ACCOUNTABILITY IN ALGORITHMIC COPYRIGHT ENFORCEMENT*

Maayan Perel** & Niva Elkin-Koren***

CITE AS: 19 STAN. TECH. L. REV. 473 (2016)

ABSTRACT

Recent years demonstrate a growing use of algorithmic law enforcement by online intermediaries. Facilitating the distribution of online content, online intermediaries offer a natural point of control for monitoring access to illegitimate content, which makes them ideal partners for performing civil and criminal enforcement. Copyright law has been at the forefront of algorithmic law enforcement since the early 1990s when it conferred safe harbor protection to online intermediaries who remove allegedly infringing content upon notice under the Digital Millennium Copyright Act (DMCA). Over the past two decades, the Notice and Takedown (N&TD) regime has become ubiquitous and embedded in the system design of all major intermediaries: major copyright owners increasingly exploit robots to send immense volumes of takedown requests and major online intermediaries, in response, use algorithms to filter, block, and disable access to allegedly infringing content automatically, with little or no human intervention.

Algorithmic enforcement by online intermediaries reflects a fundamental shift in our traditional system of governance. It effectively converges law enforcement and adjudication powers in the hands of a small number of mega platforms, which are profit-maximizing, and possibly biased, private entities. Yet notwithstanding their critical role in shaping access to online content and facilitating public discourse, intermediaries are

* We thank Oren Bracha, Miriam Marcowitz-Bitton, Jane Ginsberg, Ellen Goodman, Eldar Haber, Lital Helman, Ethan Katsh, Shelly Kreiczer-Levy, Edward Lee, Neil Netanel, Gideon Pharchomovsky, Orna Rabinovich-Einy, and Tal Zarsky, as well as the participants of the "Trust and Empirical Evidence in Law Making and Legal Process" conference at the University of Oxford on June 19-20, 2015, and the participants of the "Openness and Intellectual Property" conference at the University of Pennsylvania on July 22-24, 2015, for their insightful comments. We are further grateful to Dalit Kan-Dror, Esq., and Nati Perl for their academic assistance. This research was supported by I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation.

** Dr. Maayan Perel, Lecturer, Netanya Academic College; Reserach Fellow, Haifa Center for Law & Technology, University of Haifa Faculty of Law; S.J.D., University of Pennsylvania Law School.

*** Professor Niva Elkin-Koren, Director, Haifa Center for Law & Technology, University of Haifa Faculty of Law.

hardly held accountable for algorithmic enforcement. We simply do not know which allegedly infringing material triggers the algorithms, how decisions regarding content restrictions are made, who is making such decisions, and how target users might affect these decisions. Lessons drawn from algorithmic copyright enforcement by online intermediaries could offer a valuable case study for addressing these concerns. As we demonstrate, algorithmic copyright enforcement by online intermediaries lacks sufficient measures to assure accountability, namely, the extent to which decision makers are expected to justify their choices, are answerable for their actions, and are held responsible for their failures and wrongdoings.

This Article proposes a novel framework for analyzing accountability in algorithmic enforcement that is based on three factors: transparency, due process and public oversight. It identifies the accountability deficiencies in algorithmic copyright enforcement and further maps the barriers for enhancing accountability, including technical barriers of non-transparency and machine learning, legal barriers that disrupt the development of algorithmic literacy, and practical barriers. Finally, the Article explores current and possible strategies for enhancing accountability by increasing public scrutiny and promoting transparency in algorithmic copyright enforcement.

TABLE OF CONTENTS

I. INTRODUCTION

> [A]lgorithms take over from the messy, human process of democratic decision-making. Citizens become beholden to them, unsure of how they work, but afraid to disregard their guidance. This creates a sort of prison of 'invisible barbed wire' which constrains our intellectual and moral development, as well as our lives more generally.[1]

Why is certain content automatically blocked from appearing on certain

1. John Danaher, *Rule by Algorithm? Big Data and the Threat of Algocracy*, PHIL. DISQUISITIONS (Jan. 6, 2014, 9:55 AM), http://philosophicaldisquisitions.blogspot.co.il/2014/01/rule-by-algorithm-big-data-and-threat.html [https://perma.cc/JL3L-V8T4].

online platforms? Why is the same content nonetheless approved on other online platforms? When an online hosting facility receives a complaint of copyright infringement, how does it evaluate the complaint? When Google announces an anti-piracy policy that will push copyright infringers down in the rankings, who is considered an infringer? Do online platforms consider fair uses of copyrighted material? Detailed doctrines in copyright law are carefully designed to guide traditional, human law enforcement agents in addressing these questions. But increasingly, it is mostly algorithms—not humans—that enforce the rights of copyright owners. Does algorithmic copyright enforcement effectively comply with copyright doctrines? Unfortunately, we really do not know.

As evident in anecdotal reports, online algorithmic copyright enforcement is chaotic.[2] It has blocked a ten-year-old boy's self-authored original video starring his LEGO mini-figures and garbage truck despite the fact that he used royalty-free music.[3] It has also facilitated the removal of a home video uploaded by Stephanie Lenz, which featured her children dancing in the kitchen along to Prince's "Lets Go Crazy" song, although the video was obviously protected as fair use.[4] Algorithmic copyright enforcement has also allowed governmental agents in the U.S. Department of Homeland Security to remove a conspiracy video about President Obama.[5] Why would the U.S. Department of Homeland Security issue a takedown if it can't really own a copyright? Content restrictions of this sort may be censoring legitimate speech. Yet because we do not know what criteria enforcement algorithms employ to determine online copyright infringement, it is largely impossible for us to assess their practices. Unable to understand the practices of law enforcement algorithms, we cannot further challenge and correct their flaws.

Copyright law was at the forefront of algorithmic law enforcement beginning

---

2. In a recent complaint against Google Inc., Viacom Inc., and several others, the plaintiff, Benjamin Ligeri, alleged that defendants' use of Content ID unlawfully restricted his content, which he claimed to be protected as fair use. Among other contentions, he argued that

> Content ID is an opaque and proprietary system where the accuser can serve as the judge, jury and executioner. Content ID allows individuals, including Defendants other than Google, to steal ad revenue from YouTube video creators en masse, with some companies claiming content they don't own deliberately or not. The inability to understand context and parody regularly leads to fair use videos getting blocked, muted or monetized.

Complaint at 4-5, Ligeri v. Google, Inc., No. 1:15-cv-00188-M-LDA (D.R.I. May 7, 2015).

3. Dineen Wasylik, *Take Down Abuse: From Harry Potter to LEGOs*, DPW LEGAL: INTELL. PROP. & APPEALS (Feb. 7, 2014), http://ip-appeals.com/take-down-abuse-from-harry-potter-to-legos [https://perma.cc/D5QP-8C7G].

4. Lenz v. Universal Music Corp., 572 F. Supp. 2d 1150, 1151-52 (N.D. Cal. 2008). The uploaded video was a twenty-nine second recording in which Prince's song was heard playing in the background. Stephanie Lenz uploaded the video to YouTube to share it with her friends and family. This use of Prince's copyrighted material was fair because it was completely non-commercial, it used only a small portion of the song, and it did not replace the original song in any manner that could affect the potential market for Prince's original song.

5. Mike Masnick, *Homeland Security Issuing Its Own DMCA Takedowns on YouTube to Stifle Speech*, TECHDIRT (Aug. 1, 2012, 10:06 AM), https://www.techdirt.com/articles/20120720/02530219774/homeland-security-issuing-its-own-dmca-takedowns-youtube-to-stifle-speech.shtml [https://perma.cc/K8TE-UHXC].

in the early 1990s, conferring safe harbor protection to online intermediaries who removed allegedly infringing content upon notice under the Notice and Takedown (N&TD) procedure designed by the Digital Millennium Copyright Act[6] (DMCA). Over the past two decades, N&TD has become ubiquitous and embedded in the system design of all major intermediaries. To confront the immense volume of takedown notices sent by copyright owners[7]—many of which are sent simultaneously and automatically by robots that scan the web for allegedly infringing content, major online intermediaries use algorithms to filter, block, and disable access to allegedly infringing content automatically, with little or no human intervention.[8] Major platforms, such as Google, Facebook, and Twitter, thereby engage in algorithmic copyright enforcement on a daily basis, applying various algorithms to perform qualitative determinations, including the discretion-based assessments of copyright infringement and fair use.[9]

Algorithmic copyright enforcement carries some obvious advantages. It is often more efficient, saving the cost of hiring staff and renting office space.[10] It may further ensure consistency in applying legal doctrines and eliminate the hassle of human review.[11] While online intermediaries occasionally use algorithms to automatically implement the DMCA safe harbor provisions, some go beyond the statutory requirements, using voluntary measures to further block the distribution of infringing materials before they become available online.[12] A classic example is YouTube's Content ID, which can automatically block access to online content ex ante, and not upon receiving a particular notice of copyright

6. Digital Millennium Copyright Act, 17 U.S.C. § 1201 (2012).

7. In 2012, Google's 441,370 notices contained over 54 million individual takedown requests. In 2013, the company processed over 230 million takedown requests. In 2014, it processed 345 million requests. In 2015, Google received over half a billion removal requests. *See* Daniel Seng, *The State of the Discordant Union: An Empirical Analysis of DMCA Takedown Notices*, 18 VA. J.L. & TECH. 369, 444, 460-61 (2014); Ernesto, *Google Asked to Remove 558 Million "Pirate" Links in 2015*, TORRENTFREAK (Dec. 30, 2015) https://torrentfreak.com/google-asked-remove-558-million-pirate-links-2015 [https://perma.cc/S3C9-HC2Y]; Ernesto, *Google Discarded 21,000,000 Takedown Requests in 2013*, TORRENTFREAK (Dec. 27, 2013), https://torrentfreak.com/google-discarded-21000000-takedown-requests-in-2013-131227 [https://perma.cc/UZ7P-PKQB]; Joe Mullin, *Google Handled 345 Million Copyright Takedowns in 2014*, ARS TECHNICA (Jan. 6, 2015, 1:05 PM), http://arstechnica.com/tech-policy/2015/01/google-handled-345-million-copyright-takedowns-in-2014 [https://perma.cc/Y42Y-ZQVS].

8. JENNIFER M. URBAN ET AL., NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE 27-28 (2016), https://assets.documentcloud.org/documents/2779722/SSRN-id2755628.pdf [https://perma.cc/V7TP-J9GP].

9. *See infra* Part II.C.1.

10. *See* Thomas H. Davenport & Jeanne G. Harris, *Automated Decision Making Comes of Age*, 46 MIT SLOAN MGMT. REV. 83, 84 (2005).

11. Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1252-53 (2008).

12. *See infra* Part III.B (highlighting the distinction between regulated mechanisms of algorithmic copyright enforcement and voluntary mechanisms of algorithmic copyright enforcement). One example of an automated system implementing the DMCA's N&TD is DMCANotice.com. *See* DMCANOTICE.COM, http://www.dmcanotice.com [https://perma.cc/5UM4-GKP6].

infringement from the copyright owner.[13]

Despite these efficiency-related advantages, algorithmic copyright enforcement lacks sufficient measures to ensure that online intermediaries are held accountable for their actions, failures, and wrongdoings.[14] Our analysis shows that algorithmic implementations of N&TD, and even more so voluntary measures applied to detect and prevent copyright infringement, fare poorly in accountability measures. Algorithmic enforcement mechanisms are non-transparent in the way they exercise discretion over determining copyright infringement and fair use; they afford insufficient opportunities to challenge the decisions they make while failing to adequately secure due process; and they curtail the possibility of correcting errors in individual determinations of copyright infringement by impeding the opportunity for public oversight.

This Article proceeds in three parts. Part II discusses algorithmic law enforcement by online intermediaries and explains why copyright enforcement by online intermediaries makes an interesting case study for exploring algorithmic accountability. It further establishes a three-factor framework for assessing accountability in algorithmic enforcement that is based on transparency, due process, and public oversight. Part III then applies this three-factor framework to different mechanisms of algorithmic copyright enforcement to analyze their accountability deficiencies. Part IV maps the barriers for encouraging accountability in algorithmic copyright enforcement. In particular, Part IV considers the complicated and non-transparent nature of algorithms; the unpredictable nature of enforcement by constantly evolving learning machines; legal barriers that hinder the ability of the public to review and investigate algorithmic copyright enforcement; and the practical failure of existing mechanisms of public scrutiny, such as the counter notice procedure under the DMCA. Finally, Part IV explores existing and possible accountability-enhancing mechanisms, including watchdog initiatives, reverse engineering experiments, voluntary transparency reports from intermediaries, and regulatory mechanisms of mandatory disclosure.

II. ALGORITHMIC ACCOUNTABILITY

A. *The Rise of Algorithmic Enforcement by Online Intermediaries*

Recent years have seen a growing use of algorithms in performing law enforcement tasks.[15] Algorithmic law enforcement is becoming pervasive: for

13. *See infra* Part III.C.

14. Joshua A. Croll et. al., *Accountable Algorithms*, 165 U. PA. L. REV. (forthcoming 2017) (arguing that traditional accountability mechanisms and legal standards have not kept pace with technology).

15. Algorithmic law enforcement has been drawing the attention of law and technology scholars for over a decade. *See, e.g.*, Tarleton Gillespie, *The Relevance of Algorithms*, *in* MEDIA TECHNOLOGIES: ESSAYS ON COMMUNICATION, MATERIALITY, AND SOCIETY 167 (Tarleton Gillespie, et al. eds., 2014); Ian Kerr, *Digital Locks and the Automation of Virtue*, *in* FROM

example, cameras are issuing speeding tickets[16] and GPS-enabled bracelets or anklets fitted to offenders are tracking their locations, notifying their victims and the police whenever the offenders enter a prohibited area.[17] Algorithms not only enforce laws,[18] but also detect compliance with terms of use of online vendors and impose social norms on social media. Algorithmic enforcement generally involves large-scale collection of data by various sensors, data processing by algorithms, and automatic performance. It can be built to scale easily, offering an efficient means to manage, organize, and analyze today's massive amounts of data with uniformity and particularity, and then to structure decision-making accordingly.[19] Algorithms essentially permit more complex analysis of information through the "extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions."[20]

Algorithmic law enforcement is particularly ubiquitous online, where behavior is inherently mediated by computer code. Indeed, the software and hardware that cyberspace is built of create limitations on how people can behave,[21] for instance whether Internet users must enter a password to gain access, how active must they be whilst using a specific website to remain signed in, or the extent to which users can view data about each other as defined by their privacy preferences. The rise of algorithmic enforcement for online behavior was

---

"RADICAL EXTREMISM" TO "BALANCED COPYRIGHT": CANADIAN COPYRIGHT AND THE DIGITAL AGENDA 247 (Michael Geist ed., 2010); LAWRENCE LESSIG, CODE: VERSION 2.0 (2006); JONATHAN ZITTRAIN, THE FUTURE OF THE INTERNET AND HOW TO STOP IT (2008); Kenneth A. Bamberger, *Technologies of Compliance: Risk and Regulation in a Digital Age*, 88 TEX. L. REV. 669 (2010); Citron, *supra* note 11; Helen Nissenbaum, *From Preemption to Circumvention: If Technology Regulates, Why Do We Need Regulation (and Vice Versa)?*, 26 BERKELEY TECH. L.J. 1367 (2011); Anjanette H. Raymond & Scott J. Shackelford, *Technology, Ethics, and Access to Justice: Should an Algorithm Be Deciding Your Case?*, 35 MICH. J. INT'L L. 485 (2014); Michael L. Rich, *Should We Make Crime Impossible?*, 36 HARV. J.L. & PUB. POL'Y 795 (2013); Danny Rosenthal, *Assessing Digital Preemption (and the Future of Law Enforcement?)*, 14 NEW CRIM. L. REV. 576 (2011); Lisa Shay et al., *Confronting Automated Law Enforcement* (We Robot Conference Paper, 2013), http://www.rumint.org/gregconti/publications/201204_Shay_ALE.pdf [https://perma.cc/JVB6-CZ3A].

16. Florida Statutes section 316.0083, known as the Mark Wandall Traffic Safety Program, authorizes local governments to use red light cameras to enforce violations of sections 316.074(1) and 316.075(1)(c), both of which prohibit the running of red lights. *See* 2013-160 Fla. Laws 9; FLA. STAT. § 316.008(8)(a) (2011); Chris Matyszczyk, *Tickets Issued Due to Red-Light Cameras Are Illegal, Says Florida Court*, CNET (Oct. 21, 2014, 1:26 PM), http://www.cnet.com/news/tickets-issued-due-to-red-light-cameras-are-illegal-says-florida-court [https://perma.cc/L7CC-SH65].

17. See Christine Clarridge, *How GPS Bracelets Keep Track of Sex Offenders*, SEATTLE TIMES (Apr. 22, 2009, 12:00 AM), http://www.seattletimes.com/seattle-news/how-gps-bracelets-keep-track-of-sex-offenders [https://perma.cc/7EPT-DQ66].

18. See Citron, *supra* note 11, at 1263-67 for additional examples.

19. Bamberger, *supra* note 15, at 687-89.

20. THOMAS H. DAVENPORT & JEANNE G. HARRIS, COMPETING ON ANALYTICS: THE NEW SCIENCE OF WINNING 7 (2007) (discussing analytics specifically, rather than algorithms more generally).

21. LESSIG, *supra* note 15, at 124-25.

predicted in the late 1990s by information law scholars such as Joel Reidenberg, who described the Lex Informatica technological standards that offer technological solutions for information policy rules,[22] and Lawrence Lessig, who coined the term "code is law" to describe how algorithms can substitute for law in regulating certain behaviors.[23] Yet the comprehensiveness and robustness of algorithmic law enforcement on the Internet were hardly foreseen.

While technology has been used to aid law enforcement for many years,[24] from locks and speed bumps to full-body scans at airports,[25] algorithmic law enforcement implemented by online intermediaries (e.g., search engines or hosting websites) makes enforcement more robust.[26] Online intermediaries have acquired an important role in managing online behavior and protecting the rights of Internet users. They offer a natural point of control for monitoring, filtering, blocking, and disabling access to content, which makes them ideal partners for performing civil and criminal enforcement.[27] Intermediaries currently manage and police the usage of a tremendous stream of online content pursuant to

22. Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553, 556-58, 562 (1998) (such as the Platform for Internet Content Selection (PICS) that was designed to accommodate different standards for content without compromising free speech, or technological mechanisms that can anonymize information that would otherwise be associated with specific users).

23. LESSIG, *supra* note 15.

24. Rosenthal, *supra* note 15, at 577.

25. *See, e.g.*, *X-ray Full-Body Scanners for Airport Security*, GREEN FACTS (2016), http://copublications.greenfacts.org/en/x-ray-full-body-scanners-for-airport-security [https://perma.cc/WEU3-F5WK].

26. Other copyright enforcement algorithms employing Digital Rights Management (DRM) or Graduate Response will hence remain outside the scope of this Article. These enforcement algorithms largely function in a bi-directional way, limiting the target user's access to or use of a copyrighted work, or penalizing her by restricting her access to the web. Their implementation does not have a robust effect on public discourse and the public sphere. For a comprehensive historical analysis of DRM, see Bill D. Herman, *A Political History of DRM and Related Copyright Debates, 1987-2012*, 14 YALE J.L. & TECH. 162 (2012), and Michael S. Sawyer, *Filters, Fair Use & Feedback: User-Generated Content Principles and the DMCA*, 24 BERKELEY TECH. L.J. 363, 380-82 (2009).

27. Extensive scholarship has focused on the role of access providers, hosting facilities, search engines, social networks, and application providers as gatekeepers. *See, e.g.*, JACK GOLDSMITH & TIM WU, WHO CONTROLS THE INTERNET?: ILLUSIONS OF A BORDERLESS WORLD (2006); Patricia Sánchez Abril, *Private Ordering: A Contractual Approach to Online Interpersonal Privacy*, 45 WAKE FOREST L. REV. 689 (2010); Annemarie Bridy, *Graduated Response and the Turn to Private Ordering in Online Copyright Enforcement*, 89 OR. L. REV. 81 (2010); Stacey L. Dogan, *Trademark Remedies and Online Intermediaries*, 14 LEWIS & CLARK L. REV. 467 (2010); Mark MacCarthy, *What Payment Intermediaries Are Doing About Online Liability and Why It Matters*, 25 BERKELEY TECH. L.J. 1037 (2010); Ronald J. Mann & Seth R. Belzley, *The Promise of Internet Intermediary Liability*, 47 WM. & MARY L. REV. 239 (2005); Joel R. Reidenberg, *States and Internet Enforcement*, 1 U. OTTAWA L. & TECH. J. 213 (2004); Jonathan Zittrain, *A History of Online Gatekeeping*, 19 HARV. J.L. & TECH. 253 (2006); Paul Sholtz, *Transaction Costs and the Social Cost of Online Privacy*, FIRST MONDAY (May 7, 2001), http://journals.uic.edu/ojs/index.php/fm/article/view/859/768 [https://perma.cc/4FZU-9T2V].

different laws, including the laws of security,[28] privacy,[29] defamation,[30] and intellectual property.[31]

Algorithmic enforcement by online intermediaries effectively converges law enforcement and adjudication powers, reflecting a profound transformation in our traditional system of governance by law. While traditional law enforcement involves detection, prosecution, adjudication, and meting out punishment, algorithmic enforcement combines all functions, focusing primarily on detection and prevention.[32] As we demonstrate in this Article, the convergence of law enforcement and adjudication powers in the hands of a small number of mega platforms, and the robustness of algorithmic enforcement by private, online intermediaries raise critical challenges to the notions of trust and accountability that are inherent to reliable systems of law enforcement.[33]

B. *Accountability of Online Intermediaries Engaging in Algorithmic Law Enforcement*

Accountability refers to the extent to which decision-makers are expected to justify their choices to those affected by these choices, be held answerable for their actions, and be held responsible for their failures and wrongdoings.[34] Basically,

---

28. USA PATRIOT Act of 2001, Pub. L. No. 107-56, § 215, 115 Stat. 272, 287-88 (2001) (facilitating government access to customer data held by service providers).

