalous information-control has persisted. In response, technologists and hackers have joined the academic and legal critics of copyright.

The history of copyright enforcement measures and counter-measures thus provides a domestic analog and preview of Internet censorship in other contexts.

1.1 Squeezing Filesharing

Online copyright debates took hold in the mid 1990s, as Internet connectivity spread, "rippers" and MP3 compression enabled the public to extract and save digital tracks from music CDs, and sites arose to help people exchange music. Early music-sharers operated through central servers, depositing files and retrieving others from BBSs, FTP servers, and websites. Beyond simple file-exchange, My.MP3.com recognized CDs from a user's drive and transferred copies of their tracks to an online virtual "locker." As all of these methods involved copying, the music studios successfully argued that the unauthorized reproductions of their copyrighted works infringed copyright. [7, 5] Centralized architecture made these early sites easy to find and squash.

Napster claimed both technical and legal innovation when it was released in 1999. The peer-to-peer software distributed the burdens of file storage and the sharing activity, directing peer users to transfer files to one another so Napster itself never copied the files. Yet the Ninth Circuit found that architecture insufficient to avoid copyright liability. Because the company maintained a central directory of files and routing information, its owners were liable for contributory and vicarious infringement of copyright: Napster knowingly materially contributed to infringement, and it profited from infringing activity it had the right or ability to control. [6]

The next generation of peer-to-peer software decentralized further still: Morpheus, KaZaA, and Grokster moved the directory and routing information to supernodes nominated from among peer computers, requiring only a bootstrap download to join the network. After the Ninth Circuit found this architecture escaped Napster's

secondary liability, the Supreme Court attached a new form of liability, yet more indirect, for "inducement":

> We hold that one who distributes a device with the object of promoting its use to infringe copyright, as shown by clear expression or other affirmative steps taken to foster infringement, is liable for the resulting acts of infringement by third parties. [11]

Each of these legal attacks targeted a central corporate entity. Each subsequent round of software reduced the role of that fixed point until its architecture completely distributed its functions. Thus even as Grokster the company was found liable for copyright infringement and put out of business, its software could continue to work. The current leader among file-sharing technology, BitTorrent, launched not as a network but a protocol, leaving users to create their own networks around torrent trackers and their files. BitTorrent, Inc. has avoided litigation through more judicious corporate advertising emphasizing its software's substantial non-infringing uses. It is widely used for distribution of free and open source software, for the authorized distribution of music and movies, and for unauthorized, copyright infringing, reproduction.

Mid-way through the cat-and-mouse game with the companies promoting peer-to-peer, U.S. entertainment industries looked to other solutions to their "piracy" problems: targeting technologies and their individual end-users.

1.2 Regulating Technology

The Digital Millennium Copyright Act's anticircumvention provisions forbid using or trafficking in technologies to circumvent technological measures controlling access to or copying of copyrighted works. Although the technological locks in widespread use have been broken relatively quickly (from DeCSS through Fairplay to the PS3 hack), anticircumvention continues to restrict technological innovation around media. [26] Attempts to silence dissemination of the keys or code to decrypt and thereby circumvent mass-media encryption have often spurred even more gleeful, widespread distribution. [28, 17] Anticircumvention thus constrains the industrial producers and modes of public development, but it does little to restrain "pirates."

Even the law's built-in exceptions are one-offs. The statutory provision for triennial rulemaking for exemptions by Librarian of Congress excuses only the act of circumvention, not the distribution of circumvention tools. [3] So when the 2010 rulemaking exempted "jailbreaking" smartphones that are locked against software installs or carrier-switching, that exemption could not extend to permit third parties to offer jail-break tools or services.

The legal threats force disaggregation, as anyone too large and central in an information exchange faces legal challenge, costly even if he could eventually win. The advantage shifts to smaller, more nimbly distributed actors who trade mass reach for persistence.

1.3 Chasing End-Users

Chasing down individuals costs more effort and expense than using centralized proxies, so when copyright holders take this strategy, they often couple it with press trying to amplify the lawsuits' impact.

