NISKANEN CENTER

Research Paper

ACES in the Hole? Automated Copyright Enforcement Systems and the Future of Copyright Law

Regina Zernay
Adjunct Fellow
The Niskanen Center

Ryan Hagemann
Director of Technology Policy
The Niskanen Center

June 6, 2017

Executive Summary

In the last few decades, the arts have seen an explosion of creative content thanks to new tools for digital content creation and distribution. Coupled with this tremendous output of material is a corresponding increase in both potential and actual copyright infringement. Yet despite the significantly increased risk of infringing activity, copyright laws have not been modified at a rate to match.

As part of the effort to address this increased risk, the Digital Millennium Copyright Act ("DMCA") was enacted. The DMCA aims to serve the dual purpose of "promoting electronic commerce and the distribution of digital works by providing copyright owners with legal tools to prevent widespread piracy" while still "maintaining the integrity of the statutory limitations on the exclusive rights of copyright owners." The DMCA has done this by leaving the policing of

online infringement up to the two groups most affected by online infringement—copyright owners and the online service providers that host user-generated content.

Each group's ability to police infringement, however, has been drastically impacted by an inability to filter the enormous amount of material constantly uploaded online. In response to this dilemma, an interesting private-sector solution has emerged—the use of automated systems, or "content recognition systems" ("CRSs"), by private parties to identify and police infringement.

Two primary types of automated CRSs are currently in use. The first are automated systems specifically used for identifying potentially infringing works and remedying disputations through non-legal channels. These types of systems offer copyright owners several automatic response options when infringement is detected: block the infringing content, permit it to remain available on the host's website, or monetize a share of the revenue. The second type of systems utilize automated CRSs for the specific purpose of enforcing copyright protections. Examples include content recognition software used by universities to scan student papers for plagiarism.

To protect themselves, groups using content recognition software may set the software's parameters at a very high level of responsiveness, sometimes going beyond the measures required by the DMCA. The effect of these modifications is that CRSs sometimes flag material that is not actually infringing under the current legal standard.

Thus, the current copyright enforcement system appears to be a combination of statutes and precedent more suited to a pre-Internet world, and a self-policing system that may improperly apply the law. This paper discusses the legal implications of modifying existing private sector CRSs to create a government-adopted automated copyright enforcement system that more accurately reflects the law.

Such an automated system may help resolve two of the biggest shortcomings in the current copyright regime. First, it provides a more appropriate system of evaluation, as it addresses a technology-based problem with a technology-based solution. Second, it reshapes the existing private sector solution by closely applying the legal standard of infringement, providing a more accurate evaluation of what is (and is not) infringing. In short, it bridges the gap between the systems of creation and delivery, and the system of evaluation.

Notwithstanding the potential benefits of such a system, this paper also acknowledges that the technical complexities in creating such a system may make its implementation difficult. It also recognizes that, on net, the DMCA has been a benefit to both content creators and rightsholders, allowing for an explosive outgrowth of new content to dominate a newly interconnected world.

Ultimately, this paper concludes that an automated copyright enforcement system may eventually be a viable step in the evolution of copyright law. However, the creation of such a system may rest in the distant future, largely due to the technical limitations inherent in

accurately detecting infringing content. Given the current ecosystem, we argue that promoting CRSs and private sector, voluntary solutions—all within the framework of the DMCA—is a far better alternative for ensuring the continued progress of the useful arts.

Introduction

Over the last few decades, technology has developed at an astounding rate. New technological innovations have enhanced nearly all fields, including those involved with the generation and dissemination of copyright-eligible content. The arts have seen an explosion of tools for digital content creation, digitization of existing works, and digital delivery of content, encompassing everything from the written word to 3D art and film.

Copyright laws, however, have not been modified at a rate that matches the rapid influx of digital content.[1] While existing copyright laws may have been reasonably effective before the technological revolution, the law in its current state has not changed quickly enough to meet the increased risk of copyright infringement in the digital world.[2]

In an effort to address this issue, perhaps the most significant addition to copyright law was the Digital Millennium Copyright Act ("DMCA").[3] The DMCA has attempted to serve the dual purpose of "promoting electronic commerce and the distribution of digital works by providing copyright owners with legal tools to prevent widespread piracy" while still "maintaining the integrity of the statutory limitations on the exclusive rights of copyright owners."[4] One way the DMCA achieves these purposes is by offering online content hosts, or "online service providers"[5] ("OSPs"), safe harbor protections from secondary liability for copyright infringement when their users upload infringing content. To receive this protection, OSPs must follow a set of specific steps, including the "expeditious" removal of infringing content upon notification.[6]