29. *Id.* §§ 210-212, 115 Stat. at 283-85 (granting service providers immunity from damages if they, in good faith, produce data for an investigation undertaken to protect against international terrorism or clandestine intelligence activities).

30. The Communications Decency Act, 47 U.S.C. § 231 (2012) (criminalizing the online provision of indecent materials to minors, unless the initiator had undertaken a good faith effort to determine the age of the person on the other end of the network)

31. Digital Millennium Copyright Act, 17 U.S.C. § 1201 (2012).

32. For instance, proposals to promote wearable cameras among policemen in the US have followed the incident in Ferguson, Missouri, where a police officer shot an unarmed teenager named Michael Brown. Presumably, body cameras worn by police officers in the Ferguson aftermath would reduce police violence and diminish the need to detect and prosecute police abuse of power. *See Michael Brown Shooting: Ferguson Police to Get Body Cameras*, CBCNEWS (Aug. 31, 2014), http://www.cbc.ca/news/world/michael-brown-shooting-ferguson-police-to-get-body-cameras-1.2752146 [https://perma.cc/K6G5-9JKL]; *see also* Rich, *supra* note 15, at 803.

33. *See* NAT'L ACADEMY OF SCI., TRUST IN CYBERSPACE (Fred B. Schneider ed., 1999) (discussing trust in the context of IT systems).

34. Michael W. Dowdle, *Public Accountability: Conceptual, Historical, and Epistemic Mappings*, *in* PUBLIC ACCOUNTABILITY: DESIGN, DILEMMAS AND EXPERIENCES 1, 3 (Michael W. Dowdle ed., 2006) ("persons with public responsibilities should be answerable to 'the people' for the performance of their duties."); Adam M. Samaha, *Government Secrets, Constitutional Law, and Platforms for Judicial Intervention*, 53 UCLA L. REV. 909, 916 (2006) (explaining that democracies should allow citizens to "appreciably influence the direction of government, and . . . have an opportunity to assess progress and assign blame"); *see also* Danielle Keats Citron & Frank Pasquale, *Network Accountability for the Domestic Intelligence Apparatus*, 62 HASTINGS L.J. 1441 (2011) (focusing on accountability as a measure to cure the problems generated by the growing use of Fusion Centers); Tal Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1530, 1533 (2013).

accountability ensures that decision-makers exert power in a fair and effective manner. It can be produced by ex ante mechanisms that limit the power of decision-makers through structured guidelines and standards, by ex post mechanisms of transparency that permit review of the actions of decision-makers and the outcomes of their decisions, or by both.[35] Furthermore, accountability can be achieved through formal mandates, such as legal rules or regulations that restrain decision-makers' enforcement power,[36] and/or through informal means, such as market forces that check decision-makers' discretion and promote voluntary disclosure in relation to their choices and related outcomes.[37]

In the context of algorithmic enforcement by online intermediaries, generating accountability through these different mechanisms is rather challenging. Even where formal, structured guidelines exist (i.e., the standards set by the DMCA[38]), private players (online intermediaries) translate them into non-transparent algorithms.[39] Together with their rapid scalability, these algorithms are essentially a "black box"[40]—we do not know, and hence cannot foresee, how exactly they exercise their power to regulate our online behavior. Nor does producing accountability through formal or informal mechanisms of transparency currently seem promising. Indeed, as private, profit-maximizing entities, online intermediaries who apply enforcement algorithms to manage their users' behavior are not required to disclose to the public what content they remove, and for what specific reason.[41] They may legitimately craft their own terms of use to manage the content they distribute. Subjecting private policies of content management designed by law abiding intermediaries to legal intervention is controversial, and may raise objections similar to those asserted against proposals to interfere with the editorial discretion of publishers of the daily news or the media.[42]

But when online intermediaries engage in fundamental law enforcement,

---

35. Orna Rabinovitch-Einy, *Technology's Impact: The Quest for a New Paradigm for Accountability in Mediation*, 11 HARV. NEGOT. L. REV. 253, 260 (2006).

36. *See* Susan P. Sturm, *Second Generation Employment Discrimination: A Structural Approach*, 101 COLUM. L. REV. 458, 475 (1998) (describing the rule enforcement approach as a "fixed code of specific rules or commands that establishes clear boundaries governing conduct").

37. Rabinovitch-Einy, *supra* note 35, at 261.

38. *See infra* Part III.A.

39. *See infra* Part IV.A.

40. *See infra* Part IV.A.

41. However, under the Open Internet Transparency Rule, ISPs are required to disclose information about "network management practices, performance, and commercial terms of service." The Rule applies to service descriptions, including expected and actual broadband speed and latency. The Rule also applies to pricing, including monthly prices, usage-based fees, and any other additional fees that consumers may be charged. Additionally, it covers providers' network management practices, such as congestion management practices and the types of traffic subject to those practices. This does not seem to apply to a provider's copyright policy. *See Open Internet Transparency Rule*, FED. COMMC'NS COMM'N, https://www.fcc.gov/guides/open-internet-transparency-rule [https://perma.cc/AZB9-3E67].

42. *See* Assoc. Press v. United States, 326 U.S. 1, 20 n.18 (1945) (holding that antitrust law cannot "compel [the Associated Press] or its members to permit publication of anything which their 'reason' tells them should not be published"); *see also infra* Part II.C.

they effectively act like judges[43] who must adhere to the provisions set by the governing law they enforce. When performing public functions meant to serve the public at large by a formal or informal delegation of power from the government—and law enforcement is obviously one of these functions—online intermediaries effectively act like public administrative agencies.[44] Just as judicial enforcement procedures are required to ensure due process and facilitate public scrutiny to strengthen public trust and promote the rule of law, so too should algorithms be employed by online intermediaries to enforce the rights of Internet users.[45] Otherwise, because online intermediaries resemble a "company town"[46] more than they do a small, hardly influential newspaper, they could exercise disproportionate power, which power may ultimately shape public discourse and even violate users' fundamental rights. Indeed, if YouTube removes a video, it is unlikely to be seen; if Google blocks a link to an online business website, it may literally die. Consequently, interested audiences might be deprived of access to relevant information provided by individuals or businesses.[47]

We argue that accountability is what distinguishes content *management* determinations that online intermediaries make in their private capacities from content *adjudication* determinations they make in their administrative capacities as law enforcers. Adjudication of online content without accountability may lead to manipulation and abuse of power, create new barriers to open competition and market innovation, and challenge civil rights.[48] Through the prism of

---

43. In relation to the implementation of the Right To Be Forgotten, Google European Communications Director Peter Barron stated that: "[Google] never expected or wanted to make [these] complicated decisions that would in the past have been extensively examined in the courts, [but are] now being made by scores of lawyers and paralegal assistants [at Google]." *See* Aoife White, *Google EU Ruling Response Vetted as Complaints Pile Up*, BLOOMBERG (Sept. 18, 2014, 6:04 AM), http://www.bloomberg.com/news/articles/2014-09-18/google-eu-ruling-response-vetted-as-complaints-pile-up [https://perma.cc/UQ6F-UHUE].

44. Edward Lee, *Recognizing Rights in Real Time: The Role of Google in the EU Right to Be Forgotten*, 49 U.C. DAVIS L. REV. 1017, 1055-73 (2016).

45. *Id.* Lee explains that in relation to implementing the EU's recent decision about the Right To Be Forgotten in Case C-131/12, *Google Spain SL v. Agencia Española de Protección de Datos (AEPD) (Costeja)*, 2014 EUR-Lex 62012CJ0131 (May 13, 2014), which acknowledged the right of individuals in the EU "to request search engines to remove, from the search results for an individual's name, links to web content that contains personal information about the individual that is 'inadequate, irrelevant or excessive in relation to the purposes of the processing,' 'not kept up to date,' or 'kept for longer than is necessary,'" Google functions similarly to a government agency or administrative body. *Id.* at 1022 (citing *Costeja*, ¶¶ 92, 94).

46. MARGARET JANE RADIN, BOILERPLATE: THE FINE PRINT, VANISHING RIGHTS, AND THE RULE OF LAW 33 (2013) (using the term "democratic degradation" to describe a situation where firms displace state regulation); *see also* M. Todd Henderson, *The Nanny Corporation*, 76 U. CHI. L. REV. 1517, 1535-37 (2009); Tal Zarsky, *Social Justice, Social Norms and the Governance of Social Media*, 35 PACE L. REV. 154, 166 (2014).

47. As explained by Grimmelman, the free speech interest of a business, which is affected by its placement in search results, is derivative of users' free speech, and users' free speech is harmed when users are deprived of access to the speech offered by the website. *See* James Grimmelmann, *Some Skepticism About Search Neutrality*, *in* THE NEXT DIGITAL DECADE: ESSAYS ON THE FUTURE OF THE INTERNET 435, 441-42 (Berin Szoka & Adam Marcus eds., 2010).

48. *See infra* Part II.C; *see also* Lee, *supra* note 44, at 42-47 (counting several

accountability, the general public can criticize the manner in which online intermediaries regulate their content and shape their reactions accordingly. Particularly, the public may choose not to use the services of platforms that appear to abuse their power against fundamental rights such as equality, free speech, and freedom to conduct business, namely, the ability to reach potential customers in free and competitive markets without undue interference.

C. *Accountability Matters in Algorithmic Copyright Enforcement*

Algorithmic copyright enforcement offers an excellent case study for studying the challenges involved in governance by algorithms. Copyright law has been at the forefront of digital law enforcement since the early 1990s. The ease of digital copying and mass distribution gave rise to digital locks, digital rights management (DRM) systems, and technological protection measures (TPM), which enable rights-holders to technically prevent unauthorized access to and use of their copyrighted works.[49] Yet, confronted with the threats of dispersed mass piracy, rights-holders further increased their pressure on online service providers (OSPs) to actively participate in fighting online infringement.[50] Rights-holders pushed towards active involvement of OSPs in copyright enforcement in exchange for limited immunity from liability for copyright infringement committed by their users.[51] These developments eventually shaped the intermediary safe harbor regime under the DMCA.[52]

Arguably, copyright enforcement by online intermediaries, pursuant to the DMCA's N&TD framework, offers an efficient alternative to the cumbersome, often impracticable traditional enforcement of copyright through the legal system, which barely keeps up with the accelerated pace of technological change. Indeed, the legal system is often understaffed, slow to act, and costly for litigants and for society,[53] compared to the cheap, instant, scalable, and robust system of

---

accountability drawbacks arising from giving Google the primary responsibility of deciding the contours of the recently recognized Right To Be Forgotten: private anonymous employees that do not reflect users' diversity; possible bias of employees in favor of access to information; minimal due process afforded to affected users; and the inevitable result of mistaken legal determinations); Zarsky, *supra* note 46, at 156 (arguing that the notion that a small group of managers unilaterally sets the rules regulating the social discourse is daunting and may impact users' core rights, including their ability to engage in free speech or invoke their right to privacy).

49. *See, e.g.*, Pamela Samuelson, *Intellectual Property and the Digital Economy: Why the Anti Circumvention Regulations Need to Be Revised*, 14 BERKELEY TECH. L.J. 519, 534-35 (1999).

50. *See infra* Part III.A.

51. Niva Elkin-Koren, *After Twenty Years: Revisiting the Copyright Liability of Online Intermediaries*, *in* THE EVOLUTION AND EQUILIBRIUM OF COPYRIGHT IN THE DIGITAL AGE 29 (Susy Frankel & Daniel J Gervais eds., 2014) ("Digital networks have led to an 'enforcement failure' in copyright-related industries, turning online intermediaries into key players in enforcement efforts.").

52. 17 U.S.C. § 512 (2012).

53. For similar arguments in relation to risk management, *see* Bamberger, *supra* note 15, at 685.

online copyright enforcement. This explains why much of today's online copyright enforcement is embedded in the system design of online intermediaries, using algorithms not only to remove allegedly infringing content upon notice of copyright infringement, but also to monitor, filter, block, and disable access to content automatically flagged as infringing.[54]

But should online intermediaries be held accountable for algorithmic copyright enforcement? Indeed, as we mentioned earlier, when intermediaries choose to filter allegedly infringing materials or to remove some materials upon notice, they may simply be making private choices regarding content that is made available on their platforms.[55] At the same time however, when online intermediaries monitor, filter, block, and remove allegedly infringing materials they engage, in de facto copyright enforcement. The ubiquity of algorithmic copyright enforcement by online intermediaries makes the case for accountability even stronger. In many respects, the N&TD regime under the DMCA, and even more so voluntary mechanisms of algorithmic copyright enforcement,[56] effectively privatize governmental functions while blurring the public/private divide.[57] Private intermediaries act as both a judge and an executioner, performing functions of great importance to the public which are normally reserved to authorized governmental bodies.[58]

Indeed, the original purpose of codifying a safe harbor under the N&TD procedure was to encourage the cooperation of private, online intermediaries in combatting online copyright infringement.[59] Today, most online copyright enforcement practices take place on privately owned platforms, and not in traditional legal forums, such as courts and law offices.[60] Thus, there is a growing concern that the legitimate business interests of online intermediaries would compromise the duties involved in performing law enforcement tasks, such as unprejudiced treatment, transparency, and due process.[61] Holding online intermediaries accountable for algorithmic copyright enforcement may therefore introduce the necessary checks and oversight for the use of semi-governmental power.

---

54. *See infra* Part III.A.

55. James Grimmelmann, *Speech Engines*, 98 MINN. L. REV. 868, 870-71 (2014).

56. *See infra* Part III.B.

57. Michael D. Birnhack & Niva Elkin-Koren, *The Invisible Handshake: The Reemergence of the State in the Digital Environment*, 8 VA J. L. & TECH. 6, para. 66 (2003) (explaining that the use of private sector companies for government censorship may allow government officials, with a light hand on the trigger, to prevent content from being uploaded to the web, thus violating freedom of expression without any legal scrutiny); *see also* Jody Freeman, *Private Role in Public Governance*, 75 N.Y.U. L. REV. 543, 547 (2000).

58. *See* Lee, *supra* note 44 (describing the significant role that Google is playing in the development of the European Union's Right To Be Forgotten, and positing that Google is functioning like a private administrative agency).

59. *See infra* Part III.A.

60. Sharon Bar-Ziv & Niva Elkin-Koren, *Uncovering the Invisible: Studying Algorithmic Online Copyright Enforcement* (forthcoming) (on file with authors).

61. Freeman, *supra* note 57, at 574-75.

Specifically, three explanations support our proposition that accountability matters in algorithmic copyright enforcement by online intermediaries: the first concerns the rule of law and the need to ensure fairness and predictability in exercising enforcement functions; the second relates to freedom of speech and the free flow of information, and the third is about preserving sound copyright policy.

1. *Accountability and the Rule of Law*

Algorithmic enforcement, like any other law enforcement activity, should comply with the rule of law. This principle requires that any exercise of power is duly delegated, and that rules are clear so that people can develop reliable expectations and make autonomous choices accordingly.[62] The case of online intermediaries performing copyright enforcement tasks is no exception: unless operating pursuant to the authorization of the law (such as the DMCA), intermediaries may lack copyright enforcement authority. Yet, as explained in this Subpart, determining whether mechanisms of algorithmic copyright enforcement employed by online intermediaries abide by the rule of law is far from straightforward.

Unlike automated enforcement mechanisms that essentially detect strictly defined unlawful activities, such as red light crossing,[63] algorithmic copyright enforcement often involves implementation of flexible legal standards. Many of the most serious issues in copyright law involve discretion, including determining the degree of "originality" required to establish copyrightability;[64] deciding what amounts to "substantial similarity" to establish infringement;[65] or considering what constitutes "permissible use" under fair use.[66] Resolving these flexible issues largely requires a qualitative process of assessment and balancing that ought to be decided on a case-by-case basis.[67]

Translating doctrinal law and policy into code may result in significant, albeit unintentional, alterations of meaning,[68] partly because the artificial languages

62. Citron, *supra* note 11, at 1297.

63. When such detectors issue a ticket, clearly they have detected a car crossing in red light. This sort of decision-making does not involve any qualitative determinations and therefore, even when implemented mechanically, the output (a ticket is issued) reveals sufficient information about the decision-making process (the algorithm identified a car driving in red light; red light driving is forbidden; therefore, a ticket was issued).

64. *See generally* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991) (explaining that compilations of pre-exiting facts demand some degree of originality in their selection and arrangement to satisfy the minimum constitutional standards for copyright protection).

65. *See, e.g.*, Ideal Toy Corp. v. Fab-Lu Ltd., 266 F. Supp. 755, 756 (S.D.N.Y. 1965), *aff'd*, 360 F.2d 1021 (2d Cir. 1966).

66. Cambridge Univ. Press v. Becker, 769 F.3d 1232 (11th Cir. 2014).

67. Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 561, (1985). (explaining that fair use determinations demand a thorough case-by-case analysis).

68. *See* AUSTL. ADMIN. REVIEW COUNCIL, AUTOMATED ASSISTANCE IN ADMINISTRATIVE DECISION MAKING: ISSUES PAPER 35 (2003),

intelligible to computers have a more limited vocabulary than human languages.[69] One risk is that the code being used may not fully capture the nuances of a particular policy,[70] which, like fair use, may require case-by-case evaluation. The code may even alter the original law, either by accident or convenience, with no prior delegation of public power.[71] Obviously, "encoded rules that change established policy cannot be understood by affected individuals or reviewed by more democratically accountable superiors. In that regard, rulemaking by code writers is ultra vires even as it is inevitable."[72]

This is not to say, however, that algorithms cannot process reliable decisions about copyright infringement and fair use.[73] While some scholars insist that this is the case,[74] others contemplate that machine-learning algorithms may actually

---

http://www.arc.ag.gov.au/Documents/Automated+Assistance.pdf [https://perma.cc/C43P-4FPR].

69. James Grimmelmann, *Regulation by Software*, 114 YALE L.J. 1719, 1728 (2005).

70. *See* Graham Greenleaf et al., *Representing and Using Legal Knowledge in Integrated Decision Support Systems: DataLex Workstations*, 3 ARTIFICIAL INTELLIGENCE & L. 97, 127 (1995).

71. Citron, *supra* note 11, at 1297.

72. *Id.*

73. Relying exclusively on algorithms to execute adequate fair use determinations seems, at times, unrealistic. For instance, Professors Burk and Cohen, in discussing DRM systems, stress that "building the range of possible uses and outcomes into computer code would require both a bewildering degree of complexity and an impossible level of prescience. There is currently no good algorithm that is capable of producing such an analysis." Dan L. Burk & Julie E. Cohen, *Fair Use Infrastructure for Rights Management Systems*, 15 HARV. J.L. & TECH. 41, 56 (2001); *see also* Mark Lemley, *Rationalizing Internet Safe Harbors*, 6 J. ON TELECOMM. & HIGH TECH. L. 101, 110-11 (2007) ("Image-parsing software may someday be able to identify pictures or videos that are similar to individual copyrighted works, but they will never be able to determine whether those pictures are fair uses, or whether they are legitimate copies or displays made under one of the many statutory exceptions, or whether the individual pictured is 16 rather than 18 years of age.").

Similarly, some scholars argue that filtering algorithms cannot make reliable qualitative determinations, such as deciding "the purpose and character of the use" or the "nature of the copyrighted work." Furthermore, algorithms are unable to consider information external to the content itself when analyzing fair use. Yet, in order to assess the fourth factor, "the effect of the use upon the potential market for or value of the copyrighted work," the technology must consider external information about the market. Additionally, because filtering algorithms make mechanical decisions, some contend that they cannot look at the allegedly infringing material as a whole. Consequently, they may unlawfully ban transformative uses. 17 U.S.C. § 107 (2015); Sawyer, *supra* note 26 **Error! Bookmark not defined.**, at 389; Edward W. Felten, *A Skeptical View of DRM and Fair Use*, COMM. A.C.M., Apr. 2003, at 58.

These allegations seem to be outdated in light of today's algorithms' learning capacities. In fact, the more pertinent question should be authority-related: does copyright law permit delegating the substantial discretion originally accorded to courts to a non-transparent mechanical process? Determining whether algorithms are both authorized and capable of making reliable fair use determinations is a normative question that we prefer to leave outside the scope of this Article. We take the positivistic state of affairs as given: intermediaries effectively employ different copyright enforcement algorithms and these algorithms are empowered to execute infringement and fair use determinations.

74. Ira S. Nathenson, *Civil Procedures for a World of Shared and User-Generated Content*, 48 U. LOUISVILLE L.R. 912, 938-44 (2010) (describing how content ID procedures may compromise fair use); Sawyer, *supra* note 26, at 388-90 (2009) (arguing that fair use considerations cannot be

enable other algorithms to make smarter decisions based on learned patterns of data.[75] The problem is that accomplishing algorithmic accountability becomes very challenging when algorithms execute discretion-based decisions whose processing is a "black box." In other words, even if self-learning algorithms can be created to engage in case-by-case applications of legal standards, the lack of transparency remains a serious problem. Most consumers simply do not hold the necessary expertise to understand the complex technical terminology embedded in algorithmic enforcement mechanisms.[76] And as long as online users cannot comprehend how enforcing algorithms effectively detect online copyright infringement, they cannot determine whether algorithmic copyright enforcement effectively complies with what it is authorized to do under the law.