In 2003, record companies launched their first copyright lawsuits against individual filesharers. First they had to obtain names. While most filesharers at this time were not taking active steps to anonymize their activity, only their ISPs had identity information beyond IP address and self-chosen username; their ISPs had to be made to provide the link from IP to individual (and the individual so identified might still not be the person responsible for the filesharing). Early attempts to centralize this enforcement failed: ISPs defeated subpoenas demanding hundreds of names at once under §512(g) on the grounds that peer-to-peer activity was not covered there, only ISP-based hosting [9]; individuals fought the lumping together of hundreds of "John Does" in venues far distant from their alleged wrongdoing. Once they addressed these procedural aspects, the RIAA members had to sue in smaller batches, paying separate filing fees for each, but they filed complaints against more than 30,000 before ending this phase of their campaign. [20]

These end-user pursuits are not just time consuming, they're often wrong. ISP or complainants' records may not be sufficiently detailed to match an IP to its user at the time of alleged infringement. An IP address is not a person – even if the match is made to its subscriber, the subscriber may not be the one using the connection. Famous early cases targeted computer-less grandmothers and even dead people. Moreover, as a University of Washington team showed in threats sent to their networked laser printer, IP addresses may be incorrectly identified with infringement. [23]

Copyright holders or their representatives have also taken out-of-court measures. In 2003, as it started suing, the RIAA also invited individuals to join an "amnesty" program, that it dropped a year later. More recently, Righthaven, US Copyright Group, and ACSlaw (UK) built businesses around copyright settlements, notwithstanding challenges to the legal validity of their complaints. Some copyright claimants and agents have engaged in extra-legal activity, attempting to disrupt file-

sharing through spoofed files and faked peers.

These user-focused copyright-attacks prompt responses aimed at dispersing the points of identification: use of anonymous channels, development of private and so-called "small world" networks connecting users to vetted and trusted peers, use of ephemeral pointers. Users create spoof lists, user-generated IP address blacklists and reputation systems to warn of spoofed files and block connections to adversaries masquerading as peers.

1.4 Making Demands on Intermediaries

Finally, copyright enforcers generalize their attacks, targeting intermediaries and infrastructure providers. Under the Digital Millennium Copyright Act, Internet hosts and search engines ("information location tools") are encouraged to implement "notice and takedown." [4] They are immunized from lawsuits if they take down material in response to notices of claimed copyright infringement – and so providers take-down to avoid the risks of suit, even if they would have faced no liability. As a result, copyright holders find that a takedown demand to providers is sufficient to get material removed from hosting or search results, even though the providers have no active involvement in users' activity, but are necessary intermediaries to communication. These takedowns now number in the thousands a week. [1] Service providers often serve as effective chokepoints because the DMCA shifts their incentives toward takedown. [24] Murky rules, especially around fair use, increase the likelihood of self-censorship shy of lawful conduct. [29]

Enforcers even seek to re-shape networks. The Higher Education Opportunity Act tethers federal education funding to implementation of "technology-based deterrents" as part of a plan to "effectively combat" filesharing on campus networks. Implementing these plans often leaves networks less flexible to student experimentation, including non-infringing and research uses.

1.4.1 Domain Names and Beyond

Domain names appear as another potential chokepoint. Legislation first proposed in late 2010 and re-introduced as the PROTECT-IP Act in 2011 targets sites associated with copyright and trademark infringement through their DNS. If passed, PROTECT-IP will authorize the Attorney General to sue any "non-domestic domain name used by an Internet site dedicated to infringing activities," and to enjoin DNS servers from resolving those names. The bill permits *in rem* suits – against a name itself when its owner cannot be found for U.S. jurisdiction.

Even without getting specific Congressional authority, the Department of Homeland Security's Immigration and Customs Enforcement (ICE) has seized several sets of domain names through the US-based .com, .net, and .org domain registries. Although many of the sites were based outside the US, browsers everywhere found the websites replaced by banners warning that "willful copyright infringement is a federal crime." Among the sites whose domains were seized, sports linking site `rojadirecta.org` had been found not liable for copyright infringement in Spain, where its `rojadirecta.es` domain continues to resolve. Rojadirecta's proprietors have sued for the return of their domains.

Striking at addressing infrastructure does not take allegedly infringing content off-line, and may not even make it much harder to find – if popular, it shows up in search engines with an IP address rather than a domain name. Yet it poses legal and technical problems: stinting due process and breaking unrelated systems unexpectedly. [27, 15] In one response, developers created a quick browser plug-in to list alternate domain names, outside U.S. jurisdiction, as back-ups for redirection in the case of domain seizure. Mozilla reported that it refused a Department of Homeland Security demand to remove the plug-in from its repository. [25]

1.4.2 ISP "Co-operation"

In July 2011, shortly after an OECD high-level meeting on "The Internet Economy" endorsed "limited intermediary liability" and multi-stakeholder co-operation [13] major U.S. ISPs and entertainment industry groups announced a new joint program. Under the popularly-termed "five strikes" regime, ISPs agreed to send copyright alerts to and impose "mitigation measures" on subscribers accused of online infringement. [14] To challenge these measures, which include sharp reduction in bandwidth, the subscriber must pay a $35 fee and close access to his or her network, since the system permits only one-time defense of "unauthorized use of account" and no justifications such as "authorized unsupervised use" that would permit greater flexibility. The entertainment companies seek to do by private contract and "co-operation" what they cannot achieve by public law: to create a chain of private enforcers around every Internet connection. If an IP address does not identify the user, copyright claimants will nevertheless try to hold its assignee *responsible* for its associated activity.