Removing infringing content soon after discovering its existence is not as simple as it sounds. The difficulty stems from the sheer volume of user-generated material hosted by online sites like YouTube, Facebook, Vimeo, and others. For example, in July 2015, YouTube CEO Susan

---

[1] *See, e.g.*, Navin Katyal, *The Unauthorized Dissemination of Celebrity Images on the Internet . . . in the Flesh*, 2 TUL. J. TECH. & INTELL. PROP. 1 (2000) ("[T]he Internet poses a new threat for copyright holders, because technology seems to have outpaced current copyright laws.").

[2] *See, e.g.*, Vito A. Costanzo, *Building an IP Litigation and Content Protection Strategy for Clients with Internet-Related Copyright and Trademark Issues*, *in* MANAGING INTELLECTUAL PROPERTY ISSUES IN CYBERSPACE: LEADING LAWYERS ON DEVELOPING AN EFFECTIVE INTERNET IP STRATEGY (2012) ("The Internet not only makes it much easier to access, duplicate, and distribute copyrighted material, it also makes it easier to sue for copyright infringement.").

[3] *Executive Summary: Digital Millennium Copyright Act, Section 104 Report*, U.S. COPYRIGHT OFFICE, https://www.copyright.gov/reports/studies/dmca/dmca_executive.html (last visited Jan. 14, 2017) ("The Digital Millennium Copyright Act of 1998 (DMCA) was the foundation of an effort by Congress to implement United States treaty obligations and to move the nation's copyright law into the digital age.").

[4] *Id.*

[5] The DMCA defines a "service provider" as "an entity offering the transmission, routing, or providing of connections for digital online communications, between or among points specified by a user, of material of the user's choosing, without modification to the content of the material as sent or received." 17 U.S.C. 512(k) (2011), https://www.copyright.gov/title17/92chap5.html#512.

[6] 17 U.S.C. 512(c) (2011), https://www.copyright.gov/title17/92chap5.html#512.

Wojcicki "revealed that 400 hours of video are now uploaded to YouTube every minute."[7] To be sure, this remarkable wealth of content should be viewed as a valuable benefit to society, demonstrating that the technological revolution has led to the progress of useful arts. Despite newfound difficulties in copyright management and enforcement, it is undeniable that we are living in an age of unparalleled content creation. The digital age has seen a blossoming of content online—from open source software development and self-published authors to homebrewed video producers and independent musicians reaching wider audiences than they ever would have been able to before the advent of the Internet. Far from stymying creativity, the current system has allowed content to bloom in a way never before thought possible. Much of that has been due to the compromises driven by the DMCA.

However, the vast amount of user-generated material uploaded online has made filtering out infringing content more challenging for the parties affected. Two groups propel the DMCA's procedure for identifying and removing infringing online content—copyright owners (or their representatives) and the OSPs themselves.[8] Each group's ability to police infringing content has been affected by the enormous amount of material hosted online.[9] Those enforcing the rights of copyright owners may be incapable of manually searching for every OSP hosting infringing content.[10] By the same token, if OSPs experience a dramatic increase in notifications about infringing content, they may not have the resources to manually evaluate and expeditiously remove the content.[11]

Interestingly, a private-sector solution has emerged—the use of automated systems.[12] This paper will refer to those systems more generally as content recognition systems ("CRSs"). For the sake of clarity, this paper will refer to the software incorporated into these systems simply as "content recognition software." One example includes the popular smartphone app *Shazam*.[13] *Shazam* uses smart phones to "listen" to a few seconds of a song and identify the artist and song title.[14] CRSs are capable of identifying, with various degrees of technical accuracy, specific content. This technology, applied in an automated fashion to the realm of copyright, is the topic of this paper.

---

[7] Greg Jarboe, *VidCon 2015 Haul: Trends, Strategic Insights, Critical Data, and Tactical Advice*, TUBULAR INSIGHTS (July 27, 2015), http://tubularinsights.com/vidcon-2015-strategic-insights-tactical-advice/.

[8] JENNIFER M. URBAN, JOE KARAGANIS & BRIANNA L. SCHOFIELD, NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE 31 (2016), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2755628.