2. *Accountability and the Public Sphere*

Algorithmic enforcement by prominent online intermediaries who play a central role in shaping public discourse[77] carries direct implications for the public sphere, which further reinforce the need to secure algorithmic accountability. By offering an alternative medium for information dissemination to which people increasingly resort—especially when more traditional outlets of distribution are unavailable—online intermediaries are becoming global arbiters of free speech. Indeed, it was YouTube Movies, Google Play and Microsoft Zbox Video that offered online streaming of Sony Pictures' movie, *The Interview*, after its release in theaters was suspended following threats of terrorist attacks.[78] Online intermediaries also made the caricatures of the magazine *Charlie Hebdo* available

encoded in filtering technologies, and therefore when the burden to monitor copyright infringements shifts from copyright owners to OSP, fair use is likely to be disregarded).

75. NICHOLAS DIAKOPOULOS, ALGORITHMIC ACCOUNTABILITY REPORTING: ON THE INVESTIGATION OF BLACK BOXES 3 (2013), http://www.nickdiakopoulos.com/wp-content/uploads/2011/07/Algorithmic-Accountability-Reporting_final.pdf [https://perma.cc/V9B2-WS9A].

76. Julie E. Cohen, *DRM and Privacy*, 18 BERKELEY TECH. L.J. 575, 615 (2003).

77. YOCHAI BENKLER, THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM 130 (Yale Univ. Press 2006); *see also* Niva Elkin-Koren, *User-Generated Platforms*, *in* WORKING WITHIN THE BOUNDARIES OF INTELLECTUAL PROPERTY 111, 114-15 (2010).

78. *The Interview*, a 2014 comedy by Sony Pictures mocking North Korean leader Kim Jong-Un and depicting a plot for his assassination, was scheduled to be released on Christmas Day 2014. Following a massive cyberattack, and after receiving threatening messages that any theater screening the film would be physically attacked, Sony had decided to cancel the release of the movie in theaters. This controversial decision was widely criticized as self-censorship by free speech advocates. The decision was publicly denounced by President Obama, artists, and activists that accused Sony of "caving in" to terrorism and sacrificing free speech. Online intermediaries came to the rescue of free speech, offering *The Interview* via online streaming on YouTube Movies, Google Play and Microsoft Xbox Video. *See The Interview: A Guide to the Cyber Attack on Hollywood*, BBC NEWS (Dec. 29, 2014), http://www.bbc.com/news/entertainment-arts-30512032 [https://perma.cc/RXX4-PQWU]; *'The Interview' Made Available Online After Cyber Attack*, IRISH TIMES (Dec. 24, 2014, 6:05 PM), http://www.irishtimes.com/news/world/us/the-interview-made-available-online-after-cyberattack-1.2048469 [https://perma.cc/Z96D-6PGT].

after some news outlets decided not to re-publish them.[79]

In other cases, however, online intermediaries have played a different role, enabling speech control and generating de facto censorship. For instance, actress Cindy Lee Garcia was successful in raising a doubtful copyright claim against Google, causing the removal from YouTube of a provocative, anti-Islamic film, *The Innocence of Muslims*, based on her insignificant, five-second performance in the film.[80] It took Google fifteen months to convince the Court of Appeals that Garcia's "weak copyright claim cannot justify censorship in the guise of authorship,"[81] and to rescind the order requiring it to take down the controversial video. In a different case, YouTube facilitated the removal of a documentary film, *India's Daughter*, based on the gang rape of a twenty-three-year-old student, the screening of which was banned in India due to copyright infringement allegations.[82] YouTube also allowed the censorship of the satirical show *Fitnah* when it complied with DMCA takedown notices sent by the primary, state-funded Saudi TV channel, "Rotana."[83]

From a political perspective, Gannett Co., Inc., a massive media corporation that owns the *Courier-Journal* in Kentucky, successfully caused the removal from YouTube of a forty-second interview with the Democratic candidate for the

79. A few weeks later, in January 2015, a horrific massacre occurred at the Paris offices of Charlie Hebdo, a magazine which has published satires of the prophet Mohammed. Many journals have decided not to re-publish the latest Charlie Hebdo caricature for fear they will also be targeted, and in some countries local stores were reluctant to sell the magazine, fearing violence (Israel) or in compliance with a government ban (Turkey). CNN, along with other news outlets, has chosen to censor the controversial cartoons that ran in the magazine. *See* Barak Ravid et al., *Lieberman Tells Party Activists: Distribute Charlie Hebdo, Israel Must Not Turn Into ISIS*, HAARETZ (Jan 25, 2015, 9:03 AM), http://www.haaretz.com/news/national/1.638836 [https://perma.cc/RN8D-WZEG]; *Charlie Hebdo Attack: Three Days of Terror*, BBC NEWS (Jan. 14, 2015), http://www.bbc.com/news/world-europe-30708237 [https://perma.cc/2C32-U6M5]; Constanze Letsch, *Charlie Hebdo: Turkish Court Orders Ban on Web Pages Featuring Front Cover*, GUARDIAN (Jan. 14, 2015, 11:13 AM), http://www.theguardian.com/world/2015/jan/14/charlie-hebdo-turkey-block-web-pages-front-cover-muhammad [https://perma.cc/UT99-MTPV]; Alex Stedman, *CNN Explains Decision to Censor Charlie Hebdo Muslim Cartoons*, VARIETY (Jan. 7, 2015, 1:19 PM), http://variety.com/2015/tv/news/cnn-addresses-censoring-of-charlie-hebdo-cartoons-1201395044 [https://perma.cc/K6P5-GM6V].

80. Garcia v. Google, Inc., 766 F.3d 929, 940 (9th Cir. 2014), *rev'd en banc*, 786 F.3d 733 (9th Cir. 2015).

81. Garcia v. Google Inc., 786 F.3d 733, 743 (9th Cir. 2015) (en banc) (reasoning that "treating every acting performance as an independent work would not only be a logistical and financial nightmare, it would turn a cast of thousands into a new mantra: copyright of thousands").

82. YouTube removed most copies of the film soon after they became available due to copyright infringement allegations made by the British Broadcasting Corporation (BBC), which made the original broadcast from which the uploaded copies were taped. *See YouTube Removes India's Daughter Videos After BBC Copyright Request*, TRADEMARKS & BRANDS ONLINE (Mar. 9, 2015), http://www.trademarksandbrandsonline.com/news/youtube-removes-india-s-daughter-videos-after-bbc-copyright-request-4289 [https://perma.cc/X3NN-TAQA].

83. *Copyright Law as a Tool for State Censorship of the Internet*, BEFORE IT'S NEWS, (Dec. 3, 2014, 11:22 AM), http://beforeitsnews.com/libertarian/2014/12/copyright-law-as-a-tool-for-state-censorship-of-the-internet-2589350.html [https://perma.cc/WY27-WBBV].

Senate, Alison Lundergan Grimes, in which she desperately tried to avoid admitting she voted for President Obama, who was unfavorable in Kentucky.[84] Notwithstanding the video being a non-infringing fair use that transformed the original factual work into a political one,[85] YouTube enabled its censorship, at least temporarily, less than a month before elections, and exactly when its impact could have been most powerful.[86]

These examples suggest that algorithmic copyright enforcement may be used to remove content for reasons that presumably have very little to do with copyright infringement, turning the DMCA into a tool for global censorship.[87] Unfortunately, although copyright law includes built-in mechanisms that safeguard freedom of speech, such as the idea/expression dichotomy and the fair use doctrine,[88] online intermediaries have several incentives to go beyond these

---

84. Corynne McSherry, *For Shame: Gannett Abuses DMCA to Take Down Political Speech*, ELEC. FRONTIER FOUND. (Oct. 10, 2014), https://www.eff.org/deeplinks/2014/10/shame-gannett-abuses-dmca-take-down-political-speech [https://perma.cc/A2NS-6QK8].

85. Under 17 U.S.C. § 107 (2015),

> in determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—
> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
> (2) the nature of the copyrighted work;
> (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
> (4) the effect of the use upon the potential market for or value of the copyrighted work.

Accordingly, the demonstrative case of Courier-Journal involves a restriction on fair use: (1) the purpose of using a portion of the interview was completely political, not commercial; (2) the nature of the interview was highly factual, as it included Grimes's actual answer to a question she was asked; (3) the poster took the amount needed to accomplish its political purpose—only 40 seconds of the interview (a shorter clip might not have managed to achieve the poster's purpose); (4) the clip did not provide any financial gain to the poster, which could have otherwise deprived the copyright owner from income. If at all, the clip may only have had a positive effect upon the interview's commercial value, in highlighting its "juicy" parts.

86. For other examples, see Ryan Singel, *YouTube Flags Democrats' Convention Video on Copyright Grounds*, WIRED (Sept. 5, 2012, 12:10 AM), http://www.wired.com/2012/09/youtube-flags-democrats-convention-video-on-copyright-grounds [https://perma.cc/2HK9-TH58] (highlighting the removal of First Lady Michelle Obama's speech from YouTube due to erroneously identifying it as an infringing video).

87. Rossi v. Motion Picture Ass'n of Am., 391 F.3d 1000, 1002 (9th Cir. 2004) (describing the use of the DMCA's notice and takedown provisions to induce an ISP to take down a website from which illegal content could not be downloaded); Online Policy Group v. Diebold, Inc., 337 F. Supp. 2d 1195, 1204-05 (N.D. Cal. 2004) (detailing the use of DMCA notices to induce ISPs to take down websites containing internal memoranda that embarrassed a voting machine manufacturer, even though the websites were, in fact, protected fair use). The tactic of using the DMCA for silencing speech was also discussed in a Forbes article, which explained that the common corporate wisdom for dealing with Internet critics is to: "ATTACK THE HOST. Find some copyrighted text that a blogger has lifted from your Web site and threaten to sue his Internet service provider under the Digital Millennium Copyright Act. That may prompt the ISP to shut him down." Daniel Lyons, *Fighting Back*, FORBES (Nov. 14, 2005, 12:00 AM), http://www.forbes.com/forbes/2005/1114/128sidebar.html [https://perma.cc/MQ9H-45NB].

88. Eldred v. Ashcroft, 537 U.S. 186, 221 (2003) ("To the extent such assertions raise First Amendment concerns, copyright's built-in free speech safeguards are generally adequate to address them.").

embedded safeguards and over-enforce copyrights. First, the DMCA has provided Internet Service Providers, search engines and other intermediaries, with strong incentives to take down or block access to allegedly infringing content (otherwise, they may face liability for their users' infringements).[89] Second, online intermediaries may enter into licensing arrangements with prominent entities in copyright-heavy industries,[90] making them vulnerable to business-related pressures to strictly enforce their partners' copyrights.[91]

Over-enforcement of copyrights by disregarding fair use, or by removing works that are in the public domain, does not only limit the right of individuals who seek to share content online to express themselves and enjoy the fruits of such expressions liberally,[92] but also deprives the public as a whole of the benefit of consuming erroneously restricted speech in the marketplace of ideas. Information posted by an artist,[93] an activist,[94] or a politician[95] could raise public awareness of an issue and help communities mobilize around it. But once access to materials posted online is blocked or removed, a story may not unfold. Furthermore, the right of speakers to participate in the conversation is compromised. This triggers concerns regarding the appropriate restraints on freedom of speech and the compelling need to devise appropriate mechanisms for algorithmic accountability.

Finally, beyond these critical concerns regarding freedom of expression, unaccountable enforcement of online content raises additional concerns regarding other civil rights. From an economic perspective, control over what

---

89. 17 U.S.C. § 512 (2012) (exempting OSPs from liability for mistaken yet good faith removals of material); *see also* Seth F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, 155 U. PA. L. REV. 11, 23 (2006) (discussing the dangers of using proxy censors on free speech); Neil Weinstock Netanel, *First Amendment Constraints on Copyright After Golan v. Holder*, 60 UCLA L. REV. 1082, 1120-27 (2013).

90. *See supra* Part IV.C (discussing YouTube's Content ID).

91. Kreimer, *supra* note 89, at 29-30 ("Putting the censorship decision in the hands of the intermediary allows commercially powerful blocs of customers a potential veto on the speech of others.").

92. *See* Elisa Kreisinger, *The Impending Death of the YouTube Mashup*, DAILY DOT (June 27, 2014, 9:22 AM CT), http://www.dailydot.com/opinion/youtube-mashup-remix-copyright-universal [https://perma.cc/5XSA-75SN] (describing the limited ability of artists to upload mashups to YouTube).

93. *See, e.g.*, Parker Higgins, *Houston, We Have a Public Domain Problem*, MEDIUM (June 24, 2014), https://medium.com/@xor/houston-we-have-a-public-domain-problem-bd971c57dfdc [https://perma.cc/6P2T-C6KC] (reporting about an individual who had received a bogus copyright takedown notice for using public domain audio on SoundCloud).

94. For instance, a video in which Tom Cruise proclaims, in part, that Scientologists are the only experts on the mind, was removed by YouTube at the request of the Church of Scientology as part of a long-standing effort to keep copyrighted material from appearing on the Internet. *See* Robert Vamosi, *Anonymous Hackers Take on the Church of Scientology*, CNET (Jan. 24, 2008, 3:20 PM), http://www.cnet.com/news/anonymous-hackers-take-on-the-church-of-scientology [https://perma.cc/3EJN-SMSC].

95. *See* Jacqueline Klimas, *Online Campaign Ads May Prove Decisive in Midterm Elections*, WASH. TIMES (Sept. 28, 2014), http://www.washingtontimes.com/news/2014/sep/28/online-campaign-ads-may-prove-decisive-in-midterm-/?page=all [https://perma.cc/273B-5V68].

information becomes available may shape the consumer preferences, creating demand for some content while diminishing demand for other types of content. Google, for instance, penalizes sites repeatedly accused of copyright infringement, making them appear lower in Google's search results.[96] Users may make an unjustifiable use of Google's online mechanism of copyright enforcement to file bogus copyright complaints against their competitors. This may strongly affect the business opportunities of targeted sites and their respective owners' occupation rights, without any warrant or any finding by a court of copyright infringement. As private, profit-maximizing entities, intermediaries may potentially abuse their enforcement power due to commercial bias: they may favor their business partners and other powerful repeat players over weak Internet users.[97] Manipulations of this sort impose serious threats on open competition and market innovation,[98] further fortifying the importance of holding online intermediaries accountable for algorithmic copyright enforcement.

3. *Accountability and Copyright Policy*

Another reason to hold algorithmic copyright enforcement by online intermediaries accountable relates to its robustness, which may effectively alter settled copyright policy. Since algorithmic copyright enforcement implemented by online intermediaries affects a considerable number of people and a large volume of material, a particular implementation of rules by an algorithm (i.e., filtering, removal, blocking) can effectively shape copyright policy. For instance, algorithmic copyright enforcement may circumvent the objectives of copyright law by changing the default[99]: although copyright policy assumes that copyrighted materials are publicly available unless proven infringing, materials detected by copyright enforcement algorithms remain unavailable unless explicitly authorized by the copyright owner.[100] Similarly, the purpose of copyright law is to promote the creation of new works for the public benefit by providing authors with an economic incentive to create. Yet if algorithmic copyright enforcement fails to detect copyright infringement, it may lead to under-enforcement[101] (false negative), consequently depriving rights-holders of sufficient incentives to create. Similarly, erroneously removing or blocking access to non-infringing materials

96. *An Update to Our Search Algorithms*, GOOGLE INSIDE SEARCH (Aug. 10, 2012), https://search.googleblog.com/2012/08/an-update-to-our-search-algorithms.html [https://perma.cc/9VHH-874W].

97. *See infra* note 195 and accompanying text.

98. Jody Freeman, *Private Parties, Public Functions and the New Administrative Law*, 52 ADMIN. L. REV. 813, 845-849 (2000) (acknowledging the potential dangers to democratic accountability that private actors pose in mixed administration).

99. Jennifer Urban & Laura Quilter, *Efficient Process or Chilling Effects: Takedown Notices Under Section 512 of the Digital Millennium Copyright Act*, 22 SANTA CLARA COMPUTER & HIGH TECH. L.J. 621, 636 (2006).

100. *See infra* Part III.A.

101. *See* Shay, *supra* note 15, at 30.

through over-enforcement[102] (false positive) may further compromise copyright goals in promoting access to knowledge and encouraging creativity.

Of course, there is no error-free enforcement system,[103] and some level of error is inevitable in any procedure, including legal procedures.[104] Even automated systems, which are often perceived to be mistake-resistant, may be biased.[105] Yet the ubiquity of algorithmic copyright enforcement, derived from its "codish," automatic implementation, creates pervasive opportunities for error, which cannot possibly be corrected promptly.[106] Additionally, because algorithmic copyright enforcement evolves on private grounds, without the participation of citizens, public officials, or judges, it is "less likely to encounter repeal or amendment, compared to laws enforced through traditional means."[107] Indeed, for an opponent of current copyright laws to contest YouTube's preemptive enforcement policy, it is necessary to additionally oppose its effectiveness, "because, as a result of this effectiveness, the law will be less likely to change."[108] Hence, if we wish to preserve the traditional objectives of copyright law and ensure they are not distorted, we must be able to determine how copyright enforcement is being implemented on the ground.

D. *The Virtues of Accountability—A Three-Factor Framework*

The ubiquitous system of embedded copyright governance raises many concerns. It could effectively change the default rule of copyright law, diminish freedom of speech, and shape power relations. Online intermediaries are using algorithms to guard against copyright infringement, but are there adequate mechanisms ready to guard the guardians? To begin answering this question, we seek to explore how algorithmic mechanisms of copyright enforcement rank in accountability measures through the prism of public scrutiny. At this preliminary stage of algorithmic accountability research, we wish to examine the extent to

---

102. *See* Rich, *supra* note 15, at 812 ("[A]ny technology that seeks to prevent criminal conduct will inevitably also prevent some non-criminal conduct. This over-breadth may occur either by design or by mistake."); Rosenthal, *supra* note 15, at 594 ("Preemption will likely make many mistakes in enforcing laws that require subjective, case-specific inquiries to determine liability," such as "laws that are designed to be enforced only at the discretion of a private party, like many copyright restrictions.").

103. Shay, *supra* note 15, at 30 n.73.

104. Indeed, in *In re Verizon Internet Servs., Inc.*, the trial court downplayed the significance of such errors: "[S]uch mistakes are possible using evolving technology, but there is nothing to suggest they will cause substantial chilling of expression on the Internet." 257 F. Supp. 2d 244, 264 n. 23 (D.D.C. 2003), *rev'd sub nom.* Recording Indus. Ass'n of Am., Inc. v. Verizon Internet Servs., Inc., 351 F.3d 1229 (D.C. Cir. 2003).

105. Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO. SYS. 330, 332 (1996).

106. *Id.* at 331 ("Computer systems, for instance, are comparatively inexpensive to disseminate, and thus, once developed, a biased system has the potential for widespread impact. If the system becomes a standard in the field, the bias becomes pervasive.").

107. Rosenthal, *supra* note 15, at 597-98.

108. *Id.* at 598.

which the *public* can monitor algorithmic copyright enforcement by online intermediaries. Indeed, public review has proved its effectiveness in generating copyright reform. The successful backlash to the Stop Online Piracy Act (SOPA)[109] and the Preventing Real Online Threats to Economic Creativity and Theft of Intellectual Property Act (PIPA)[110] in January 2012, for instance, demonstrates how public outcry can influence copyright lawmaking.[111] The speedy success of the SOPA/PIPA contest[112] shows that public outcry may sometimes be very powerful in its ability to perpetuate reform.

Due to the dual capacity of online intermediaries who, on the one hand, make private, business-related decisions regarding content management, and on the other hand, fulfill governmental duties of law enforcement when engaging in content adjudication,[113] we leave the aspect of legal scrutiny for future research. In the remaining discussion, we explore whether members of the public understand online copyright enforcement policies; whether they enjoy sufficient opportunities to challenge such policies; and whether they have the capacity to correct erroneous decisions about online content.

To create a useful accountability toolbox in the context of algorithmic enforcement by private players, we shift away from traditional administrative law scholarship, whose inquiry into accountability "focuses inordinately on formal accountability to the three branches of government,"[114] towards an alternative, decentralized model of decision making.[115] Perceiving private actors as regulatory resources capable of promoting the efficacy and legitimacy of administration leaves room for additional, occasionally informal mechanisms of accountability.[116] For instance, greater specificity as to the terms enforced by private players, as well as preservation of minimal administrative procedures, such as notice and hearing requirements, might play an important role in oversight.[117] Other possible, non-formal mechanisms of accountability include voluntary disclosures, which enable the public to access information about enforcement by private actors and generate market pressure for them to improve

109. Stop Online Piracy Act, H.R. 3261, 112th Cong. (2011).

110. Preventing Real Online Threats to Economic Creativity and Theft of Intellectual Property Act of 2011, S. 968, 112th Cong. (2011) (also known as the Protect IP Act of 2011).