Network users have responses even to these measures, using VPNs to service providers who have not joined the content-control business, and anonymizing networks such as I2P. Yet, the industry hopes, more people will be driven toward authorized services (and it may be the ISPs' participation in the profits from such services that drove them to cooperate in policing).

2 Learning from Copyright Censorship and Anti-Censorship

Copyright and other forms of censorship are clearly not strictly equivalent; their motivations differ: to provide incentive for creative activity, or to support political goals. The lines sometimes cross. Copyright's power may be used for political censorship or to gain commercial advantage, dramatically, in 2008, when broadcasters' copyright complaints caused the removal from YouTube of several McCain for President videos, despite strong fair use defenses to the political use of short clips. [24] Both copyright and political censorship represent attempts by those with power in one sphere to extend it over others; both operate, at the extreme, by disrupting networked communications.

Thus after more than a decade of serious online copyright debate, we can draw lessons from this version of information control. In particular, copyright's history provides a catalog of mechanisms of enforcement, responses, and counter-responses. Further, it provides policy and pragmatic arguments against blunt-instrument information blocking, and a set of well-tested example modes for anti-censorship technologies.

Copyright dissidents – and those wary of being wrongly targeted as such – have tended to respond by spreading and fragmenting their efforts. Instead of centralized servers, whether file directories or DNS, users move to distributed trackers and lists. Unable to trust or depend upon central directories, users develop alternative means of authenticating and verifying the integrity of connections and resources, sacrificing some simplicity for greater resilience.

Along with technical circumvention, copyright dissidents have turned to political and cultural modes, of production as well as demand. Creative Commons offers both an alternative means of licensing copyrighted works – licenses for sharing rather than for control – and a badge marking the author as an active or latent member of a movement. [19] Artists who announce their works under CC license tap into that potential affiliation with an audience, who may share their works via social media, finance them through crowd-sourced means such as Kickstarter, and seek them out because of art *and* politics.

2.1 Patterns

Patterns in copyright censorship include the censors' search for chokepoints and centralization to narrow the locus for liability pressure, use of blocking or filtering at these chokepoints, and delegation of censorship. Copyright censorship shows the lack of transparency inherent to information-control: Block-lists are kept secret lest they become menus of infringing content, but then cannot be examined for accuracy. It displays an expansion of mandate, from content, to technology, to infrastructure, and causes an accompanying loss of generality of the tools for disruptive innovation outside the copyright realm. Information-control spans layers and domains as its deployers try to outpace those they would control.

Censorship in the copyright space is often ineffective at thwarting infringement, but not ineffectual. It distorts the information environment, raises costs for speakers and infrastructure providers, and burdens technology innovation. Moreover, as it moves further from the end-user, from the accused direct infringer into the infrastructure, its operation becomes more difficult to challenge because it lacks transparency. Censorship that spans layers divides its harms from the direct incentive or expertise to oppose it.

Patterns appear in the responses as well: those opposing information-control distribute and decentralize, they encrypt and obfuscate. They make the tools of copyright infringement and of free expression hard to distinguish.

2.2 Salient distinctions

As Biddle et al. describe, their "darknet" hypothesis, that "any content protection system will leak popular or interesting content into the darknet" through which it will be redistributed, starts from a premise that the content to be shared or confined is of mass interest. [17] The joint interest of many creates critical mass and density for efforts to decrypt and circulate popular material. A first-run movie or popular song meets this criterion, an item of political dissent may not – if it sustained mass interest, it might not be in dissent. Unpopular expression may not find enough nodes of interest to gain this foothold. Where a blockbuster movie's advertising helps people to find it either in the theaters or on the darknet, political dissent must build its own buzz: dissidents may have to convince audiences to want to hear something different as well as showing how to find it.

Distributed small-worlds networks are harder to censor, or at least make it more difficult to find and censor all of them, but centralized broadcast media has its place too, to attract mass attention and lend perceived legitimacy to the voices broadcast.

Yet perhaps here too politics can take lessons from copyright by allying the political speech with entertainment. Political speakers share a stake in the infrastructure one layer down, shared with the posters and viewers of cute cat videos [31] and even with their opposition. Cross-domain alliances are not new. Freenet [18] drew on Publius for censorship-resistant publishing, and was adopted by some for copyright-resistance. General-purpose technologies that can be used to spread a view-

point *or* its opposite, and even permit dialogue between them, are more generative [30] platforms than are propaganda sites. A content-agnostic, end-to-end network serves the posters of banal status updates and the reporters on the events of #jan25 in Tahrir Square – and further permits users to migrate from one use to the other as their circumstances warrant.