[9] *Id.*

[10] *Id.*

[11] *Id.* at 31–32.

[12] *Id.*; Michael W. Carroll, *Pinterest and Copyright's Safe Harbors for Internet Providers*, 68 U. MIAMI L. REV. 421 (2014) ("In recent years, some copyright owners have invested in automated enforcement by relying on computer algorithms to identify the presence of a copyrighted work . . . and to send an automated take-down notice to service providers . . . . Some service providers, such as YouTube, have invested in automated processes to respond to such notices.").

[13] *Shazam – Name Any Song in Seconds*, SHAZAM, https://www.shazam.com/ (last visited Mar. 21, 2017).

[14] *About Us*, SHAZAM, https://www.shazam.com/company (last visited Mar. 21, 2017).

There are two types of automated CRSs discussed in this paper. In order to avoid conflating the two, their operations and effects need to be clearly differentiated. Both types utilize content recognition software as identification mechanisms, but aim to achieve different objectives.

The first are automated systems specifically used for *identifying* potentially infringing works and *remedying* disputations through non-legal channels, which we will refer to as automated identification and remediation systems ("AIRSs"). Once potentially infringing content is identified, these systems invoke a pre-chosen process selected by the rightsholder: block the infringing content, permit it to remain available on the host's website, or monetize a share of the revenue (often through associated advertisements, as in the case of YouTube's Content ID software). These systems have been developed by private sector firms operating under the constraints imposed by the DMCA. Their voluntary adoption has largely helped ease many of the contentious issues that would otherwise have existed between rightsholders and content creators in the digital age.

AIRSs are already in use online for a variety of purposes. For instance, every time a video is uploaded to YouTube, the video is scanned for the presence of copyrighted material owned by other users. If potentially infringing material is found, the content may be monetized, blocked, or left up, as determined by the owner of the infringed content.[15]

The second type utilizes automated CRSs for the specific purpose of enforcing copyright protections. Examples of the second type of automated system include content recognition software used by schools and universities to scan student papers automatically for plagiarism.[16] This paper will refer to these as automated copyright enforcement systems ("ACESs"). These are distinct from AIRSs in that the ultimate goal of an ACES is to "enforce" copyright claims, not remedy disputes through voluntary or non-legal channels.

Though CRSs are used for many purposes, very few detailed explanations of how they work are available.[17] Perhaps the most comprehensive description is found in a recent report released by the non-profit group Engine, entitled *The Limits of Filtering: A Look at the Functionality & Shortcomings of Content Detection Tools ("The Limits of Filtering")*.[18] *The Limits of Filtering* describes three methods used to determine whether content is infringing: metadata filtering, hash-based filtering, and content fingerprinting.[19] As *The Limits of Filtering* notes, infringers can

---

[15] *How Content ID Works*, YOUTUBE, https://support.google.com/youtube/answer/2797370?hl=en (last visited Mar. 21, 2017).

[16] *Professors Use Technology to Fight Student Cheating*, U.S. NEWS & WORLD REPORT (Oct. 3, 2008, 5:26 PM), https://www.usnews.com/education/articles/2008/10/03/professors-use-technology-to-fight-student-cheating.

[17] URBAN, KARAGANIS & SCHOFIELD, *supra* note 8 at 3 ("Some secrecy is defensible. Rightsholders expressed concern that publicizing details of their enforcement practices might enable pirates to circumvent them. OSPs worried that revealing their practices could subject them to negative attention from rightsholders, targets, or other OSPs, forcing them to change their practices.").

[18] *See* EVAN ENGSTROM & NICK FEAMSTER, THE LIMITS OF FILTERING: A LOOK AT THE FUNCTIONALITY & SHORTCOMINGS OF CONTENT DETECTION TOOLS (2017), http://www.engine.is/the-limits-of-filtering.

[19] *Id.* at i–ii.

avoid detection under the first method, metadata filtering systems, by simply altering the metadata information of the infringing file.[20] Similarly, even minor alterations of an infringing file, such as shortening the file or converting it to a different file type, may cause the second method, hash-based filtering systems, to miss a match.[21] Since the third method, content fingerprinting, more closely evaluates the actual content of files,[22] this paper's analysis will focus on that method.