111. *See* Yafit Lev-Aretz, *Copyright Lawmaking and Public Choice: From Legislative Battles to Private Ordering*, 27 HARV. J.L. & TECH. 203, 204 (2013) ("Over one hundred thousand websites took part in the strike, during which some were effectively closed, while others featured information about the Bills and directed users to action centers to communicate their worries to Congress. Users zealously responded and fulminated against the Bills through posts on social networks, online petitions, and e-mails and phone calls to Congress.").

112. *Id.* ("[T]he stated positions by members of Congress on SOPA and PIPA shifted overnight from 80 for and 31 against to 55 for and 205 against.").

113. *See infra* Part III.B.

114. Freeman, *supra* note 57, at 549.

115. *Id.* at 548.

116. *Id.*

117. *Id.* at 608.

their decision-making process.[118] Accordingly, our accountability toolbox identifies three proxies for the public's practical ability to (1) understand the algorithmic decision-making process; (2) enjoy sufficient opportunities to challenge such processes; and (3) correct erroneous/improper decisions about online content.

The first proxy representing the public's ability to understand the implementation of algorithmic law enforcement is transparency.[119] Indeed, public scrutiny depends on public literacy. Without knowing that specific conduct took place, it is impossible for the public to render judgment on the merits of such conduct. In other words, transparency creates public literacy, which is necessary to establish a demand for fairness and efficiency.[120] Transparent decision-making processes expose decision makers to the risk of shaming.[121] Fearing that the public will learn about their missteps, decision makers who function in a transparent environment are discouraged from engaging in problematic conduct.[122] Furthermore, transparency also ensures that consumers exercise meaningful choice regarding which intermediary to use, and put market pressure on intermediaries to accommodate their interests.

A second proxy in our accountability toolbox, which signifies the ability of online users to challenge and contest algorithmic decisions about their online content, is due process. Scholars have increasingly acknowledged the important role of due process in facilitating algorithmic accountability. Professor Citron, for instance, coined the term "technological due process"[123] to refer to procedures designed to ensure that predictive algorithms satisfy some standard of review and revision to confirm their fairness and accuracy. Other scholars have relied on Citron's contribution and expanded its application to scoring systems[124] and predictive systems of privacy harms.[125]

Accordingly, we examine formal measures of due process (i.e., algorithmic implementations of the counter notice procedure set by the DMCA) and informal measures (i.e., voluntary dispute procedures) that purport to ensure procedural

---

118. *Id.* at 614.

119. Mark Fenster, *The Opacity of Transparency*, 91 IOWA L. REV. 885, 894 (2006) (explaining that transparency fosters an informed public debate and generates trust in and legitimacy for government).

120. Zarsky, *supra* note 34, at 1533-34.

121. *Id.* at 1534; Lawrence Lessig, *Against Transparency*, NEW REPUBLIC (Oct. 8, 2009), http://www.newrepublic.com/article/70097/against-transparency [https://perma.cc/LVT4-UHB5].

122. Zarsky, *supra* note 34, at 1534.

123. Citron, *supra* note 11, at 1301-13.

124. Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 20 (2014).

125. Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93 (2014) (relying on a "technological due process" model to address Big Data's predictive privacy harms); Neil M. Richards & Jonathan H. King, *Three Paradoxes of Big Data*, 66 STAN. L. REV. ONLINE 41, 43 (2013) (calling for a "[t]echnological [d]ue [p]rocess" solution to governmental and corporate decision-making by Big Data predictions).

fairness. The fact that copyright enforcement by online intermediaries happens to flourish on privately owned grounds should not deem procedural due process safeguards inapplicable.[126] Copyright enforcement algorithms are sovereign over crucial expressive aspects of individual lives, so we must ensure they "giv[e] subjects basic rights."[127] Affected individuals should have the rights to inspect, correct, and dispute what they believe to be inaccurate adjudication decisions made with respect to their online conduct. Otherwise, "[i]f law and due process are absent from this field, we are essentially paving the way to a new feudal order of unaccountable reputational intermediaries."[128]

Finally, a third proxy of accountability is the extent to which public oversight may effectively result in correcting errors made by algorithms. For appropriate public review depends not only on adequately explaining and justifying decision-makers' activities to the public, but also on making available accompanying mechanisms for public sanctions and corrections.[129] Indeed, public pressure can potentially force the reversal of erroneous content restrictions more quickly than reversal through the DMCA's counter notice procedure,[130] and certainly more quickly than reversal through a lawsuit. Of course, as restricted material becomes more time-sensitive and newsworthy, the role of public outcry in correcting improper restrictions of free speech becomes more crucial.[131]

In the following Part, we demonstrate that algorithmic copyright enforcement by online intermediaries ranks poorly in the accountability measures described above. Specifically, current systems employed by online intermediaries are non-transparent; they afford insufficient opportunities for affected individuals to challenge enforcement decisions; and they largely evade public oversight.

---

126. Martin H. Redish & Lawrence C. Marshall, *Adjudicatory Independence and the Values of Procedural Due Process*, 95 YALE L.J. 455, 478–89 (1986) (explaining that the underlying values of due process include transparency, accuracy, participation, and fairness).

127. Citron & Pasquale, *supra* note 124, at 19. Securing due process vis-á-vis algorithmic copyright enforcement by online intermediaries also finds support in the work of other scholars concerned about the extraordinary power of private entities: e.g., LORI ANDREWS, I KNOW WHO YOU ARE AND I SAW WHAT YOU DID: SOCIAL NETWORKS AND THE DEATH OF PRIVACY 189–91 (2012) (concluding with a proposal for a "Social Network Constitution"); REBECCA MACKINNON, CONSENT OF THE NETWORKED: THE WORLDWIDE STRUGGLE FOR INTERNET FREEDOM 240–41 (2012) (proposing ten principles of network governance); Jeffrey Rosen, *Madison's Privacy Blind Spot*, N.Y. TIMES (Jan. 18, 2014), http://www.nytimes.com/2014/01/19/opinion/sunday/madisons-privacy-blind-spot.html?_r=0 [https://perma.cc/D2HZ-LCAY] ("What Americans may now need is a constitutional amendment to prohibit unreasonable searches and seizures of our persons and electronic effects, whether by the government or by private corporations like Google and AT&T. . . . [O]ur rights to enjoy liberty, and to obtain happiness and safety at the same time, are threatened as much by corporate as government surveillance.").

128. Citron & Pasquale, *supra* note 124, at 19.

129. Jennifer Shkabatur, *Transparency With(out) Accountability: Open Government in the United States*, 31 YALE L. & POL'Y REV. 79, 80-81 (2013).

130. 17 U.S.C. § 512(g)(2)(B) (2012).

131. Sawyer, *supra* note 26**Error! Bookmark not defined.**, at 392.

III. EXPLORING ACCOUNTABILITY IN ALGORITHMIC COPYRIGHT ENFORCEMENT SYSTEMS

Does algorithmic copyright enforcement by online intermediaries contain adequate measures to ensure its accountability? In this Part, we apply our three-factor accountability framework to different systems of algorithmic copyright enforcement employed by online intermediaries to reveal their accountability rankings. We proceed in three Subparts. Subpart A presents the legal process of N&TD set by the DMCA, which establishes the baseline for the development of algorithmic copyright enforcement systems by online intermediaries. Our analysis shows that the statutory standards form a deficient starting point in terms of accountability because they fail to ensure adequate transparency, they only partly secure due process, and they provide insufficient room for public oversight.

Subpart B uses our three-factor accountability framework to analyze the accountability of two types of algorithmic copyright enforcement systems: (1) systems that largely implement the standards set by the DMCA's N&TD procedure, performing ex post removals of content; and (2) systems that go beyond N&TD, enabling blocking, filtering and demoting of content based on secret, undisclosed proprietary codes.[132]

Finally, Subpart C analyzes the accountability of YouTube's monetizing system of Content ID as an interesting hybrid of algorithmic copyright enforcement. Combining an ex ante mechanism of algorithmic content blocking with an ex post mechanism of content removal, we find the analysis of Content ID's accountability to be particularly important, considering YouTube's dominant market position and robust impact on free speech and public discourse.

A. *The Deficient Accountability Standards of the DMCA*

Since the early days of the Internet, online intermediaries were perceived as potential gatekeepers against the distribution of infringing materials.[133] Given the ever-growing threats of dispersed mass piracy, copyright owners aimed to shift some of the burden and costs of monitoring, detecting, and enforcing copyrights to online intermediaries. But the latter, who had facilitated the exchange and dissemination of user generated content (UGC), sought to avoid any cost or burden of online enforcement and to minimize potential barriers to the free flow of information, which was seen as essential to the development of their business models.[134] This battle between rights-holders and intermediaries

132. For a comprehensive overview of such voluntary mechanisms, see Annemarie Bridy, *Copyright's Digital Deputies: DMCA-Plus Enforcement by Internet Intermediaries*, *in* RESEARCH HANDBOOK ON ELECTRONIC COMMERCE LAW 185 (John A. Rothchild ed., 2016).

133. Niva Elkin-Koren, *After Twenty Years: Revisiting the Copyright Liability of Online Intermediaries*, *in* THE EVOLUTION AND EQUILIBRIUM OF COPYRIGHT IN THE DIGITAL AGE 29, 29 (Susy Frankel & Daniel J Gervais eds., 2014) ("[D]igital networks have led to an 'enforcement failure' in copyright-related industries, turning online intermediaries into key players in enforcement efforts.").

134. *Id.* at 29-30.

shaped the intermediary safe harbor regime under the DMCA.[135] Essentially, this legislation "help[ed] copyright owners ensure rapid removal of allegedly infringing material from the Internet while guaranteeing compliant OSPs[136] a safe harbor from liability for their users' acts of copyright infringement."[137]

The N&TD procedure established by the DMCA requires OSPs to respond "expeditiously" to notices of infringement by removing or disabling access to allegedly infringing material when certain conditions are met.[138] Hosting services (websites, social networks) are further required to take "reasonable steps promptly to notify the subscriber that it has removed or disabled access to the material"[139] and promptly forward any counter notices from alleged infringers back to the original complainant.[140] If, after ten to fourteen days following receipt of a counter notice, the complainant does not notify the OSP that she has filed a lawsuit, then the OSP must reinstate the contested material.[141] To maintain their immunity under the N&TD regime, OSPs cannot have actual knowledge that infringing content is on their systems or be "aware of facts or circumstances from which infringing activity is apparent."[142] Moreover, they should not receive a direct financial benefit from any infringing activity which they have the right and ability to control.[143] Finally, the DMCA further encourages compliance with N&TD by exempting OSPs from liability for mistaken yet good faith removals of material.[144]

The DMCA sets minimum standards that afford copyright owners a short,

---

135. 17 U.S.C. § 512 (2012); Jennifer M. Urban & Laura Quilter, *Efficient Process or "Chilling Effects"? Takedown Notices Under Section 512 of the Digital Millennium Copyright Act*, 22 SANTA CLARA COMPUTER & HIGH TEC. L. J. 621, 621 (2005).

136. To maintain immunity from monetary liability for material that is transmitted over networks, cached on a server, linked to, or stored at the direction of a user, OSPs were required to adopt and implement certain policies. In particular, OSPs must comply with two preliminary policies. First, they must adopt and reasonably implement a policy to terminate the accounts of repeat infringers and must notify users of this plan. Second, they must also accommodate "standard technical measures" used by copyright owners to identify infringing material. *See* 17 U.S.C. § 512(a), (b), (c), (d), (i).

137. Urban & Quilter, *supra* note 99, at 622.

138. 17 U.S.C. § 512(b)(2)(E)(i)-(ii), 512(c)(1)(C).

139. *Id.* § 512(g)(2)(A).

140. *Id.* § 512(g)(2)(B). A counter-notification must include the following: (A) a physical or electronic signature; (B) identification of the material removed and its former location; (C) statement under penalty of perjury that the user has a good faith belief the material was mistakenly removed; (D) the user's name, address, and phone number; and (E) consent to the jurisdiction of Federal District Court. *See id.* § 512(g)(3).

141. Search engines, on the other hand, are not required to notify the alleged infringer of removal because they are not expected to have any service relationship with the alleged infringer. 17 U.S.C. § 512(d); *see also* Urban & Quilter, *supra* note 99, at 626.

142. 17 U.S.C. § 512(c)(1)(A). If OSPs later become aware of such content, they must expeditiously remove it from their systems.

143. *Id.* § 512(c)(1)(B).

144. *Id.* § 512(g)(1) (Intermediaries that fail to act in good faith may lose safe harbor and may be required to pay damages to content providers whose material was unlawfully removed under the intermediaries' stated terms of use).

clear and efficient procedure for removing infringing materials from OSPs' platforms. Although the drafters of N&TD arguably did not foresee the robustness of online copyright enforcement, this procedure articulates a de facto baseline for automatic, algorithmic online copyright enforcement. Unfortunately, however, the minimum standards set by the DMCA constitute a deficient starting point in facilitating accountability, irrespective of their algorithmic application, for three reasons: (1) they are not entirely transparent; (2) they only partly secure due process; and (3) they are insufficiently exposed to public oversight.

First, the DMCA's N&TD regime generates transparency primarily by establishing a duty to notify targeted users about removals of their content.[145] A hosting service (e.g., a website, a social network) is thus required to take "reasonable steps promptly to notify the subscriber that it has removed or disabled access to the material."[146] This legal requirement, however, is insufficient to foster transparency since it fails to apply to major functions of OSPs—linking and caching—notwithstanding their significant implications for public discourse.[147]

When users search the web for content, the search results they retrieve often shape their expectations and subsequently, their knowledge of the subject of inquiry. Thus, removal of links from search indexes makes it harder for users to find and access the allegedly infringing material.[148] Because the DMCA encourages online intermediaries who receive prompt takedown notices to remove allegedly infringing content automatically,[149] online copyright enforcement mechanisms are likely to remove content when removal is inappropriate and even in cases where the claim is based on erroneous misidentification.[150] If intermediaries do not promptly notify content providers about the removal of their links, it may be too late for the providers to find out about it independently. Content providers are unlikely to check routinely that all of the content they have posted online appears in major platforms' search results. Consequently, by the time providers independently discover that their links no

145. 17 U.S.C. § 512(g)(2)(A).

146. *Id.*

147. *Id.* § 512(d); Urban & Quilter, *supra* note 99, at 681 (the counter-notification provision under section 512(g)(A) of the DMCA requires notifying a "subscriber of the service provider" whose "material is residing at the direction of the subscriber on a system or network controlled or operated by or for the service provider, or to which access is disabled by the service provider." It seems inapt to relate to the owner of a site whose content has been linked or indexed by the service provider as the provider's "subscriber" for the simple reason that search and caching services do not have account holders or subscribers); *see also* Perfect 10, Inc. v. Google, Inc., No. CV 04–9484 AHM (SHx), 2010 WL 9479059, at * 4 (C.D. Cal. Jul. 26, 2010). Furthermore, the information location links provided by § 512(d) service providers do not fit appropriately into the definition of "material residing at the direction" of the subscriber because they are generated by the service provider. Seng, *supra* note 7, at 428-29.

148. Urban & Quilter, *supra* note 99, at 682.

149. 17 U.S.C. § 512(c)(1)(C) (2013).

150. This is not to say the humans cannot make mistakes too. The argument is that in black-or-white cases, where takedown notices are unequivocally erroneous (because they identify the wrong copyrighted work, for instance), humans are likely to get it right and decline the claim.

longer appear in the search results of major platforms, the harm to their businesses may be irreparable. Furthermore, content providers cannot count on public outcry to contest the removal of the links to their content from the search results. Indeed, unless users are looking for specific content, they are unlikely to realize that such materials are no longer available.[151] Users of search engines may sometimes lack precise knowledge of which content they are looking for, otherwise they might have accessed it directly. Hence, by the time content providers could appreciate that links to their content are no longer available through major search engines, they could illegitimately lose a lot of traffic.

Furthermore, the DMCA's provisions not only deprive link providers of promptly learning that their links have been removed, but they also deprive them of learning the cause for the removal. Recognizing that a specific link has been removed and understanding that the cause for removal was copyright-related are both essential preconditions for targeted users, as well as third parties, to contest the removal of their links. Links may also be removed due to non-copyright related reasons, such as child pornography, defamation or simply because a bug caused them to be inaccessible. Hence, at least with respect to removals that are based on § 512(d), a genuine transparency problem is embedded in the DMCA's framework, making it difficult for targeted information providers to acquire knowledge about the removal of their links.

The second reason why the DMCA's framework constitutes a deficient starting point in facilitating accountability relates to its mistreatment of due process. To begin with, the DMCA's counter notice procedure[152] provides impractical challenging opportunities for alleged infringers operating in a robust sphere of online copyright enforcement. Consider for example the domain 4shared.com, which received an average of 1166 removal requests a week from August 2016 to September 2016, amounting to less than 5% of the domain's total indexed URLs.[153] It seems like a heavy burden for the domain owner to employ

151. Niva Elkin-Koren, *Let the Crawlers Crawl: On Virtual Gatekeepers and the Right to Exclude Indexing*, 26 U. DAYTON L. REV. 179, 180-81 (2001).

152. 17 U.S.C. § 512(g)(2)(A). A counter-notification must include the following: (A) a physical or electronic signature; (B) identification of the material removed and its former location; (C) statement under penalty of perjury that the user has a good faith belief the material was mistakenly removed; (D) the user's name, address, and phone number; and (E) consent to the jurisdiction of Federal District Court. *Id.* § 512(g)(3). If after ten to fourteen days, the complainant does not notify the webhost that it has filed a lawsuit, then the webhost must reinstate the contested material. Otherwise the webhost risks losing its safe harbor and it may be found liable for the damages suffered by users whose content had been unlawfully restricted.

153. *See* Google, *Specified Domain: 4shared.com*, GOOGLE TRANSPARENCY REPORT (Sept. 2016)
https://www.google.com/transparencyreport/removals/copyright/domains/4shared.com [https://perma.cc/8RUS-NCLC] (this domain seems to be legitimate as more than 95% of its indexed URLs are not targeted as infringing by reporting organizations. Nevertheless, it still suffers from inadequate challenging opportunities under which the domain owner must dispute each takedown request independently); *see also* Colette Bennett, *Nintendo Issues DMCA Takedown Notice for Hundreds of Fan-Made Games*, DAILY DOT (Sept 6, 2016, 8:06 AM), http://www.dailydot.com/parsec/nintendo-pulls-fan-games [https://perma.cc/SUZ9-YBKC]

the counter notice procedure for each of these requests. Accordingly, granting a statutory counter notice right simply does not seem to fit today's vigorous world of massive, predominantly automatic,[154] online copyright enforcement.

But even for a more moderate reality of online copyright enforcement, the statutory counter notice procedure does not seem to offer adequate challenging opportunities.[155] Removing material that may qualify as fair use before notifying the alleged infringer and before giving her the opportunity to contest the removal in a hearing may result in "an extra-judicial temporary restraining order, based solely on the copyright holder's allegation of copyright infringement."[156] Requiring OSPs to reinstate materials that were subject to a counter notice in cases where complainants fail to file suit within fourteen days does not seem to cure this problem, at least not when removals concern time-sensitive expressive material.[157] Furthermore, the DMCA's framework does not require OSPs to include any explanation of the legal ramifications of filing a counter notice (e.g., the OSP's duty to reinstate the removed material, unless the complainant files a copyright infringement suit within fourteen days). Nor does the DMCA's N&TD procedure require OSPs to disclose the identity of the copyright owner and the details of the allegedly infringed copyrighted work. This obviously frustrates the ability of alleged infringers to contest content removals, which depends largely on the comprehensiveness of the takedown notice they receive.[158]

Moreover, the DMCA's safe harbor provisions effectively encourage this denial of users' "due process"[159] because they exempt OSPs from liability for mistaken yet good faith removal of material,[160] without affording a parallel protection for failure to act on a notice in good faith.[161] While copyright owners should only request that intermediaries remove content when they have a good

(describing how Game Jolt, an open source indie gaming community received hundreds of DMCA takedown notice for allegedly infringing uses of Nintendo products).

154. Today, instead of human review of content that may be infringing, there are companies, such as Total Wipes, that employ robots to automatically send takedown requests to some of the world's most famous online services, including Skype, Tor, Dropbox, LibreOffice, Python, and WhatsApp. *See*, *e.g.*, Jamie Williams, *Absurd Automated Notices Illustrate Abuse of DMCA Takedown Process*, ELEC. FRONTIER FOUND. (Feb. 24, 2015), https://www.eff.org/deeplinks/2015/02/absurd-automated-notices-illustrate-abuse-dmca-takedown-process [https://perma.cc/NEK2-QARG].

155. Urban & Quilter, *supra* note 99, at 639.

156. *Id.*

157. *Id.* at 637.

158. For instance, Canada employs a regime of "Notice and Notice" under which ISPs who receive a notice from a copyright holder that their subscribers may be infringing copyright must forward the notice to the subscriber. In the 2012 amendments to the Canadian Copyright Act, this regime has been codified and became mandatory as of January 2nd 2015. *See Notice and Notice Regime*, OFF. OF CONSUMER AFF. (Jan. 20, 2015), http://www.ic.gc.ca/eic/site/oca-bc.nsf/eng/ca02920.html [https://perma.cc/ZWD5-F3GG].

159. JAY DRATLER, JR., CYBERLAW: INTELLECTUAL PROPERTY IN THE DIGITAL MILLENIUM § 6.05(1)(c) (2007); *see also* discussion *supra* Part II.C.3 regarding accountability and copyright policy.

160. 17 U.S.C. § 512(g)(1) (2012).