3 Conclusion

The lessons of copyright as censorship work in both directions: those seeking censorship-circumvention learn from the copyright evaders, while those seeking to censor learn from the copyright enforcers. Interests in free expression, and in the political freedom it supports, should lead us to reject the extremes of copyright censorship too. Even democratically chosen restrictions on information exhibit implementation flaws, unavoidable tendencies to overreach and to squelch expression outside their mandate. Democratic regimes should reject this information-control mode and the tools and examples it gives their non-democratic counterparts.

References

[1] Chilling Effects Clearinghouse. `http://www.chillingeffects.org/`.

[2] Reno v. American Civil Liberties Union, 1997.

[3] Digital Millennium Copyright Act, 17 U.S.C. §1201, 1998.

[4] Digital Millennium Copyright Act, 17 U.S.C. §512, 1998.

[5] UMG Recordings, Inc. v. MP3. Com, Inc., 2000.

[6] A & M Records, Inc.. v. Napster, Inc., 2001.

[7] Arista Records, Inc. v. MP3Board, Inc., 2002.

[8] Eldred v. Ashcroft, 2003.

[9] Recording Industry Association of America v. Verizon Internet Servs., 2003.

[10] 321 Studios v. Metro Goldwyn Mayer Studios, Inc., 2004.

[11] Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd., 2005.

[12] Golan v. Holder, 2010.

[13] Communiqué on Principles for Internet Policy-Making. `http://www.oecd.org/dataoecd/40/21/48289796.pdf`, 2011.

[14] Memorandum of understanding. `http://www3.buzzmakerdev.net/~cci/sites/default/files/Momorandum%20of%20Understanding.pdf`, 2011.

[15] Security and Other Technical Concerns Raised by the DNS Filtering Requirements in the PROTECT IP Bill, 2011.

[16] BENKLER, Y. Free as the Air to Common Use: First Amendment Constraints on Enclosure of the Public Domain. *New York University Law Review 74* (1999), 354.

[17] BIDDLE, P., ENGLAND, P., PEINADO, M., AND WILLMAN, B. The darknet and the future of content distribution. In *ACM Workshop on Digital Rights Management* (2002).

[18] CLARKE, I., SANDBERG, O., WILEY, B., AND HONG, T. W. Freenet: A distributed anonymous information storage and retrieval system. In *INTERNATIONAL WORKSHOP ON DESIGNING PRIVACY ENHANCING TECHNOLOGIES: DESIGN ISSUES IN ANONYMITY AND UNOBSERVABILITY* (2001), Springer-Verlag New York, Inc., pp. 46–66.

[19] CREATIVE COMMONS. *The Power of Open*. 2011.

[20] ELECTRONIC FRONTIER FOUNDATION. RIAA v. The People: Five Years Later. `http://www.eff.org/wp/riaa-v-people-years-later`, 2008.

[21] LESSIG, L. Free (ing) Culture for Remix. *Utah L. Rev.* (2004), 961.

[22] NETANEL, N. Locating Copyright Within the First Amendment Skein. *Stan. L. Rev. 54* (2001), 1.

[23] PIATEK, M., KOHNO, T., AND KRISHNAMURTHY, A. Challenges and directions for monitoring P2P file sharing networks-or: why my printer received a DMCA takedown notice. In *Proceedings of the 3rd conference on Hot topics in security* (2008), USENIX Association, pp. 1–7.

[24] SELTZER, W. Free Speech Unmoored in Copyright's Safe Harbor: Chilling Effects of the DMCA on the First Amendment. *Harvard Journal of Law and Technology 24* (2010), 171.

[25] SELTZER, W. In DHS Takedown Frenzy, Mozilla Refuses to Delete MafiaaFire Add-On, 2010.

[26] SELTZER, W. The Imperfect is the Enemy of the Good: Anticircumvention Versus Open User Innovation. *Berkeley Technology Law Journal 25* (2010), 909.

[27] SELTZER, W. Exposing the flaws of censorship by domain name. *IEEE Security and Privacy 9* (2011), 83–87.

[28] TOURETZKY, D. S. Gallery of CSS Descramblers. http://www.cs.cmu.edu/ dst/DeCSS/Gallery.

[29] TUSHNET, R. Copy This Essay: How Fair Use Doctrine Harms Free Speech and How Copying Serves It. *Yale law journal 114*, 3 (2004), 535–592.

[30] ZITTRAIN, J. The generative internet. *Harvard Law Review* (2006), 1974–2040.

[31] ZUCKERMAN, E. The cute cat theory. `http://www.ethanzuckerman.com/blog/2008/03/08/the-cute-cat-theory-talk-at-etech/`, 2008.