One caveat: It is significant to note that cryptographic hash functions are often used to generate digital fingerprints of files, and as a result, some content fingerprinting systems could be considered hash-based.[23] Though cryptographic hash functions are a type of hash function, they are described as a "more powerful" form of hash-based filtering, and differ from other hash functions because they are much less likely to result in two different files returning the same output.[24] *The Limits of Filtering* makes a further distinction, remarking that because content fingerprinting systems "rely on algorithms that process the underlying media content of a given file, they are naturally constrained to a small subset of copyrightable content."[25] Thus, while some refer to cryptographic hash functions as simply "hash functions," this is technically incorrect, since the term "hash functions" serves as a generic description that encompasses both cryptographic hash functions (like those used for content fingerprinting) and "other sorts of algorithms like cyclic redundancy checks."[26] While some descriptions of content fingerprinting may confusingly refer to the use of "hashes," this likely references the particular subset of cryptographic hash functions, and not hash-based systems in general.

On a basic level, most content fingerprinting software systems perform the five following steps:

(1) Copyrighted material is submitted for inclusion in the software's database;
(2) The software analyzes the material for distinctive features and creates a "digital fingerprint" of the material based on this analysis;
(3) The digitally fingerprinted material is added to the software's database;
(4) New material being evaluated for a match is similarly scanned and compared against material in the software's database;

[20] *Id.* at 12.
[21] *Id.* at 13.
[22] *Id.* at 13–14.
[23] *See, e.g.*, Simson Garfinkel, *Fingerprinting Your Files*, MIT TECHNOLOGY REVIEW (Aug. 4, 2004), https://www.technologyreview.com/s/402961/fingerprinting-your-files/; David Ashfield, *Why Are Cryptographic Hash Functions Important in Digital Forensics?,* CCL GROUP (Jan. 30, 2014), https://www.cclgroupltd.com/cryptographic-hash-functions-important-digital-forensics/; *What is Cryptographic Hashing? MD5, SHA, and More*, TIPTOP SECURITY (Dec. 15, 2014), https://tiptopsecurity.com/what-is-cryptographic-hashing-md5-sha-and-more/.
[24] JOHN EDWARD SILVA, AN OVERVIEW OF CRYPTOGRAPHIC HASH FUNCTIONS AND THEIR USES 3 (2003), https://www.sans.org/reading-room/whitepapers/vpns/overview-cryptographic-hash-functions-879.
[25] ENGSTROM & FEAMSTER, *supra* note 18.
[26] Tim Fisher, *Cryptographic Hash Function*, LIFEWIRE (updated Oct. 30, 2015), https://www.lifewire.com/cryptographic-hash-function-2625832.

(5) If the software recognizes enough of a similarity between the new material and material in its database, a match is identified and flagged.[27]

What constitutes a "match" depends on how the software has been programmed. Factors such as the types and percentage of similarities serve as parameters for the software's evaluation. The parameters used and their level of sensitivity are features usually influenced by the needs of the user licensing or commissioning the software.[28]

To protect themselves, groups using content recognition software may set parameters at a very high level of responsiveness. For example, some OSPs have adopted automated procedures that "go beyond measures required by section 512 [of the DMCA]."[29] This type of cautious behavior has been observed by both rightsholders and OSPs attempting to preserve their rights and protections under the DMCA. The effect of these modifications is that systems sometimes flag material that is not actually infringing under the current legal standard. At least one report noted several issues resulting from this incorrect application of the law, finding "nearly 30% of [automated] takedown requests were of questionable validity."[30] This suggests that rightsholders and OSPs are engaging in a system of de facto self-policing that is less effective than it could be, given prevailing legal standards.[31]

Harms caused by erroneous DMCA takedown notices resulting from this self-policing include:

- **Clearly mismatched material.** For instance, a notice was sent on behalf of Paramount accusing Paramount's *Anchorman: The Legend of Ron Burgundy* of infringing the Paramount movie *An Officer and a Gentleman*.[32]
- **Failure to consider fair use or "legitimate political speech."** NBC issued a takedown notice for a "satirical Obama campaign video, which used archival footage of Tom Brokaw to announce the 'bad news' that Senator John McCain had been elected president." The video was "[i]ntended to motivate Obama supporters to 'get out the vote,'" but "was taken down days before a critical voter registration deadline."[33]
- **Improper takedowns caused by mistaken claims of ownership.** In 2012, YouTube footage of NASA's Curiosity rover landing on Mars, a public domain video posted by NASA on its own YouTube page, was removed because of a takedown notice sent by Scripps News Service incorrectly claiming Scripps' ownership of the video.[34]