161. Urban & Quilter, *supra* note 99, at 638.

faith belief that a user's material is infringing their copyright[162]—a requirement that was recently interpreted by the Ninth Circuit in *Lenz v. Universal Music Corp.*[163] as mandating copyright holders to consider the fair use defense before sending a notice of removal—it would be naïve to expect copyright owners to make objective, court-like determinations regarding the fair use of their own, private creations. Moreover, the fear of OSPs that in reviewing content and questioning frivolous claims they will risk losing the safe harbor because they "knowingly" host infringing material,[164] makes it more likely for OSPs to err on the side of unnecessary removals, further diminishing due process. Finally, to avail themselves of the benefits of judicial review, alleged infringers must first submit a counter notice, upon which the complainant must then file suit.[165] Unless the complainant files suit, courts do not review DMCA content takedowns, and even when they do, it may often be too late, especially when the removal concerns expressive material.[166]

The third reason why the DMCA is a deficient starting point in promoting accountability in online copyright enforcement is its failure to facilitate sufficient public oversight and proper opportunities for correcting erroneous removals.[167] The DMCA's framework theoretically enables the correction of erroneous content removals by framing an ex post regime of N&TD that allows the removal of content only *after* it first appears online. This is insufficient to foster public participation. Upon receiving a notice of infringement from copyright owners, intermediaries must remove the allegedly infringing material "expeditiously."[168] Considering the immense volume of takedown requests facilitated by the DMCA's speedy enforcement regime,[169] it would be impossible for the public as a whole to promptly identify and contest inappropriate content removals, especially when the content removed does not seem to have any groundbreaking potential.[170]

B. *Regulated Versus Voluntary Mechanisms of Algorithmic Copyright Enforcement*

The previous Subpart showed that the procedural standards of N&TD set by the DMCA fail to offer a sufficient framework for accountability in online copyright enforcement. The following discussion demonstrates that the

162. 17 U.S.C. § 512(g)(1).

163. Lenz v. Universal Music Corp., 801 F.3d 1126, 1135-36 (9th Cir. 2015).

164. *See supra* note 142 and accompanying text.

165. Urban & Quilter, *supra* note 99, at 628.

166. *Id.* at 637.

167. *See supra* discussion about the virtues of accountability in Part II.D.

168. 17 U.S.C. § 512(c)(1)(C) (2013).

169. *See supra* note 7.

170. Some private initiatives have attempted to make it easier for target users and third parties to keep track with DMCA content removals, but they too seem insufficient to enhance public review over such a ubiquitous system. *See infra* Part IV.B.

algorithmic implementation of online copyright enforcement makes the lack of accountability even more severe. Particularly, the lack of accountability arises in three aspects: (1) the non-transparent implementation of algorithmic copyright enforcement imposes serious limitations over public literacy; (2) its automatic application restrains the ability of users to challenge it, further diminishing due process; and (3) its robustness further reduces the likelihood of public-driven corrections.

Regulated mechanisms of algorithmic copyright enforcement refer to algorithms that largely implement the DMCA's N&TD regime, and therefore seek to comply with its settled accountability standards. Dominant intermediaries, such as Google,[171] Facebook,[172] and Twitter,[173] apply the DMCA's framework of N&TD using algorithms. Facebook, for instance, allows rights-holders (and their legal representatives) to file a notice of copyright infringement using an online form, containing all of the detailed information required by 17 U.S.C. § 512(c)(3)(A).[174] When it does remove content, Facebook sends a warning to the user who posted the content, notifying her that content posted to Facebook was removed due to a notice of copyright infringement.[175] Facebook also provides the alleged infringer with the claimant's contact information and allows her to submit a counter-notification, *if* the content was removed under the notice and counter notice procedures of the DMCA.[176] Google,[177] Twitter[178] and YouTube[179] apply

171. *See Removing Content from Google*, GOOGLE SUPPORT (2016), https://support.google.com/legal/troubleshooter/1114905?rd=2 [https://perma.cc/T5M9-NTV4].

172. *See Reporting Copyright Infringements*, FACEBOOK HELP CTR. (2016), https://www.facebook.com/help/400287850027717 [https://perma.cc/Y6RG-PVC5].

173. *See Copyright Policy*, TWITTER HELP CTR. (2016) https://support.twitter.com/articles/15795-copyright-and-dmca-policy [https://perma.cc/8N82-2VME].

174. Facebook, for instance, requires the following information to be included in a notice of copyright enforcement:

> The fastest and easiest way to submit a claim of copyright infringement to us is to use our online form. It may be required by law that you include the following information: Your complete contact information (full name, mailing address and phone number). Note that we regularly provide your contact information, including your name and email address, the name of your organization or client who owns the right in question, and/or the contents of your report to the person who posted the content you are reporting. You may wish to provide a professional or business email address for contact by users. A description of the copyrighted work that you claim has been infringed. A description of the content on our site that you claim infringes your copyright. Information reasonably sufficient to permit us to locate the material on our site. The easiest way to do this is by providing web addresses (URLs) leading directly to the allegedly infringing content. A declaration that you have a good faith belief that use of the copyrighted content described above, in the manner you have complained of, is not authorized by the copyright owner, its agent, or the law, the information in your notice is accurate, and you declare, under penalty of perjury, that you are the owner or authorized to act on behalf of the owner of an exclusive copyright that is allegedly infringed. Your electronic signature or physical signature.

See What Should I Include when Submitting a Report to Facebook Alleging Infringement of My Copyright?, FACEBOOK HELP CTR. (2016), https://www.facebook.com/help/400287850027717 [https://perma.cc/74TF-72DK].

175. *Id.*

176. This suggests that Facebook also removes content by algorithms that do not

similar algorithmic implementations of the DMCA's N&TD procedure.

Nevertheless, these prominent platforms occasionally go beyond the framework of the DMCA, exploiting algorithms to filter, block and demote content, sometimes before it ever becomes available online. While these voluntary mechanisms shift the heavy burden of policing online copyright infringement from copyright owners and courts to algorithms, they create greater challenges to accountability because they are not subject to any regulation—not even that of the DMCA.

Among these voluntary algorithms are filters that allow the ex ante blocking of content. In late 2007, some UGC[180] sites, including MySpace, Veoh, DailyMotion, and Soapbox (via Microsoft), collaborated with large content companies, including Disney, CBS, NBC Universal, and Viacom, in recommending that UGC sites use copyright filtering technologies.[181] These technologies essentially compare uploaded material against samples of copyrighted material (Reference Material) provided by copyright owners. When users attempt to upload content that matches any Reference Material, the content can be blocked before it ever becomes available online.[182]

Additionally, search engines implement algorithms, such as Google's Pirate algorithm, to combat online copyright infringement.[183] Updated in 2014,[184] this algorithm essentially changes the search algorithm so that it downgrades the

---

implement the DMCA's framework of ex post content removal.

177. *See Submit a Copyright Takedown Notice*, YOUTUBE HELP (2016), https://support.google.com/youtube/answer/2807622 [https://perma.cc/E2ND-5BTM].

178. *See How Do I File a Copyright Claim*, TWITTER HELP CTR. (2016), https://support.twitter.com/entries/15795#5 [https://perma.cc/P8AT-PDZB].

179. *See supra* note 177.

180. For analysis of the legal issues related to user-generated content, see Niva Elkin Koren, *Governing Access to User-Generated-Content: The Changing Nature of Private Ordering in Digital Networks*, *in* GOVERNANCE, REGULATIONS AND POWERS ON THE INTERNET 318-43 (2012); Tom W. Bell, *The Specter of Copyism v. Blockheaded Authors: How User-Generated Content Affects Copyright Policy*, 10 VAND. J. ENT. & TECH. L. 841 (2008) (providing an economic view of UGC and predicting that UGC will drive down the costs and increase accessibility for all content); Greg Lastowka, *User-Generated Content and Virtual Worlds*, 10 VAN. J. ENT. & TECH. L. 893 (2008); Edward Lee, *Warming Up to User-Generated Content*, 5 U. ILL. L. REV. 1459 (2008).

181. PRINCIPLES FOR USER GENERATED CONTENT SERVICES, http://ugcprinciples.com [https://perma.cc/Q6Y3-BUFS].

182. Sawyer, *supra* note 26, at 365. A classic example for a filtering technology is the algorithm applied by YouTube's system of Content ID, which automatically scans videos as they are uploaded for copyrighted material and blocks access to videos that contain material which copyright owners have asked YouTube to block. *See infra* Part III.C.

183. Glenn Gabe, *Google's Pirate Algorithm and DMCA Takedowns: Exploring the Impact Threshold*, G-SQUARED INTERACTIVE (Dec. 9, 2013), http://www.hmtweb.com/marketing-blog/google-pirate-algorithm-dmca [https://perma.cc/EED9-RPWG]; Stuart Long, *Google Rolls Out DMCA Algorithm Update After More than 2 Years*, BRANDED3 (Oct. 27, 2014), http://www.branded3.com/blogs/google-rolls-dmca-algorithm-update-2-years [https://perma.cc/3ZR5-ZFAW].

184. *Continued Progress on Fighting Piracy*, GOOGLE PUB. POL'Y BLOG (Oct. 17, 2014), http://googlepublicpolicy.blogspot.co.il/2014/10/continued-progress-on-fighting-piracy.html [https://perma.cc/RLX9-QDUG].

ranking of allegedly pirated websites, placing them at the bottom of the search results list. Unfortunately, this method leaves the door open for manipulation by strategic players. For example, competing websites could abuse this mechanism as a tool for sabotage. By filing sham notices of copyright infringement against their competitors, they can increase the likelihood that a content-demoting algorithm will kick in and make their competitors' websites appear lower in the search engine's search results list. Presumably, politicians could also leverage this mechanism as a reputation management strategy.[185]

We now turn to show that algorithmic implementations of the statutory N&TD procedure do not adequately facilitate the three aspects of accountability, and that voluntary measures of algorithmic copyright enforcement, which go beyond N&TD, fare even worse. While advocates of voluntary measures of algorithmic copyright enforcement maintain that they solve many of the problems associated with online infringements,[186] we demonstrate that they leave almost no room for public scrutiny, notwithstanding their robust implications for public discourse.[187]

1. *Transparency*

Algorithmic copyright enforcement by online intermediaries imposes serious limitations on transparency because both regulated and voluntary measures fail to fully disclose how they exercise their power. It is unclear what precise circumstances trigger the underlying enforcing algorithms. One could expect that algorithmic N&TD methods would be slightly more predictable than voluntary methods, as they largely adhere to the DMCA framework. However, these methods still retain some unrevealed discretion to determine when copyright infringement has occurred and to decide when to apply the statutory N&TD takedown procedure automatically.

For instance, although the statutory language uses the imperative language "acts expeditiously to remove, or disable access"[188] when defining the responsibilities of an intermediary seeking to enjoy the DMCA safe harbor in relation to allegedly infringing content on its platform, Facebook admits that after receiving a claim of copyright infringement, it *may* remove the reported content from its platform.[189] When exactly this would happen is largely a mystery.

185. For example, the DMCA has been used to target political content on YouTube. Elliot Harmon, *Once Again, DMCA Abused to Target Political Ads*, ELEC. FRONTIER FOUND. (Nov. 17, 2015), https://www.eff.org/deeplinks/2015/11/once-again-dmca-abused-target-political-ads [https://perma.cc/FQ5G-U8JL] (reporting on abuse of the DMCA to takedown a television ad supporting a controversial proposal to regulate short-term property rental services like Airbnb).

186. Amir Hassanabadi, *Viacom v. Youtube: All Eyes Blind: The Limits of the DMCA in a Web 2.0 World*, 26 BERKELEY TECH. L.J. 405, 438 (2011).

187. *See supra* Part II.C.2.

188. 17 U.S.C. § 512(b)(2)(E)(i)-(ii), 512(c)(1)(C) (2012).

189. *See What Happens After I Submit a Claim of Copyright Infringement to Facebook?*, FACEBOOK HELP CTR. (2016), https://www.facebook.com/help/400287850027717

Algorithmic implementations of N&TD are not as straightforward as we would have expected automated systems to be. Occasionally, they disregard takedown notices and elect not to remove allegedly infringing content.[190]

The use of voluntary measures makes the situation described above even less transparent. Algorithms are employed to filter, block, and demote content based on an entirely undisclosed, self-determined threshold. Without understanding the nuances of their decision-making processes, it is impossible to hold such algorithms accountable.

2. *Due Process*

Holding algorithmic mechanisms of copyright enforcement accountable also depends on the availability of adequate channels to contest content restrictions. Nonetheless, neither regulated mechanisms of algorithmic law enforcement nor voluntary ones facilitate adequate challenging opportunities, consequently diminishing due process. In the case of regulated mechanisms of N&TD, this is because they often fail to comply with the already-deficient standard of counter notice set by the DMCA.[191] For instance, under YouTube's algorithmic implementation of the DMCA N&TD, the counter notice feature only becomes available if the alleged infringer elects to dispute the infringement claim.[192] But an alleged infringer who receives a Content ID claim of copyright infringement[193] may elect not to dispute the claim, and instead either acknowledge it, remove the allegedly infringing material, swap out the allegedly infringing audio track with one of YouTube's free-to-use songs, or share revenue with the copyright owner.[194] By choosing one of these alternative options, rather than disputing the claim, the content provider effectively forgoes his ability to file a counter notice. But under the DMCA, counter notices should be available in theory and in practice, regardless of the intermediary's internal business model.

Furthermore, anecdotal evidence proves that sometimes, counter notices fail to fulfill their statutory role in ensuring intermediaries repost content they had

---

[https://perma.cc/RW7G-EPQ8].

190. For instance, Twitter was accused of disregarding DMCA takedown notices filed by an artist named Christopher Boffoli, who created the popular "Disparity Series," consisting of photographs featuring miniature figures in funny poses on various types of food. Since his photographs went viral, Boffoli has been sending takedown requests under the DMCA to individuals and sites like Facebook, Pinterest, and Google. In his complaint to a U.S. district court, Boffol accused Twitter of inducing copyright infringement and failing to disable access to copyrighted material even after being notified about infringing uses. *See* Jon Brodkin, *Twitter Won't Take Down "Giant Food" Photos, so Artist Sues*, ARS TECHNICA (Sept. 11, 2012, 2:52 PM), http://arstechnica.com/tech-policy/2012/09/11/twitter-wont-take-down-tiny-food-photos-so-artist-sues [https://perma.cc/7RAC-E5DN].

191. *See supra* Part III.A.

192. As we explain henceforth in Part III.C, YouTube's Content ID procedure is a hybrid of regulated and voluntary copyright enforcement.

193. *See infra* note 213 and accompanying text.

194. *What Is a Content ID Claim?*, YOUTUBE HELP (2016), https://support.google.com/youtube/answer/6013276 [https://perma.cc/H6FW-MUMW].

previously removed due to a DMCA notice of infringement. Consider the following message that an artist named John McKelvey received from YouTube as a response to a counter notice he had submitted after his video, which included short portions of songs taken from Eric B. & Rakim's record for critical purposes, was taken down due to a copyright infringement notification filed by Universal: "Thank you for your counter-notification. The complainant has reaffirmed the information in its DMCA notification. YouTube has a contractual obligation to this specific copyright owner that prevents us from reinstating videos in such circumstances. Therefore, we regretfully cannot honor this counter-notification."[195]

YouTube states outright that, it will not repost videos that allegedly infringe content owned by "partner" companies following their removal, even if their uploaders file a counter notice showing they constitute fair use, due to YouTube's contractual obligations to the copyright owners involved.[196] Considering that their counter notices may be disregarded due to YouTube's internal business relations, it should not surprise anyone why potential target users rarely bother to file counter notices.

Additionally, regulated mechanisms of copyright enforcement often impede users' ability to file counter notices.[197] Consider the following example of an actual notice of removal sent by Facebook's algorithm to a user:

> We have removed your video entitled (no title) uploaded at 9:00am February 3rd, 2014. This video may include copyrighted material (such as a clip or audio) that you do not have the right to share.
> If you think your video should not have been removed because:
> (1) you are the copyright owner, or
> (2) you have permission from the copyright owner to upload and distribute this material on Facebook, or

---

195. Andy, *YouTube's Deal with Universal Blocks DMCA Counter Notices*, TORRENTFREAK (Apr. 5, 2013), http://torrentfreak.com/youtube-deal-with-universal-blocks-dmca-counter-notices-130405 [https://perma.cc/6UZB-YJ3A]; *see also* Mike Masnick, *YouTube Won't Put Your Video Back up, Even if It's Fair Use, if It Contains Content from Universal Music*, TECHDIRT (Apr. 5, 2013, 11:52 AM), https://www.techdirt.com/articles/20130405/01191322589/youtube-wont-put-your-video-back-up-even-if-its-fair-use-if-it-contains-content-universal-music.shtml [https://perma.cc/3HKH-Y7LT].

196. Andy, *supra* note 195. YouTube claims the following:

> YouTube enters into agreements with certain music copyright owners to allow use of their sound recordings and musical compositions. In exchange for this, some of these music copyright owners require us to handle videos containing their sound recordings and/or musical works in ways that differ from the usual processes on YouTube. Under these contracts, we may be required to remove specific videos from the site, block specific videos in certain territories, or prevent specific videos from being reinstated after a counter notification. In some instances, this may mean the Content ID appeals and/or counter notification processes will not be available. Your account will not be penalized at this time.

Id.

197. Under YouTube's application of N&TD, the process of counter notice only becomes available if the alleged infringer elects to dispute the Content ID claim, and the copyright owner can respond to a dispute in several different ways: (1) release the claim; (2) uphold the claim; (3) take down the video by submitting a formal notice. *See infra* note 216 and accompanying text.

> (3) you otherwise believe you are legally entitled to upload and distribute this material on Facebook you may visit the link below to video an appeal requesting that it be reinstated:
> [link redacted]
> If you do not want to appeal, there is no need to take any action. Please be careful about videos you upload in the future. If they are identified as possibly containing copyright infringing material, they may also be removed. This could result in us temporarily or permanently blocking your ability to upload videos, or permanently disabling your account.[198]

This notice does not foster due process. To begin with, an alleged infringer receiving such a notice does not necessarily understand its merit.[199] Even if she does, it would be hard for her to dispute this notice because it omits crucial information that an alleged infringer must have in order to properly dispute content removal; most importantly, the details of the allegedly infringed work and the identity of the copyright owner requesting the takedown.

With voluntary algorithmic copyright enforcement, due process is completely non-existent. Since algorithms replace copyright owners in policing allegedly infringing content, users cannot issue a counter notification and threaten suit against copyright owners.[200] Indeed, when content is removed due to a notice of copyright infringement according to the statutory N&TD regime, content providers are notified about the removal (except link providers) and based on this notification, they can file a counter notice, dispute the removal and force reposting of removed content within ten to fourteen business days. However, when content is filtered automatically by filtering technologies and without a takedown notice, content providers cannot threaten suit against copyright owners, for the owners played no role in the removal.[201] Nor can content providers bring suit against the intermediaries for removing the content because online intermediaries are usually protected from such suits under their terms-of-use agreements.[202] Going beyond the statutory framework, voluntary mechanisms of algorithmic copyright enforcement do not afford alleged infringers with even the minimum due process protections set by the DMCA: they do not grant alleged infringers the right to contest content restrictions through a counter notice procedure, and they do very little in terms of validating copyright ownership rights.[203]

Moreover, although the statutory N&TD regime requires complainants to include in their takedown request an affidavit validating their copyright

198. Wasylik, *supra* note 3.

199. Lemley, *supra* note 73, at 115.

200. Sawyer, *supra* note 26, at 385.

201. *Id.*

202. *Id.*

203. The DMCA establishes a procedure—filing a takedown notice requires an affidavit. 17 U.S.C. § 512(c)(3) (2012). If there is no procedure of that sort in voluntary measures, some materials will be removed without any sufficient legal ground, which could make removal difficult for the user to contest.

ownership,[204] voluntary mechanisms occasionally restrict content without any sufficient legal ground. For instance, in a recent complaint filed against YouTube under its terms of use, the plaintiff alleged that Content ID removed his parody of the film *The Girl With the Dragon Tattoo*, which Content ID designated as being owned by Pirateria, although Pirateria is not the owner of the rights to this film.[205] The plaintiff further argued that when he posted, under fair use, a critique of the 2014 remake of *Teenage Mutant Ninja Turtles*, a Content ID claim was made with YouTube on behalf of Viacom, although Viacom is not the true copyright owner.[206]

Unfortunately, as mentioned above, target users cannot sue copyright owners for improper content restrictions, as copyright owners play no active role in detecting copyright infringement under voluntary regimes. Nor can they sue the intermediaries for restricting access to their content, since intermediaries, as private, profit-maximizing entities, can easily prevent these sorts of suits under their terms-of-use. As a result, target users are left without proper legal recourse against illegitimate content restrictions by voluntary mechanisms of algorithmic copyright enforcement.

3. *Public Oversight*

By removing material after the fact, after receiving a notice of copyright infringement from a copyright owner, regulated mechanisms of algorithmic enforcement seem to facilitate public oversight better than voluntary mechanisms, which restrict access to content ex ante. Nevertheless, given the immense volume of takedown requests facilitated by algorithmic copyright enforcement, it may be impossible for the public to follow, inspect and contest inappropriate content removals, even when content is removed ex post, pursuant to the DMCA's N&TD procedure. Indeed, according to Google's Transparency Report, in June 2015 Google alone received requests from copyright owners to remove 39,013,486 URLs.[207] That is more than one million removal requests per day. Reviewing the practices of such a robust system of online copyright enforcement is simply impractical.