---

[27] Jonathan Strickland, *How Content-Recognition Software Works,* HOWSTUFFWORKS (July 16, 2007), http://computer.howstuffworks.com/content-recognition.htm.
[28] *Id.*
[29] URBAN, KARAGANIS & SCHOFIELD, *supra* note 8 at 1–2.
[30] *Id.* at 3.
[31] *Id*. at 2–3.
[32] *Id.* at 91.
[33] *NBC Issues Takedown on Viral Obama Ad*, ELECTRONIC FRONTIER FOUNDATION, https://www.eff.org/takedowns/nbc-issues-takedown-viral-obama-ad (last visited Mar. 20, 2017).
[34] Alexander Higgins, *NASA's Mars Rover Crashed into a DMCA Takedown*, BEFORE IT'S NEWS (Aug. 7, 2012, 1:48 PM), http://beforeitsnews.com/alternative/2012/08/nasas-mars-rover-crashed-into-a-dmca-takedown-2448814.html.

- **Abuse of the DMCA Notice and Takedown process.** In one case, a political candidate allegedly used the DMCA to shut down the YouTube account of People for the American Way's Right Wing Watch ("RWW") by sending "a series of takedown notices claiming that RWW's use of clips of his program infringed his copyrights," though "RWW was [] posting them as part of its critical work." The "bogus complaints" caused RWW's entire YouTube account "to be taken offline—twice."[35]

These and other examples suggest the current copyright enforcement system is a combination of statutes and precedent more suited to a pre-Internet world, and a self-policing system that often improperly applies the law.

This paper discusses the legal implications of a government-adopted ACES. Based on systems that already exist, but modified for the purpose of more closely applying the law, such an automated system may help resolve two of the biggest shortcomings in the current copyright system. First, it provides a more appropriate system of evaluation, as it addresses a technology-based problem with a technology-based solution. Second, it reshapes the existing private sector solution by closely applying the legal parameters of infringement, offering a more accurate evaluation of what is (and is not) infringing. In short, it bridges the gap between the systems of creation and delivery, and the system of evaluation.

However, it is worth noting that the technical complexities in creating such a system may make its implementation difficult. We also argue that the DMCA has, on net, been a benefit to both content creators and rightsholders, allowing for an explosive outgrowth of new content to dominate a newly interconnected world. As a result, we note that the technology to maintain a repository of various copyrighted content may, in the future, help better identify rightsholders and assist in detection of infringing materials. In the meantime, however, the existing safe harbor system, buttressed by innovative technological solutions like CRSs, is working well, and any changes to the system should aim to maximize the options for rightsholders when their works are used by others.

Part I of this paper describes the current state of the copyright enforcement system. Part II examines CRSs currently in use by the private sector. Part III considers potential uses for an ACES. Part IV outlines recommended steps to design and implement an ACES for general use. Finally, Part V addresses potential obstacles to implementation.

Ultimately, this paper concludes that an ACES may eventually be a viable step in the evolution of copyright law. The structure of an ACES has already been created through private sector development. Thus, its implementation could possibly be achieved through the licensing and modification of existing automated systems. However, the creation of such a system may rest in the distant future, largely due to the technical limitations inherent in accurately detecting infringing content. Given the current ecosystem, we argue that promoting CRSs and private

[35] *Attempt to Silence the Political Speech at Right Wing Watch*, ELECTRONIC FRONTIER FOUNDATION, https://www.eff.org/takedowns/attempt-silence-political-speech-right-wing-watch (last visited Mar. 20, 2017).

sector, voluntary solutions—all within the framework of the DMCA—is a far better alternative for ensuring the continued progress of the useful arts.

I. The Current State of Copyright Enforcement

The current copyright enforcement system leaves a great deal of uncertainty about what content copyright owners actually possess and how well this content will be protected. This uncertainty enhances the attractiveness of an automated system that consistently applies the law, which may lead to more predictable outcomes.

Between the long duration of copyright ownership[36] and the unclear line separating copyrightable content from the public domain,[37] it is often unclear what material is protectable. Additionally, the standard for infringement has been inconsistently applied, making it difficult for copyright owners to determine what constitutes a violation.[38]

For those who own copyrights, it is difficult to protect and fully exploit their property. For those who have created profitable new content, earnings are jeopardized by the higher likelihood of infringement suits. Those issues, however, must necessarily be compared to the considerable benefits associated with the recent content explosion made possible by the era of digital communications. In the years following the passage of the DMCA, the global economic value of

[36] Currently, the duration of the copyright term for single authors covers the life of the author plus 70 years. 17 U.S.C. 302 (2011). For a more in-depth discussion of how copyright duration has increased over time, see Chapter 6 of TOM W. BELL, INTELLECTUAL PRIVILEGE: COPYRIGHT, COMMON LAW, AND THE COMMON GOOD (2014). *See also* Paul Belleflamme, *The Economics of Copyright Protection*, IPDIGIT (Oct. 2, 2013), http://www.ipdigit.eu/2013/10/the-economics-of-copyright-protection/.

[37] For instance, *Copyright & Fair Use* warns of several types of works that, while based on public domain materials, are not in the public domain due to any number of reasons, including modification, protection under trademark law, or protection in other countries. *See* Rich Stim, *Public Domain Trouble Spots*, COPYRIGHT & FAIR USE: STANFORD UNIVERSITY LIBRARIES, http://fairuse.stanford.edu/overview/public-domain/trouble-spots/ (last visited Jan. 6, 2017).

[38] *See, e.g.*, Rachel Isabelle Butt, *Appropriation Art and Fair Use*, 25 OHIO ST. J. ON DISP. RESOL. 1055 (2010), https://kb.osu.edu/dspace/bitstream/handle/1811/76940/OSJDR_V25N4_1055.pdf?sequence=1 ("The current system to determine a fair use through litigation is ineffective as applied to conflicts involving visual arts because the courts have misused the fair use doctrine. The first problem is that inconsistent case law regarding fair use and copyright infringement fails to provide guidance for artists. The second problem is that the judicial concept of copyright infringement does not comport with the accepted norms in the art world; artists historically have borrowed and copied existing expression without objection or conflict."); Tonya M. Evans, *Sampling, Looping, and Mashing . . . Oh My!: How Hip Hop Music is Scratching More Than the Surface of Copyright Law*, 21 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 843, 847–48 (2011), http://ir.lawnet.fordham.edu/iplj/vol21/iss4/1 ("different infringement standards are being applied by the circuit courts to sound recording infringement cases . . . . [t]he resulting incongruent decisions reflect an inconsistent application of federal law."); Mark A. Lemley, *Our Bizarre System for Proving Copyright Infringement*, 57 J. COPYRIGHT SOC'Y U.S.A. 719, 741–42 (2010) (criticizing the existing standard for demonstrating copyright infringement in 2010, stating "[o]ur rules for proving copyright infringement make little sense."); Pamela Samuelson, *A Fresh Look at Tests for Nonliteral Copyright Infringement*, 107 NW. L. REV. 1821, 1823 (2013), http://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1045&context=nulr ("One reason why conventional tests for judging nonliteral copyright infringement are problematic is that there are too many tests and not enough guidance about which one to use in what kinds of cases.").

the entertainment industry has grown by 66 percent, "and U.S. households spent almost 15 percent more on entertainment as a percentage of income in 2008 than in 2000."[39] As noted in *The Limits of Filtering*, "[i]n terms of actual content creation, the years since the DMCA's passage have been amongst the most prolific in history. [OSPs] have seen similar growth, as the [I]nternet sector contributed nearly $1 trillion to the U.S. economy—nearly 6 percent of real GDP—in 2014."[40] Nonetheless, legal uncertainty persists in the copyright realm. A discussion of the current state of copyright enforcement is therefore worth delving into.

What follows is a discussion of factors that have contributed to the uncertainty of copyright ownership, concerns regarding regulation of online distribution, and the current self-policing system.

A. The Uncertainty of Copyright Ownership

There are a number of factors that contribute to the uncertainty of rights possessed by copyright owners.

It is not always clear what is protected by copyright. The courts have acknowledged that "[t]he determination of [copyright] infringement is one of the most difficult of all legal questions,"[41] and at least one observer has noted that the courts' rules for evaluating copyright infringement "make little sense."[42]

It is also unclear what is copyrighted and what may be freely used because it is part of the public domain. Material in the "public domain" has been defined as "creative materials that are not protected by intellectual property laws such as copyright, trademark, or patent laws. The public owns these works, not an individual author or artist. Anyone can use a public domain work without obtaining permission, but no one can ever own it."[43] However, determining whether material falls within the public domain is challenging, since the answer depends on when the work was created, whether and when the work was published, and whether the work in question includes material from an earlier work.[44]

---

[39] ENGSTROM & FEAMSTER, *supra* note 18, at 9; *see also* Joshua P. Friedlander, *News and Notes on 2015 RIAA Shipment and Revenue Statistics,* https://www.riaa.com/wp-content/uploads/2016/03/RIAA-2015-Year-End-shipments-memo.pdf (last visited April 9, 2017).