Voluntary mechanisms of algorithmic copyright enforcement employed by online intermediaries permit almost no public oversight whatsoever as they operate behind the veil of the intermediaries' proprietary code. Hosting platforms occasionally apply ex ante algorithms to block and filter content before it ever becomes publicly available,[208] whereas search engines employ algorithms that downgrade sites that link to allegedly infringing content, without signaling to

204. 17 U.S.C. § 512(c)(3)(A)(vi) (2013).

205. *See supra* note 2.

206. *See supra* note 2.

207. Google, *Requests to Remove Content Due to Copyright*, GOOGLE TRANSPARENCY REPORT (Apr. 2016), https://www.google.com/transparencyreport/removals/copyright (last visited May 8, 2016).

208. *See infra* Part III.C.

users that some materials are no longer available. Under such circumstances, public corrections are extremely unlikely. How can the public possibly contest the removal of content it has never been aware of in the first place?

C. *Corporate Copyright: YouTube's Content ID*

In the previous Subpart, we explained three different reasons that contribute to difficulties in holding copyright enforcement algorithms accountable for making copyright-related determinations. In the following paragraphs, we explore the operation of YouTube's monetizing system of Content ID as an interesting hybrid of algorithmic copyright enforcement, combining an ex ante mechanism of algorithmic content blocking with features of DMCA-style mechanisms of ex post content removal. Content ID's accountability is particularly important, not only due to YouTube's robust implications for public discourse,[209] but also because Content ID has turned algorithmic copyright enforcement into a private-financial model of copyright enforcement.[210]

Content ID codifies an advanced set of copyright policies and content management tools, which shift the burden of policing copyright infringement from copyright owners to identification technologies, while protecting copyright owners' interests beyond the basic removal process provided by the DMCA. As a practical matter, Content ID allows copyright owners to identify their works using a digital identifying code.[211] Basically, Content ID notifies subscribed copyright owners whenever a video uploaded to YouTube matches a work they own, offering them one of four choices: (1) mute audio that matches their music; (2) block a whole video from being viewed; (3) monetize the video by running ads against it; or (4) track the video's viewership statistics.[212]

When Content ID identifies a match between content provided by a target

209. YouTube is the number one national video website and the third most visited website globally. *See* Wendy Boswell, *Video Websites: The Top Ten*, ABOUT.COM (Mar. 6, 2016), http://websearch.about.com/od/imagesearch/tp/popularvideosites.htm [https://perma.cc/QQ2A-52FR]; *see also The Top 500 Sites on the Web*, ALEXA, http://www.alexa.com/topsites [https://perma.cc/XC9S-ETAZ] (last visited Apr. 20, 2016).

210. Content ID currently scans over 400 years-worth of video and utilizes more than 25 million references files of more than 5,000 partners, including US network broadcasters, record labels, and movie studios. *See Statistics*, YOUTUBE, http://www.youtube.com/yt/press/statistics.html [https://perma.cc/UV9R-HYPZ] (last visited Apr. 20, 2016).

211. The technology underlying Content ID relies on digital fingerprinting to sample an uploaded file and compare it against a database of reference files provided by participating copyright owners. *See* Brad Stone & Miguel Helft, *New Weapon in Web War over Piracy*, N.Y. TIMES (Feb. 19, 2007), http://www.nytimes.com/2007/02/19/technology/19video.html [https://perma.cc/3QVD-VA22] (discussing fingerprinting technologies for identifying audio and video). For a technical discussion of how fingerprinting is used to identify copyrighted content, see Craig Seidel, *Content Fingerprinting from an Industry Perspective*, 2009 IEEE INT'L CONF. ON MULTIMEDIA AND EXPO 1524.

212. *How Content ID Works*, YOUTUBE HELP, https://support.google.com/youtube/answer/2797370?hl=en [https://perma.cc/PUZ5-C6TR] (last visited Apr. 20, 2016).

user and a copyrighted work on file, it automatically sends a Content ID claim to the target user. At this point the user can either: (1) acknowledge the claim; (2) if the claim is for a piece of music in the video, choose to remove the song without having to edit and reload; (3) swap out the allegedly infringing song with a free-to-use song; (4) share revenue with the copyright owner; or (5) dispute the claim.[213]

The system includes a very detailed mechanism of disputing a Content ID claim. If the target user elects to dispute the claim, the copyright owner gets thirty days to respond.[214] Failure to respond promptly causes the Content ID claim to expire.[215] The copyright owner can respond to a dispute in several different ways: (1) release the claim; (2) uphold the claim; (3) takedown the video. If the copyright owner elects to uphold the claim, the user may be able to appeal her decision.[216] If the user appeals the owner's decision to uphold the Content ID claim, the copyright owner gets thirty days to respond. After the user appeals, the copyright owner can either release the claim or takedown the video.[217]

If the owner elects to takedown the video, the user receives a copyright strike in her account.[218] This means that her video has been taken down from YouTube following the legal request of the copyright owner. Receiving a copyright strike puts the user's account in bad standing, which means that she may lose access to certain YouTube features.[219] A user cannot resolve a copyright strike by deleting the allegedly infringing video. Instead, she can either: (1) wait six months for the copyright strike to expire and complete Copyright School, while receiving no

---

213. *Dispute a Content ID Claim*, YOUTUBE HELP, https://support.google.com/youtube/answer/2797454 [https://perma.cc/8FHL-HFVJ] (last visited Apr. 20, 2016).

214. *Id.*

215. *Id.*

216. *Id.* This depends on whether the user's account is in good standing on the date of the appeal and possibly on other undisclosed factors. An account is in "good standing" when it has no Community Guidelines strikes (content removals that are not copyright-related, i.e. sexual content or violence), no copyright strikes, and no more than one video blocked worldwide by Content ID. *Keep Your YouTube Account in Good Standing*, YOUTUBE HELP, https://support.google.com/youtube/answer/2797387 [https://perma.cc/4YZA-GLFR] (last visited Apr. 20, 2016).

217. *What Happens After I Dispute*, YOUTUBE HELP, https://support.google.com/youtube/answer/2797454 [https://perma.cc/DRA7-GGKL] (last visited Apr. 20, 2016).

218. *Id.*

219. Particularly, she will be restricted from uploading videos as unlisted, uploading videos that are longer than 15 minutes, uploading videos under a Creative Commons license, InVideo programming, customizing video thumbnails, live events and hangouts on Air, appealing rejected Content ID claim disputes, sharing private videos, and using YouTube Video Editor. In addition, a user receiving a copyright strike will be restricted from using certain channel features and joining the YouTube Partner Program (allowing creators to monetize content on YouTube through a variety of ways including advertisements, paid subscriptions, and merchandise). *Keep Your YouTube Account in Good Standing*, YOUTUBE HELP, https://support.google.com/youtube/answer/2797387 [https://perma.cc/4YZA-GLFR] (last visited Apr. 20, 2016).

additional copyright strikes during this period of time; (2) contact the person who claimed the video and ask him to retract his claim of copyright infringement; (3) submit a counter notification, if the user believes her video either qualifies as fair use or is not infringing.[220]

According to YouTube's website, a counter notification is:

> A legal request for YouTube to reinstate a video that has been removed for alleged copyright infringement. The process may only be pursued in instances where the upload was removed or disabled as a result of a mistake or misidentification of the material to be removed or disabled, such as fair use. It should not be pursued under any other circumstances.[221]

Only the original uploader or an agent authorized to act on her behalf may submit a counter notification.[222] YouTube forwards the counter notification to the party who submitted the original Content ID claim of copyright infringement.[223] This process takes ten days to complete.[224] Users whose accounts have been suspended for multiple copyright violations cannot access the counter notification webform, and may instead submit a free-form counter notification.[225]

The major advantage offered by Content ID is the monetizing option. Rather than simply blocking allegedly infringing materials, the monetizing feature of Content ID promotes copyright licensing agreements, enabling copyright owners to permit allegedly infringing uses of their works in return for a share in the economic benefit generated by advertisements displayed with the content.[226] The vast majority of copyright holders elect not to remove infringing content, but instead to monetize it,[227] so much so that video monetizing is becoming a successful business model.[228] Indeed, copyright owners can build a profitable online licensing business through which they can obtain revenue easily from YouTube users who make use—any use—of their copyrighted material. It is not that copyright licensing was non-existent before the emergence of Content ID; it

---

220. *How to Resolve a Copyright Strike*, YOUTUBE HELP, https://support.google.com/youtube/answer/2814000 [https://perma.cc/9NM5-5AMC] (last visited Apr. 20, 2016).

221. *Id.*

222. *Counter Notification Basics,* YOUTUBE HELP, https://support.google.com/youtube/answer/2807684?hl=en-GB [https://perma.cc/YXL8-KWS3] (last visited Apr. 20, 2016).

223. *Id.*

224. *Id.*

225. *Id.*

226. Yafit Lev-Aretz, *Second Level Agreements,* 45 AKRON L. REV. 137, 152 (2012).

227. *Id.* at 158.

228. To name one example, Rumblefish, Inc., recently acquired by SESAC, is a full service music micro-licensing platform for rights-holders providing monetization on YouTube. Bruce Houghton, *Online Video Monetization Heats up: Q&A with Rumblefish Founder/CEO on Acquisition by SESAC*, HYPEBOT.COM (Aug. 13, 2014), http://www.hypebot.com/hypebot/2014/08/online-video-monetization-heats-up-qa-with-rumblefish-founderceo-on-acquisition-by-sesac.html [https://perma.cc/6PX4-JG5M].

is just that the latter made this process shorter, clearer, and more efficient.[229]

Further, the automatic identification technology lowers the transaction costs associated with entering into licensing agreements.[230] It identifies the potential parties (a copyright owner on the one side, and a user on the other) and facilitates a contract by a click of button. Moreover, this business model effectively flips the fundamental default established by copyright law: authorize first, use later. With Content ID, owners allow users to experiment with their works without giving them any explicit license, while exercising their rights to get a share in revenue only after the (unlicensed) derivative uses become commercially viable (use first, authorize later).

Standing alone, facilitating copyright licensing in an era of mass infringement seems both welcomed and appreciated.[231] Nevertheless, Content ID may err in reporting a technological match over fair use,[232] forcing users to acquire a license where they should in fact be free to use the materials.[233] Furthermore, in the absence of a reliable procedure for validating copyright ownership, some works may be erroneously claimed by non-copyright-owners.[234] Considering the crucial role YouTube plays in shaping public discourse,[235] it is critical to hold Content ID accountable for its determinations. That said, figuring out how Content ID exercises its power is extremely challenging because it operates behind the veil of a proprietary code that primarily adheres to YouTube's business interests. Against this backdrop, we next examine the accountability of Content ID.

1. *Transparency*

Content ID allows content providers to upload digital copies of video or audio works, which YouTube's servers use to create a digital reference against which all other videos on the site are scanned. If even a portion of another video matches the sample in either its visual or audio content, the video is flagged as containing that copyrighted content.[236] It is unclear, and hence unpredictable,

---

229. Lev-Aretz, *supra* note 226, at 158.

230. Paul J. Heald, *How Copyright Makes Books and Music Disappear* 30 (Illinois Public Research Paper No. 13-54, 2013), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2290181 [https://perma.cc/6X2D-BUHV].

231. *Id.*

232. *See* Maayan Perel & Niva Elkin-Koren, *Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement*, FLA. L. REV. (forthcoming 2017).

233. Note that it is not unlikely that alleged infringers will accept an unfair deal and agree to share advertising profit with copyright owners, notwithstanding their non-infringing uses. Under Content ID's rapid takedown policy, which allows content to be removed automatically before, and sometime absent, prior notice to content providers and without adequate legal scrutiny (both adjudication before removal and review after removal), agreeing to share profit with copyright owners at least guarantees both the access to the popular distribution channel and the prospective income, whereas declining such a deal unfortunately risks losing both.

234. *See supra* note 79 and accompanying text.

235. *See supra* Part II.C.2

236. Lauren G. Gallo, *The (Im)possibility of "Standard Technical Measures" for UGC Websites*,

what exact portion of copyrighted material must be embedded in an upload to trigger the system.[237] As a result, Content ID may unlawfully flag fair uses or de minimis uses of content.[238] Because the thresholds of Content ID are not disclosed,[239] users cannot appreciate how it exercises its power. Indeed, the non-transparent nature of Content ID's identification procedure curtails the ability of alleged infringers to learn whether their videos were flagged due to an alleged copyright infringement or otherwise by reason of mistake or misidentification.

2. *Due Process*

Content ID implements the DMCA's counter notice feature in a way that diminishes due process. As noted earlier,[240] an alleged infringer must dispute a Content ID claim before availing herself of the counter notice application.[241] Yet, YouTube implicitly discourages users from filing a dispute, emphasizing that an invalid dispute may cause the copyright owner to takedown the allegedly infringing video, and when this happens, the target user's account could be subject to restrictions imposed by a copyright strike.[242] Note, that a copyright strike is against a *specific user* and not against a *specific account.* Thus, a user can open as many accounts as she wishes, so long she does not receive a total of three copyright strikes (against whichever account). If she receives three strikes, *all* her accounts are being terminated. Target users may hence be induced not to dispute a Content ID claim, especially when the contested content is timely.[243] Neglecting the counter notice process is especially problematic when the blocked content could have been allowed as fair use: target users are obviously deprived of their right to share and reap financial benefit from their own creations, but there are also serious policy concerns resulting from the fact that the public as a whole is precluded from hearing, viewing and experiencing the disputed content.[244]

YouTube further compromises due process when it fails to verify the rights

---

34 COLUM. J.L. & ARTS 283, 296 (2011).

237. *Id.*

238. *See* Perel &Elkin-Koren, *supra* note 232.

239. They may even vary among copyright owners.

240. *See supra* note 192 and accompanying text.

241. *See supra* note 192 and accompanying text.

242. *See supra* note 218 and accompanying text.

243. Urban & Quilter, *supra* note 99, at 626. For instance, during the 2008 McCain-Palin campaign, several political videos were removed from the McCain campaign's YouTube channel for copyright infringement allegations. McCain-Palin's counsel urged YouTube to make an exception for the videos posted by political candidates and campaigns, suggesting that YouTube commit to a legal review of these political videos and decline to remove clearly non-infringing material, rather than taking down and insisting on the DMCA waiting period of ten to fourteen business days. This was because the political speech that was removed from the McCain-Palin YouTube channel took place at the height of election season, making it pointless to wait for the material to be reinstated. *See* Wendy Seltzer, *Free Speech Unmoored in Copyright's Safe Harbor: Chilling Effects of the DMCA on the First Amendment*, 24 HARV. J.L. & TECH 171, 173 (2010).

244. *See supra* Part II.C.2

claimed by copyright owners.[245] When a claimant creates a reference file, either by uploading content or by claiming an already uploaded video, he effectively asserts ownership over the content in the reference file.[246] Apparently, YouTube does not require sufficient proof of ownership in claimed works. [247] This flaw is troubling: users who manage to gain admission to the Content ID system can upload any content into the system, which would later flag videos as belonging to them. As a result, swindlers may hijack ad revenue of videos from users (and possibly from the true right-holders), regardless of whether they really own any copyright interest in the videos.[248]

GoDigital Media Group[249] is a prominent example of this sort of conduct. Serving copyright owners who wish to identify and monetize their copyrighted works in UGC, GoDigital is responsible for thousands of illegitimate copyright claims on YouTube. AdSHARE,[250] one of GoDigital's five companies, is responsible for monetizing fan engagement online, by identifying, tracking and monetizing user-uploaded versions of copyright owners' content on social media websites. AdSHARE works on music compositions, sound recordings, and video, and is able to identify even short snippets of copyrighted content used on social media platforms, such as YouTube, Google, Facebook and Soundcloud. Unfortunately, AdSHARE does not always verify whether its clients actually own the rights to particular works before it begins monetizing them automatically and therefore, it ends up claiming a wide variety of royalty free, public domain, and Creative Commons content.

Finally, YouTube's Content ID arbitrarily favors dominant copyright holders over small creators, possibly depriving the latter of substantial due process. As described earlier, YouTube has publicly admitted that, when it comes to videos containing content from certain partner companies, it will not repost them following their removal, even if the video uploaders file a counter notice showing that the videos constitute fair use.[251] As a result, some alleged infringers targeted by YouTube's Content ID might be left with no meaningful recourse to confront unlawful content removals. Even though they can technically dispute content takedowns, such disputes are effectively pointless.

---

245. *See supra* notes 8 and accompanying text.

246. *What is a Policy?*, YOUTUBE HELP, https://support.google.com/youtube/answer/107383?hl=en&ref_topic=24332 [https://perma.cc/ZCC4-DHEQ] (last visited Apr. 18, 2016).

247. Patrick McKay, *YouTube Copyfraud & Abuse of the Content ID System*, FAIRUSETUBE (Nov. 23, 2011, 3:22 PM), http://fairusetube.org/youtube-copyfraud [https://perma.cc/82TY-LKM8].

248. *Id.*

249. GODIGITAL MEDIA GROUP, http://www.godigitalmg.com [https://perma.cc/B36R-X8Z7] (last visited Apr. 23, 2016).

250. AdSHARE, http://adshare.tv [https://perma.cc/DUC5-FW8Z] (last visited Nov. 15, 2016).

251. *See supra* notes 195-196 and accompanying text.

3. *Public Oversight*

Because YouTube's Content ID allows the ex-ante blocking of videos, it curtails the ability of the public to promptly render its own fair use judgment.[252] Without the opportunity to experience the content of blocked videos, almost no room is left for public pressure to demand the reversal of the system's determination. While users could resort to uploading videos protesting automatic blocking,[253] they could not include the underlying infringing material in their protest videos, so they are unlikely to generate sufficient public pressure to demand the repost of blocked videos.[254]

Hence, outside the imperfect option of filing a counter notice against allegedly unjustified takedowns, the filtering algorithm embedded in Content ID seems to remain wild and free. While it arguably makes the licensing process shorter, clearer and more efficient,[255] it imposes unwarranted restrictions on non-infringing materials and fair uses of content. By automatically locating and monetizing digital uses of copyrighted content, the system makes it unnecessary for owners to search and notify YouTube of infringing content. Consequently, Content ID results in many false positives,[256] or instances where Content ID automatically blocks or monetizes a video that does not actually contain infringing content.[257] For example, YouTube has recently even flagged a cat purring as copyright infringing music.[258] Finally, by lowering the transaction costs associated with entering into licensing agreements,[259] Content ID may result in unjust enrichment when targeted users are unjustifiably dragged into paying to share non-infringing content or fair use videos.

IV. ENHANCING ACCOUNTABILITY: BARRIERS AND STRATEGIES

In the previous Parts, we explained why accountability is important in algorithmic copyright enforcement, and why existing mechanisms fail to adequately promote it. In particular, we demonstrated that the non-transparent

---

252. Sawyer, *supra* note **Error! Bookmark not defined.**, at 394.

253. *See, e.g.*, Fred von Lohmann, *YouTube's January Fair Use Massacre*, ELEC. FRONTIER FOUND. (Feb. 3, 2009), http://www.eff.org/deeplinks/2009/01/youtubes-january-fair-use-massacre [https://perma.cc/CC2T-LTEY].

254. Sawyer, *supra* note **Error! Bookmark not defined.**, at 394.

255. Lev-Aretz, *supra* note 227, at 158.

256. Sawyer, *supra* note **Error! Bookmark not defined.**, at 382-83.

257. Ben Depoorter & Robert Kirk Walker, *Copyright False Positives*, 89 NOTRE DAME L. REV. 319, 326 (2013); *see also Endgame Removed from YouTube by False Copyright Claim*, INFOWARS (Feb. 15, 2015), http://www.infowars.com/endgame-removed-from-youtube-by-false-copyright-claim [https://perma.cc/9J73-2YZV].

258. Ernesto, *YouTube Flags Cat Purring as Copyright Infringing Music*, TORRENTFREAK (Feb. 11, 2015),
http://torrentfreak.com/youtube-flags-cat-purring-as-copyright-infringing-music-150211 [https://perma.cc/P3Z7-9PMG].

259. Heald, *supra* note 230, at 37.

nature of copyright enforcement algorithms prevents users from understanding how the algorithms exercise their power; that such algorithms further diminish users' due process by affording them inadequate challenging opportunities; and that their ubiquity and frequent ex ante application curb public oversight by diminishing the process of error correction. In this final Part, we identify the barriers for enhancing accountability in algorithmic copyright enforcement by online intermediaries and discuss various strategies for removing them.

A. *Mapping the Barriers to Algorithmic Accountability*

This Subpart divides the barriers for enhancing accountability in algorithmic copyright enforcement implemented by online intermediaries into three categories: the first category explores technical barriers, namely the non-transparent nature of algorithms, as well as the operation of copyright enforcement algorithms, which often relies on constantly evolving learning machines. The second category covers legal barriers that impede public literacy in relation to algorithmic copyright enforcement exploited by online intermediaries. The third category focuses on practical barriers that effectively negate the purported objectives of the counter notice procedure under the DMCA, resulting in insufficient dispute opportunities.

1. *Technical Barriers*

There are two intertwined technical barriers to adequate public scrutiny of algorithmic copyright enforcement. The first is the non-transparent nature of algorithms, which makes it difficult to review their decision-making processes.[260] Algorithmic copyright enforcement embeds a high degree of automation, even if it may sometimes involve some degree of human analysis.[261] While algorithms can be built to advance specific values and policies,[262] they are ultimately complex

260. This is not to say, however, that algorithmic copyright enforcement merits a higher level of transparency than what manual copyright enforcement demands. Transparency "should be applied to all steps which might compromise rights of individuals and seem arbitrary, be they automated or manual. The level of automation needs not, on its own, merit a higher level of transparency." *See* Zarsky, *supra* note 34, at 1552. Both regimes—the automated one and the human-implemented one—should eventually reach a similar degree of transparency, yet automated regimes, by their nature, inherently challenge this goal.

261. Steve Lohr, *Algorithms Get a Human Hand in Steering Web*, N.Y. TIMES (Mar. 10, 2013), http://www.nytimes.com/2013/03/11/technology/computer-algorithms-rely-increasingly-on-human-helpers.html?pagewanted=all&_r=1& [https://perma.cc/X8GF-37SA] (demonstrating how the resources of intermediaries, such as Google and Twitter, are becoming more human curated).

262. Nissenbaum, *supra* note 15, at 1373; Bruno Latour, *Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts*, *in* SHAPING TECHNLOGY/BUILDING SOCIETY 225, (Wiebe Bijker & John Law eds., 1992). Regarding copyright enforcement specifically, see R. Polk Wagner, *Reconsidering the DMCA*, 42 HOUS. L. REV. 1107 (2005) (suggesting that the total regulatory effect combines both law and technology and that changes in technology affect the law and vice versa).

codes that we (and most program developers[263]) cannot easily deconstruct. The ability to integrate an almost unlimited number of variables into automated decision rules means that algorithms "can successfully apply rules whose complexity would make them collapse under their own weight if humans were forced to apply them."[264] As Bamberger explains, "programming and mathematical idiom can shield layers of embedded assumptions from high-level firm decision makers charged with meaningful oversight and can mask important concerns with a veneer of transparency."[265]

Furthermore, enforcement algorithms used by online intermediaries effectively advance the intermediaries' own interpretation of legal norms.[266] This process of translating legal mandates into code inevitably embodies particular choices as to how the law is interpreted, which may be affected by a variety of extrajudicial considerations, including the conscious and unconscious professional assumptions of program developers, as well as various private business incentives.[267] It may even be affected by automated bias.[268] As Citron stresses, policy distortions can arise when possibly biased code writers, who lack "policy knowledge," translate policy from human language to code.[269] Some disparity between the algorithmic representation of law and the law as it operates in practice is hence unavoidable.[270] Comprehending algorithmic systems of law enforcement is hence a complex task, which demands knowing what cognitive frames of reference, as well as social, political, economic, and legal motivations shaped the choices made by those who designed them.[271]

The second technical barrier to proper accountability in algorithmic copyright enforcement relates to the learning capacities of algorithms. Machine learning enables the identification of trends, relationships and hidden patterns in disparate groups of data.[272] Some algorithms can thus adapt their code and shape

263. Frank Pasquale, *Restoring Transparency to Automated Authority*, 9 J. ON TELECOMM. & HIGH TECH. L. 235, 246 (2011).

264. Grimmelmann, *supra* note 69, at 1734.

265. Bamberger, *supra* note 15, at 727.

266. *Id.* at 675.

267. *Id.* at 675-76.

268. *Id.* at 675 (explaining that automation biases are "decision pathologies that hinder careful review of automated outcomes, especially by those with financial incentives that promote risky behavior. These very phenomena contributed to the failure of risk regulation and risk management to prevent the recent financial meltdown"); Friedman & Nissenbaum, *supra* note 105, at 333.

269. Citron, *supra* note 11, at 1261-62.

270. Harry Surden et al., *Representational Complexity in Law*, 11 INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 193, 193 (2007).

271. *See* Jay P. Kesan & Rajiv C. Shah, *Deconstructing Code*, 6 YALE J.L. & TECH. 277, 283 (2004) ("Science & Technology Studies examines how technology is shaped by societal factors such as politics, institutions, economics, and social structures.").

272. Bernhard Anrig et al., *The Role of Algorithms in Profiling*, *in* PROFILING THE EUROPEAN CITIZEN: CROSS-DISCIPLINARY PERSPECTIVES 65 (Mireille Hildebrandt & Serge Gutwirth, eds., 2008), http://link.springer.com/chapter/10.1007%2F978-1-4020-6914-7_4 [https://perma.cc/JCD7-EYK8].

performance based on experience. For instance, Google and Facebook may run dozens of different versions of an algorithm to assess their relative merits, with no guarantee that the version a user interacts with at one moment in time is the same as five seconds earlier.[273] Indeed, algorithms are often designed to be reactive and mutable to inputs.[274] The EdgeRank algorithm which Facebook uses to determine what posts, and in what particular order, to feed into each user's timeline works in concert with each individual user, ordering posts in accordance with how one interacts with "friends."[275]

These learning capacities make algorithms significantly less predictable than some computer codes that operate by means of on-off rules,[276] further diminishing their already-deficient accountability. While we can easily judge why an automated camera issues a speeding a ticket, determining whether removal, filtering, or blocking algorithms legitimately vindicate the rights of copyright owners is quite challenging.[277] For these algorithms are not merely tools for implementing the goals of the intermediaries employing them; they may practically shape the meaning of those goals themselves. Under Bamberger's interpretation, "they create a Gestalt, or world view, that alters the perceptions of the decision makers they inform."[278]

A different view may suggest that the learning capacities of law enforcement algorithms should be weighed in favor of public oversight, and not against it. Particularly, the elasticity of learning machines,[279] or their ability to reshape and adapt code as circumstances demand,[280] arguably make them more adjustable to mistake correction.[281] This is especially so due to their central locations, which enable quick changes.[282] Nevertheless, after the initial introduction of technology, its flexibility mostly vanishes as it transitions from a device "oriented toward human needs" to "an important component of the formative context."[283] Much

---

273. Rob Kitchin, *Thinking Critically About and Researching Algorithms* 16 (The Programmable City Working Paper No. 5, 2014), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2515786 [https://perma.cc/6YRN-HNZZ]

274. *Id.*

275. Taina Bucher, *Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook*, 14 NEW MEDIA & SOC'Y 1164 (2012).

276. Bamberger, *supra* note 15, at 676.

277. Pasquale, *supra* note 263, at 237.

278. Bamberger, *supra* note 15, at 676.

279. Grimmelmann, *supra* note 69, at 1730-31.

280. Bamberger, *supra* note 15, at 710.

281. *See* Hal R. Varian, *Kaizen, that Continuous Improvement Strategy, Finds Its Ideal Environment,* N.Y. TIMES (Feb. 8, 2007), http://www.nytimes.com/2007/02/08/business/08scene.html [https://perma.cc/5U3Z-4MUZ]; *see also* Sawyer, *supra* note **Error! Bookmark not defined.**, at 384 (creating a distinction between Content ID, which is much more adaptable to change because its central location enables quick changes, and Digital Rights Managements, which are much more difficult to fix without rolling out a second generation of technology).

282. Varian, *supra* note 281.

283. Claudio U. Ciborra, *De Profundis? Deconstructing the Concept of Strategic Alignment*, 9

like legislative acts, technology, too, establishes a set of guidelines for public order that may persist for many generations. Unable to understand and deconstruct the layers of codes and policies embedded in law enforcement algorithms, we expect the general public to largely work its way outside this complex regime, rather than contest it.

2. *Legal Barriers*

Several legal barriers also obstruct the public's ability to hold algorithmic mechanisms of copyright enforcement accountable. First, copyright law may hinder public literacy that could be pursued through independent, reverse engineering research.[284] Theoretically, users can engage in self-help and reverse engineer copyright enforcement algorithms to learn how they exercise their power. This is exactly what Glen Gabe, a digital marketing veteran, tried to do with Google's Pirate algorithm.[285] As part of his attempt to figure out the threshold Google uses when choosing to target a pirated domain by making it appear lower in its search results, Gabe analyzed the data on Google's Transparency Report, including copyright takedown notices, domains being specified in those takedowns and top copyright owners. After analyzing the organic search trends of affected sites, Gabe concluded that the Pirate algorithm probably kicks in when at least half of one's indexed URLs are subject to takedown requests.[286] By analyzing the output of this unknown algorithm, i.e. how many takedown requests were received by websites that were ultimately targeted by the Pirate algorithm, Gabe effectively enhanced users' literacy regarding the metric underlining Google's search engine mechanism, consequently enhancing its accountability.[287]

Notwithstanding the DMCA's important benefit in affording the public self-help measures that are capable of extracting valuable information on algorithmic governance, the DMCA's anti-circumvention provisions[288] may suppress this

---

SCANDINAVIAN J. INFO. SYS. 67, 76 (1997).

284. *See* Maayan Perel & Niva Elkin-Koren, *Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement*, 69 FLA. L. REV. (forthcoming 2017).

285. *See supra* note 183 and accompanying text.

286. *See supra* note 183 and accompanying text.

287. See Diakopoulos, *supra* note 75 (discussing similar strategies of reverse engineering).

288. 17 U.S.C. §§ 1201-1202 (2012). The most pertinent of the DMCA's anti-circumvention provisions read in part:

> (a) Violations Regarding Circumvention of Technological Measures.—
> (1)(A) No person shall circumvent a technological measure that effectively controls access to a work protected under this title.
> . . . .
> (2) No person shall manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof,
> that—
> (A) is primarily designed or produced for the purpose of circumventing .. ;
> (B) has only limited commercially significant purpose or use other than to circumvent . . . ; or
> (C) is marketed by that person or another acting in concert with that person with that person's knowledge for use in circumventing . . .

sort of reverse engineering research. Generally, the anti-circumvention provisions prohibit the use, development, or distribution of technologies which are designed to "circumvent" (e.g., hack, crack, or break) access control systems,[289] with some narrow exceptions.[290] Defenders of strong copyright called for this legal intervention, acknowledging that technological protection measures (TPM) and digital rights management systems (DRM) are not tamper-proof.[291]

Nevertheless, during their first decade of existence, the anti-circumvention provisions were thought to have a very restrictive impact.[292] At that time, the prevalent technological measures were basically encrypted computer codes that were incorporated into fixations of copyrighted material, such as CDs or music files, to prevent illegal copying and public distribution of copyrighted works.[293] Meaning, the anti-circumvention provisions initially applied to technologies that restricted access to specific content, and not to particular distribution channels.[294] Therefore, the basic assumption was that the "anti-circumvention rules do not generally affect the users of copyrighted work."[295] The overwhelming majority of users—namely lay computer users—were presumed to be in the same position after the DMCA's anti-circumvention enactment as before it.[296]

This assumption seems to collapse under algorithmic copyright enforcement by online intermediaries. The application of online technologies that restrict the distribution of allegedly infringing content effectively pushes online intermediaries' negative impact in suppressing the development of innovative circumvention technologies forward.[297] Originally designed to prohibit users

. . . .

(b) Additional Violations.—

(1) No person shall manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof,

that—

(A) is primarily designed or produced for the purpose of circumventing protection . . . ;

(B) has only limited commercially significant purpose or use other than to circumvent protection . . . ; or

(C) is marketed by that person or another acting in concert with that person with that person's knowledge for use in circumventing protection . . .

17 U.S.C. § 1201(a)-(b).

289. *See supra* note 288 and accompanying text.

290. Pamela Samuelson, *Intellectual Property and the Digital Economy: Why the Anti Circumvention Rules Need to be Revised*, 14 BERKELEY TECH. L.J. 519 (1999) (explaining that circumvention is permissible for some limited purposes, such as achieving program-to-program interoperability or engaging in encryption research and computer security testing).

291. *Id.*

292. Stephen M. Kramarsky, *Copyright Enforcement in the Internet Age: The Law and Technology of Digital Rights Management*, 11 DEPAUL-LCA J. ART & ENT. L. & POL'Y 1, 10 (2001) ("[T]he new anti-circumvention laws prevent sophisticated users from bypassing the technology.").

293. *Id.*

294. *See supra* note **Error! Bookmark not defined.** and accompanying text.

295. Wagner, *supra* note 262, at 1124.

296. *Id.* at 1125.

297. *Id.* (The DMCA "simultaneously suppresses and encourages technology. That is, on the one hand it encourages the deployment of 'access control' technologies on copyrighted

from developing ways to deconstruct a single-content access control, the anti-circumvention provisions could now be interpreted to discourage users from challenging technological measures that monitor popular channels of content distribution.[298] In contrast to DRM technologies, which are bi-directional—defining the direct relationship between a copyright owner and the holder of a specific copy of the copyrighted content—mechanisms of algorithmic copyright enforcement have far-reaching implications over public discourse.[299] Having the ability to circumvent mechanisms of algorithmic copyright enforcement thus reaches beyond the narrow interests of lawful owners of specific copies of copyrighted content and curious technologists because it enables users to contest what some scholars have characterized as the "privication" of information that would have otherwise been public.[300]

Nevertheless, when technologists have attempted to reverse engineer filtering programs used in public schools, libraries, and similar institutions to protect minors from exposure to indecent or otherwise harmful material posted online, the developers of the filtering programs have been successful in arguing that reverse engineering the encryption to analyze the list of blocked sites violated the DMCA's anti-circumvention rules.[301] In *Edelman v. N212*, the court refused to declare that reverse engineering the software was lawful, even though that information was critically important to a back-then public policy debate over whether legislatures should mandate use of filtering software in public schools and libraries.[302] By virtue of comparison, there is a genuine risk that the DMCA's anti-circumvention provisions would be invoked to ban the deconstruction of online copyright enforcement technologies employed by online intermediaries, consequently refuting its important accountability-enhancing potential.

Another legal obstacle to enhancing accountability in algorithmic copyright enforcement is created by trade secrecy law. In a famous legal battle between Viacom and YouTube, a judge refused to force YouTube to provide Viacom with the computer source code which controls both the YouTube.com search function and Google's internet search tool "Google.com."[303] The court explained that "the search code is the product of over a thousand person-years of work" and that

works . . . On the other hand, it prohibits the use, development, or distribution of 'circumvention' technologies").

298. *See supra* Part II.C.2.

299. *See supra* Part II.C.2.

300. Privication describes the possibility of private publication, where content providers distribute content on a large-scale but at the same time retain control over access. *See* Jonathan Zittrain, *What the Publisher Can Teach the Patient: Intellectual Property and Privacy in an Era of Trusted Privication*, 52 STAN. L. REV. 1201, 1203 (2000).

301. Press Release, ACLU, In Legal First, ACLU Sues over New Copyright Law: Says Blocking Program Lists Should Be Revealed (July 25, 2002), https://www.aclu.org/news/legal-first-aclu-sues-over-new-copyright-law-says-blocking-program-lists-should-be-revealed [https://perma.cc/4GSR-LK8Q].

302. 263 F. Supp. 2d 137, 138 (D. Mass. 2003).

303. Viacom Int'l, Inc. v. YouTube, Inc., No. 1:07-cv-02103-LLS, 2008 U.S. Dist. LEXIS 50614 (S.D.N.Y. Jul. 2, 2008).

"there is no dispute that its secrecy is of enormous commercial value. Someone with access to it could readily perceive its basic design principles, and cause catastrophic competitive harm to Google by sharing them with others who might create their own programs without making the same investment."[304] Although the production and examination of the source code was, according to Viacom, the only way to find out whether YouTube's search algorithm effectively increases the rank or visibility of allegedly infringing material over non-infringing material,[305] the court denied Viacom's motion to compel YouTube to produce its source code and consequently lose its trade secret.

As part of the growing recognition of the economic benefits streaming from protecting intangible assets, it would be reasonable to expect firms to claim an increasingly broad range of non-public information as trade secrets.[306] Trade secrecy law encourages intermediaries to keep their proprietary algorithms secret because once they are revealed, they lose trade secret protection as a matter of law.[307] As demonstrated by *Viacom Int'l, Inc. v. YouTube, Inc.*, protecting the secrecy of proprietary codes is necessary not only to prevent intentional infringement, but also to prevent competitors from free riding on an intermediary's economic investment in developing its codes.

Nonetheless, secrecy has major negative consequences for society because it obstructs accountability.[308] In the digital realm, which largely enhances consumers' capacity to actively engage in creative processes,[309] trade secrecy can further undermine access to information.[310] Accordingly, keeping copyright enforcement algorithms secret is particularly problematic because "they are *de facto* sovereigns over important swaths of social life."[311] Their invisible hand effectively controls what content is available and further determines its ranking and accessibility.[312] Yet so long as copyright enforcement algorithms are potentially trade secrets, investigating their misconduct may never reach a conclusive end,[313] notwithstanding their bothering impact on public discourse.

---

304. *Id.* at *8.

305. *Id.* at *10.

306. *See, e.g.*, Robert P. Merges, *One Hundred Years of Solicitude: Intellectual Property Law, 1900-2000*, 88 CAL. L. REV. 2187, 2233-40 (2000) (discussing expansions in intellectual property protection during the twentieth century).

307. Pasquale, *supra* note 263, at 237.

308. David S. Levine, *Secrecy and Unaccountability: Trade Secrets in Our Public Infrastructure*, 59 FLA. L. REV. 135, 170-77 (2007); Mary L. Lyndon, *Information Economics and Chemical Toxicity: Designing Laws to Produce and Use Data*, 87 MICH. L. REV. 1795, 1855-56 (1989).

309. Niva Elkin Koren, *Making Room for Consumers Under the DMCA*, 22 BERKLEY TECH. L.J. 1119, 1120 (2007).

310. Pasquale, *supra* note 263, at 245-46.

311. DAVID STARK, THE SENSE OF DISSONANCE: ACCOUNTS OF WORTH IN ECONOMIC LIFE 1 (2011).

312. *See supra* Part II.C.2.

313. Pasquale, *supra* note 263, at 245. Google's secrecy about its website-ranking algorithm has provoked unsuccessful investigations in Europe and the U.S., however. *See* Editorial, *The Google Algorithm*, N.Y. TIMES (July 15, 2010), http://www.nytimes.com/2010/07/15/opinion/15thu3.html (last visited May 8, 2016) ("[T]he

Finally, beyond the DMCA's anti-circumvention provisions and trade secrecy law, general "computer intrusion" and "anti-hacking" laws form another legal barrier to enhancing accountability in algorithmic copyright enforcement.[314] Indeed, state and federal statutes further protect computer network owners from hacking and unauthorized intrusions, including the Computer Fraud and Abuse Act (CFAA),[315] the Wiretap Act,[316] the Electronic Communications Privacy Act (ECPA),[317] and a variety of state computer intrusion statutes. These laws criminalize different conducts of computer hacking, generally subject to a financial damage threshold requiring that a plaintiff prove that the intrusion caused some harm.[318] This criminalization includes the penetration of computer systems to gain knowledge about their operation,[319] which could further restrain users from investigating copyright enforcement algorithms.

3. *Practical Barriers*

A third barrier to enhancing accountability in algorithmic copyright enforcement relates to the practical ineffectiveness of the counter notification procedure under the DMCA. Under § 512(g)(2)(B), hosting services are required to promptly forward any counter notices from alleged infringers back to the original complainant.[320] If after ten to fourteen days, the complainant does not notify the OSP that she had filed a lawsuit, then the OSP must reinstate the contested material. Although this is a simple self-help procedure allowing users to easily contest improper takedown removals, alleged infringers hardly employ it in practice.[321] Bruce Boyden has recently found that the Motion Picture Association of America (MPAA) sent twenty-five million takedown notices to search engines

---

potential impact of Google's algorithm on the Internet economy is such that it is worth exploring ways to ensure that the editorial policy guiding Google's tweaks is solely intended to improve the quality of the results and not to help Google's other businesses."); Richard Waters, *Unrest over Google's Secret Formula*, FIN. TIMES (July 12, 2010), http://www.ft.com/cms/s/0/1a5596c2-8d0f-11df-bad7-00144feab49a.html (last visited May 8, 2016).

314. See Elec. Frontier Found., The CFAA: Blocking Competition and Stifling Innovation, https://www.eff.org/files/filenode/cfaa-stifling-innovation.pdf [https://perma.cc/97R3-PCT3] for examples of how anti hacking laws prohibit reverse engineering, yet with emphasis on innovation, not accountability.

315. 18 U.S.C. § 1030 (2012).

316. 18 U.S.C. §§ 2510-2522 (2012).

317. 18 U.S.C. §§ 2701; 3121 (2012).

318. *Id.*; *see also* 18 U.S.C. § 1030(a)(4) (2012), which addresses the access and fraudulent use of a protected computer and is triggered if the value of the use obtained exceeds $5,000.

319. Eric J. Sinrod & William P. Reilly, *Cyber-Crimes: A Practical Approach to the Application of Federal Computer Crime Laws*, 16 SANTA CLARA HIGH TECH. L.J. 177, 181 (2000).

320. A counter-notification must include the following: (A) a physical or electronic signature; (B) identification of the material removed and its former location; (C) statement under penalty of perjury that the user has a good faith belief the material was mistakenly removed; (D) the user's name, address, and phone number; and (E) consent to the jurisdiction of a federal district court. *See* 17 U.S.C § 512(g)(3) (2012).

321. Seng, *supra* note 7, at 426.

and cyber-lockers during a six-month period in 2013, while receiving only eight counter notices challenging its requests.[322] According to Twitter's transparency report, Twitter received 2,453 DMCA takedown notices in respect of the Twitter platform during the month of January 2015, while receiving only two counter notices in respect of all Twitter services (i.e. including Vine) in return.[323]

Different explanations for this discrepancy in takedown notices and counter notices include users' disinclination to disclose their identities in a counter notice and thereby submit themselves to U.S. jurisdiction;[324] the intimidating language of takedown notices, or simply users' "ignorance or unawareness of the possible responses" to a notice of takedown.[325] For whatever reason, the bottom line is that even when a counter notice procedure is available,[326] it is rarely employed, rendering it practically useless in enhancing the accountability of algorithmic copyright enforcement mechanisms.

B. *Accountability Enhancing Strategies*

As the previous Subpart demonstrated, when aiming to enhance the accountability of algorithmic copyright enforcement mechanisms, different technical, legal, and practical obstacles must be taken under consideration. In the final Subpart of this Article, we will examine several strategies that may address some of the barriers discussed above. Generally, these strategies can be pursued by four distinct players: (1) individual users; (2) privately organized groups or "watchdogs"; (3) online intermediaries; and (4) regulators.

1. *Encouraging Public Participation*

When thinking of enhancing the accountability of algorithmic copyright enforcement, it is natural to put the initial focus on users and ask what they can do to watch copyright enforcement algorithms. The answer splits into two: first, innocent target users—namely, users whose online content has been improperly restricted—should always avail themselves of the counter notice procedure. While the counter notice procedure is insufficient by itself to generate adequate accountability,[327] it is still better than having no challenging opportunities whatsoever. Perhaps, if it were taken more seriously, it could discourage

322. Bruce Boyden, The Failure of the DMCA Notice and Takedown System: A Twentieth Century Solution to a Twenty-First Century Problem (Dec. 2013) http://cpip.gmu.edu/wp-content/uploads/2013/08/Bruce-Boyden-The-Failure-of-the-DMCA-Notice-and-Takedown-System1.pdf [https://perma.cc/33TZ-T3U4].

323. *Copyright Notices: Jan. 1-Jun. 30*, TWITTER TRANSPARENCY REPORT (2015), https://transparency.twitter.com/copyright-notices/2015/jan-jun [https://perma.cc/53SZ-SV7L].

324. Seng, *supra* note 7, at 430.

325. *Id.*

326. As we explained, the counter notice procedure is not always available. *See supra* notes 191-192 and accompanying text.

327. *See supra* Part III.A.

repetitive filers of takedown notices from harassing innocent users. Realizing that target users fight back to force intermediaries to reinstate their non-infringing or otherwise fair use content, copyright holders and their representatives might be dissuaded from expending their time and resources on filing bogus copyright infringement claims.

Second, researchers and the public in general should be encouraged to engage in reverse engineering research to reveal the essence of algorithmic enforcement systems.[328] In this regard, the DMCA's anti-circumvention provisions should be excused under 17 U.S.C § 1201(g). This exception authorizes "encryption research"—"activities necessary to identify and analyze flaws and vulnerabilities of encryption technologies applied to copyrighted works, if these activities are conducted to advance the state of knowledge in the field of encryption technology or to assist in the development of encryption products."[329] Arguably, the act of reverse engineering the code underlying copyright enforcement algorithms may be considered as "encryption research" aimed to investigate the code in order to advance the state of the knowledge in the field of algorithmic copyright enforcement.

Fortunately, it appears that reverse engineering is increasingly receiving positive treatment. For instance, when a Princeton graduate was threatened with a DMCA lawsuit after publishing a report documenting weaknesses in a CD copy-protection technology developed by SunnComm, public outcry and negative press attention caused the plaintiff to retreat from its threats.[330] Similarly, Hewlett-Packard resorted to DMCA threats when researchers published a security flaw in HP's Tru64 UNIX operating system, but after widespread press attention, HP ultimately withdrew the DMCA threat.[331]

Furthermore, when the act of reverse engineering does not result in damaging a computer system, general anti-hacking could also be excused.[332] Researchers and scholars who engage in reverse engineering of copyright enforcement algorithms should therefore avoid causing any damage to the underlying code by limiting their hacking efforts to challenging a given software, without attempting to tamper with its operation. Finally, trade secrecy law regards reverse engineering as a lawful way to acquire know-how that the product's manufacturer may claim as a trade secret.[333] Accordingly, the protection of trade secrecy law could vanish where the threshold of copyright enforcement

328. Google's secrecy about its website-ranking algorithm has provoked investigations in Europe. *See* Waters, *supra* note 313, at 22 ("Prompted by three complaints, the European Commission this year began an informal investigation, the first time that regulators have pried into the inner workings of the technology that lies at the heart of Google.").

329. 17 U.S.C. § 1201(g)(1)(A) (2012).

330. *DMCA: Ten Years of Unintended Consequences*, ELEC. FRONTIER FOUND. (Oct. 27, 2008), https://www.eff.org/wp/unintended-consequences-ten-years-under-the-dmca [https://perma.cc/57VJ-FKXZ].

331. *Id.*

332. *See* 18 U.S.C. § 1030(a)(5) (2012).

333. *See, e.g.*, Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 484-92 (1974).

algorithms is revealed as a result of independent reverse engineering efforts.[334]

2. *Watchdogs*

Another important role in enhancing accountability in algorithmic copyright enforcement includes that of the private initiatives that are committed to protecting online free speech and free flow of information. Private initiatives may boost the accountability of copyright enforcement algorithms by retrieving information about improper speech restrictions, commenting on it, and distributing it to the public.[335] A prominent example is the Electronic Frontier Foundation (EFF),[336] which is a leading nonprofit organization defending civil liberties in the digital world. Through "impact litigation, policy analysis, grassroots activism, and technology development," the EFF promotes "user privacy, free expression and innovation."[337] Armed with technological experts, activists, and attorneys, the EFF increases the awareness of policymakers, the press, and the public to online free speech violations by posting comprehensive analysis and educational guides, and engaging in transparency enhancing initiatives.[338]

For instance, its Takedown Hall of Shame[339] project collects and displays bogus copyright (and trademark) complaints that have threatened all kinds of creative expression on the Internet. By highlighting interesting stories of improper DMCA takedown requests, the EFF's Takedown Hall of Shame can raise public awareness and generate public pressure on intermediaries. In that sense, it has an ex ante shaming effect in discouraging unjustified takedown requests, as well as ex post impacts in fostering public oversight and contest.

Another important project of the EFF relates to promoting quantitative Fair Use Principles.[340] These principles recommend that content only be blocked if both the audio and video tracks match the same work and ninety percent or more of the uploaded content comes from a single work.[341] Moreover, they also advocate compliance with the DMCA's N&TD procedures when removing

334. UNIF. TRADE SECRETS ACT § 1 cmt. 2, 14 U.L.A. 437, 438 (1990).

335. Zarsky, *supra* note 34, at 1535.

336. ELEC. FRONTIER FOUND., https://www.eff.org [https://perma.cc/LVP2-YH46] (last visited May 8, 2016).

337. *About EFF*, ELEC. FRONTIER FOUND., https://www.eff.org/about [https://perma.cc/67UF-VL7S] (last visited May 8, 2016).

338. *Id.*

339. *Takedown Hall of Shame*, ELEC. FRONTIER FOUND., https://www.eff.org/takedowns [https://perma.cc/M4E7-7X8D] (last visited May 8, 2016).

340. *Fair Use Principles for User Generated Video Content*, ELEC. FRONTIER FOUND., http://www.eff.org/issues/ip-and-free-speech/fair-use-principles-usergen [https://perma.cc/U3AQ-SKTA] (last visited May 8, 2016); *See* Julian Sanchez, *EFF Seeks Mashup Makers to Fight YouTube Filtering*, ARS TECHNICA (Feb. 3, 2009, 2:44 PM), http://arstechnica.com/telecom/news/2009/02/eff-seeksmashup-makers-to-fight-youtube-filtering.ars [https://perma.cc/6TAE-EL54].

341. *Fair Use Principles for User Generated Video Content*, *supra* note 340.

content, including enabling counter notices.[342] Finally, the Fair Use Principles encourage UGC sites to enable dialogue between content owners and users to resolve fair use takedowns.[343] Unfortunately, thus far, these well-intended recommendations have not been widely implemented. As we showed, intermediaries do not publish how they determine fair use; content is occasionally blocked ex ante, in direct opposition to the DMCA; and fair use takedowns are hardly resolved through online dispute resolution systems.[344]

Another valuable private initiative that helps holding online copyright enforcement mechanisms accountable is Lumen (formerly the Chilling Effects).[345] Founded by the Berkman Center for Internet and Society at Harvard University, Lumen offers an invaluable clearinghouse for researchers and the public in general. Basically, Lumen is "an independent 3rd party research project studying cease and desist letters concerning online content," especially "requests to remove content from online."[346] Its goals are:

> to educate the public, to facilitate research about the different kinds of complaints and requests for removal—both legitimate and questionable—that are being sent to Internet publishers and service providers, and to provide as much transparency as possible about the 'ecology' of such notices, in terms of who is sending them and why [they are being sent], and to what effect.[347]

Yet, while Lumen receives a growing number of removal notices—especially from Google and Twitter, who made public commitments to publish their notices in Lumen, but also from private recipients—it does not publish all notices of removal received by all existing intermediaries.[348] Indeed, even Google states that a copy of each legal notice "*may* be sent to the Lumen project for publication and annotation,"[349] allowing that some removal notices may remain unpublished. Hence, further supporting and encouraging the participation of online intermediaries in this project is an important step in ensuring researchers have access to data that can teach them about the practices of online intermediaries engaging in algorithmic copyright enforcement, which in turn promotes public literacy and accountability.

---

342. *Id.*

343. *Id.*

344. *See supra* Part III.B.

345. LUMEN, https://lumendatabase.org [https://perma.cc/K6UD-FRN5] (last visited May 8, 2016).

346. *See About*, LUMEN, https://lumendatabase.org/pages/about [https://perma.cc/6H37-S8J3] (last visited May 8, 2016).

347. *Id.*

348. *Copyright Policy*, TWITTER HELP CTR., http://support.twitter.com/articles/15795 [https://perma.cc/BRQ8-YZLB] (last visited May 8, 2016); *Legal Removal Requests*, GOOGLE SUPPORT, https://support.google.com/legal/answer/3110420?hl=en&ref_topic=4556931 [https://perma.cc/RYE9-36JX] (last visited May 8, 2016); Seng, *supra* note 7, at 379.

349. *Legal Removal Requests*, *supra* note 348 (emphasis added).

3. *Intermediaries*

Intermediaries themselves may also assist in enhancing the accountability of algorithmic copyright enforcement because they stand in the best position to increase the transparency of their own, private mechanisms. Arguably, if intermediaries voluntarily disclosed information about their fair use policies and their quantitative thresholds, they could foster public scrutiny over their algorithms' automatic determinations. If intermediaries published all takedown notices they receive; they could also assist in generating public literacy by contributing to the clearinghouse of raw data available for researchers and policy makers. Google[350] and a few other intermediaries, including Mapbox, Medium, Reddit, Twitter, Wikimedia and Wordpress,[351] already disclose such information about DMCA takedown notices they receive in their transparency reports.

Despite the advantages recited above, it is important to stress the pitfalls of such voluntary transparency reports.[352] First, voluntary sharing of information may allow the release of partial information that might be biased and misleading. Second, establishing public oversight exclusively on intermediates' self-reporting may entrust online intermediaries with public functions, thus further strengthening their power in the public sphere. Some extent of regulatory intervention may therefore be required to ensure accurate reporting, as we suggest below.

4. *Regulators*

A fourth player in enhancing the accountability of algorithmic copyright enforcement is the regulator. Regulation might be necessary to impose some level of mandatory disclosure on online intermediaries who engage in copyright enforcement, and to set minimal standards for such disclosure duties. Specifically, given the un-scalability of the current, top-down DMCA regulation, which fails to keep pace with the robustness of algorithmic copyright enforcement, we advocate a collaborative-dynamic regulation that considers the experiences of

350. Google, *supra* note 207.

351. *See Medium's Transparency Report*, MEDIUM (2014), https://medium.com/transparency-report/mediums-transparency-report-438fe06936ff [https://perma.cc/7EXT-Y62D] (last visited May 8, 2016); Reddit, Reddit Transparency Report: Requests for User Information and for Removal of Content (Jan. 29, 2015) https://www.redditstatic.com/transparency/2014.pdf [https://perma.cc/DUA6-7P3W] (last visited May 8, 2016); *Transparency Report*, MAPBOX, https://www.mapbox.com/transparency-report [https://perma.cc/88BT-C3QK] (last visited May 8, 2016); *Transparency Report*, TWITTER, https://transparency.twitter.com [https://perma.cc/WAF3-EZ85] (last visited May 8, 2016); *Transparency Report*, WIKIMEDIA FOUND., https://transparency.wikimedia.org [https://perma.cc/SS6R-KWWR] (last visited May 8, 2016); *Transparency Report*, WORDPRESS, http://transparency.automattic.com [https://perma.cc/Q3VC-F8EJ] (last visited May 8, 2016).

352. Perel & Elkin-Koren, *supra* note 284 (explaining that disclosures cannot practically facilitate adequate accountability in a world of robust algorithmic enforcement and suggesting using the methodology of black box tinkering for extracting valuable information about the practices of algorithmic mechanisms of enforcement).

specific intermediaries in order to continuously update the standards it sets.[353] Indeed, it has been recognized that a system that produces accountability through bottom-up efforts "can curb discretion, promote consistency, allow for monitoring, and create incentives for high-quality performance."[354]

Ideally, enhancing intermediaries' accountability in enforcing copyrights through regulation should rely on three components previously described in Bamberger's work on automated risk-management technologies[355]: (1) increasing transparency as to the decisions intermediaries make in structuring their copyright enforcement systems; (2) investing in the competence of the administrative regulator itself, both in terms of technical expertise and in terms of computing capacity; and (3) encouraging cooperation between the regulated intermediaries and the regulator in the ongoing development of an effective regime of algorithmic copyright enforcement.[356] We explain each of these elements henceforth.

First, regulation may enhance the accountability of algorithmic copyright enforcement by setting standards of disclosure, such as requiring intermediaries to disclose the criteria their enforcing algorithms consider when determining copyright infringement, including their quantitative thresholds (i.e., what percentage of the copyrighted work must be used to cause content restriction) and fair use policies.[357] Perhaps it would even be advisable to standardize measures of copyright infringement and fair use to create uniformity and consistency among the different intermediaries, which may further assist in preventing users from attempting to reload their removed content on alternative platforms.[358]

---

353. Michael C. Dorf, *The Domain of Reflexive Law*, 103 COLUM. L. REV. 384 (2003) (reviewing JEAN L. COHEN, REGULATING INTIMACY: A NEW LEGAL PARADIGM (2002)). Dorf essentially proposes a model of reflective law in which insights drawn from experience at the relatively local level are continually refined and transmitted to the standard-setter, which uses these insights continually to update the standards all must meet. *Id.* We propose to encourage similar cross-fertilization between the regulator who sets the standards of disclosure and the online intermediaries that must apply them in order to develop standards that on the one hand, allow for adequate disclosure, and on the other hand, do not overly burden the intermediaries.

354. Rabinovitch-Einy, *supra* note 35, at 269 (using the term "structural accountability" to describe "a system that produces accountability through bottom-up efforts").

355. Bamberger, *supra* note 15.

356. *Id.* at 735-38.

357. In the EU, for instance, specific legal rules in the Data Protection Directive provide users with a supplemental right to receive information about the underlying logic of automated processing of their personal data. *See* Council Directive 95/46/EC, art. 12(a), 1995 O.J. (L 281) 42 (EC); *see also* Douwe Korff, *Data Protection Laws in the EU: The Difficulties in Meeting the Challenges Posed by Global Social and Technical Developments* (Centre for Public Reform, Working Paper No. 2, 2010), http://ec.europa.eu/justice/policies/privacy/docs/studies/new_privacy_challenges/final_report_working_paper_2_en.pdf [https://perma.cc/ADG5-XU96].

358. One example relates to the EFF's Fair Use Principles discussed earlier. Although these specific recommendations are yet to be implemented in any meaningful way, this sort of standardization accompanied with regulated disclosure would generally allow target users to better understand why their content uploads had been taken down or targeted for take down, while further facilitating public scrutiny over the enforcing algorithms' determinations. *See*

Furthermore, regulation can also impose reporting obligations on online intermediaries, requiring them to publish all takedown notices they receive in order to allow interested parties to follow and review their conduct. Note, however, that intermediaries should not be forced to disclose their actual source code—not only because it is extremely unlikely that lay users would be able to read and understand it, but especially because the source code may be legitimately protected under trade secrecy law.[359] As a general rule, when deciding on a specific mandatory disclosure policy, it is important not to overburden intermediaries with requirements. Too much transparency may cause intermediaries to make conservative, inefficient, and unfair decisions, fearing negative criticism.[360] Indeed, an unbalanced mandate of compulsory disclosure may result in tilting the algorithm's default postulation exceedingly towards tolerating copyright infringement.

Second, increasing the transparency of copyright enforcing algorithms must be supplemented with increasing the competence of the regulator[361] (for instance, a trusted advisory committee within the Federal Trade Commission). The immense volume of takedown notices, combined with the fact that these notices are processed by machines, calls for the hiring of enough staff with sufficient technical expertise to propose policy modifications that can be tailored to algorithms. In other words, a suitable regulator must be able to appreciate which policies algorithms can achieve, and which they objectively cannot (because they are too discretionary, for instance). This is especially so if the regulator wishes to standardize copyright infringement and fair use policies and translate them into machine-readable code.

Third, regulators should engage in "robust, albeit collaborative, participation in the ongoing development" of accountable mechanisms of algorithmic copyright enforcement.[362] Indeed, "[c]ollaboration seeks to facilitate a variety of arrangements that might capitalize upon the know-how and abilities of nongovernmental groups in ways that reconfigure those groups and their relationships, while also providing adequate accountability."[363] For instance, regulators can collect data from intermediaries regarding which algorithms work best, and how to deal with gray-area cases. By collaborating with intermediaries in developing regulations and acknowledging the practical limitations of computer codes as enforcement mechanisms, regulators might profoundly change the enforcing algorithms' default assumptions and reduce the likelihood of false positives.

---

supra notes 337-343 and accompanying text.

359. *See supra* notes 303-313 and accompanying text.

360. Zarsky, *supra* note 34, at 1537.

361. Bamberger, *supra* note 15, at 734.

362. *Id.* at 735.

363. Jody Freeman, *Collaborative Governance in the Administrative State*, 45 UCLA L. REV. 1, 31 (1997) (emphasis removed).

V. CONCLUSION

This Article engaged in a critical discussion about algorithmic governance and how it intersects with conventional proxies of accountability. Conceiving algorithms as independent players in automated enforcement regimes that employ non-transparent learning capacities demands rethinking how to accumulate the important virtues of accountability. In this Article, we focused on algorithmic copyright enforcement by online intermediaries to learn about deficiencies in algorithmic accountability. We identified the various causes that prevent adequate public scrutiny and explored current and possible strategies for enhancing accountability in algorithmic copyright enforcement.

The automation of online copyright enforcement—algorithmic applications of the statutory Notice and Takedown regime, and especially algorithmic systems that are capable of blocking allegedly infringing content ex ante—raises serious challenges to the basic notions standing at the heart of accountable enforcement regimes: transparency, due process and public oversight. Private, profit-maximizing mega-players' use of opaque codes to implement discretional legal doctrines, such as fair use and copyright infringement, can have worrying impacts on freedom of speech and the rule of law.

However efficient an invisible hand can be in coordinating online content, it may occasionally be arbitrary and even biased. To secure the free flow of information and protect online users' right to create and liberally enjoy the fruits of their own creations, we must enable adequate checks on algorithmic copyright enforcement employed by online intermediaries. Especially because online copyright enforcement affects fundamental rights, it is vital to allow affected individuals to understand how mechanisms of algorithmic copyright enforcement exercise their power, their decision-making criteria, and how their decisions may be challenged. Otherwise, individuals could be deprived of their right to choose the content they upload and the online platforms they use.

Current practices of algorithmic copyright enforcement, however, do not seem to pursue these objectives successfully. As we have demonstrated, affected individuals lack sufficient knowledge as to the specific thresholds that trigger algorithmic mechanisms of online copyright enforcement. Oftentimes, no procedural safeguards are available to allow affected individuals to challenge algorithmic determinations, and even when challenging opportunities do exist, they are frequently inefficient.

The public as a whole is also largely incapable of efficiently monitoring algorithmic copyright enforcement by online intermediaries. Because ex post content removals are so pervasive, it is impractical to expect the public to promptly review all removals and generate meaningful public pressure to cause the reposting of improperly removed content. Moreover, a large amount of content is being automatically restricted by ex ante filters, such as YouTube's Content ID, making it absolutely impossible for general members of public to know about content restrictions and subsequently contest them.

Enhancing the accountability of algorithmic copyright enforcement faces

several barriers. In addition to the inherent non-transparency of algorithms and their constantly improving learning capacities, there are several legal barriers—particularly the DMCA's anti-circumvention provisions, trade secrecy law and general anti-hacking laws—that may obstruct the public's ability to deconstruct the code underlining algorithmic enforcement mechanisms and hold those mechanisms accountable. As a result, it is imperative to hone current accountability enhancing strategies, such as private watchdogs that police and report improper content restrictions and private disclosure initiatives, and begin thinking of more structured forms of accountability—either by embedding internal scrutiny-enhancing mechanisms in the design of enforcement algorithms, or by subjecting online intermediaries to some degree of external regulatory inspection—to improve the accountability of algorithmic copyright enforcement.