[40] ENGSTROM & FEAMSTER, *supra* note 18, at 9–10.

[41] Michelle V. Francis, *Musical Copyright Infringement: The Replacement of Arnstein v. Porter–A More Comprehensive Use of Expert Testimony and the Implementation of an "Actual Audience" Test*, 17 PEPP. L. REV. 493, 494 (1990) (citing A. SHAFTER, MUSICAL COPYRIGHT 146 (2d ed. 1932)).

[42] Lemley, *supra* note 38, at 741–42.

[43] Rich Stim, *Welcome to the Public Domain*, STANFORD UNIVERSITY LIBRARIES, http://fairuse.stanford.edu/overview/public-domain/welcome/ (last visited Jan. 14, 2017).

[44] *See* Office of the General Counsel, *Copyright and Fair Use*, HARVARD UNIVERSITY, http://ogc.harvard.edu/pages/copyright-and-fair-use (last visited Jan. 14, 2017).

Perhaps most importantly, under the courts' current standard of evaluation, it is not clear what is and is not infringement. The tests courts utilize vary from one circuit to the next, which creates inherent uncertainty about how a copyright case will be resolved. In addition, the tests have been applied inconsistently, increasing uncertainty.[45] To illustrate this, it may be helpful to explain what infringement is and describe the current tests conducted by the courts.

Copyright Infringement

Copyright infringement occurs when any of the copyright owner's exclusive rights under the Copyright Act have been violated.[46] If a copyright owner believes an infringement has occurred, the owner may sue the alleged infringer to enforce the rights the owner believes have been violated.[47]

There are two requirements for demonstrating copyright infringement: (1) ownership of a valid copyright by the plaintiff, and (2) copying by the defendant.[48] The first requirement, ownership, is typically established through proof of registration with the U.S. Copyright Office.[49] The second requirement, "copying by the defendant," consists of two separate elements: "copying in fact," and "copying as a legal proposition."[50]

Element one, "copying in fact," may be proven with direct or indirect evidence.[51] If direct evidence of copying cannot be shown, copying in fact may be shown indirectly by demonstrating that the alleged infringer had access to the earlier work, and the two works in question share "probative similarity."[52]

Element two, "copying as a legal proposition," also known as "substantial similarity,"[53] is arguably the more difficult element to prove. The circuits employ different tests to evaluate substantial similarity, and, as *Nimmer on Copyright* notes,[54] "[a]lthough it is clear that the determination of substantial similarity presents an issue of fact, the correct procedure for that determination remains clouded."[55] The traditional test for substantial similarity is the "Audience Test," which *Harold Lloyd Corp. v. Witwer* states as the following:

---

[45] *See, e.g.*, *supra* note 38.
[46] Catherine Palo, *Copyright Infringement Litigation*, *in* 77 AM. JUR. TRIALS 449 § 23 (updated Feb. 2016).
[47] *Id.*
[48] MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.01 (Matthew Bender, rev. ed. 2015).
[49] *Id.*
[50] *Id.*
[51] *Id.*
[52] *Id.*
[53] *Id.* at 13-9 to -10.1.
[54] NIMMER ON COPYRIGHT is considered one of the leading treatises on U.S. copyright law. *Nimmer on Copyright*, LEXISNEXIS STORE, https://store.lexisnexis.com/products/nimmer-on-copyright-skuusSku10441 (last visited Mar. 28, 2017).
[55] NIMMER & NIMMER, *supra* note 48 at § 13.03(E)(3).

> *The question really involved in such comparison is to ascertain the effect of the alleged infringing [work] upon the public, that is, upon the average reasonable man. If an ordinary person who has recently read the story sits through the presentation of the picture, if there had been literary piracy of the story, he should detect that fact without any aid or suggestion or critical analysis by others. The reaction of the public to the matter should be spontaneous and immediate.*[56]

Many Circuits have modified the Audience test. For instance, the Second Circuit's test for substantial similarity turns on (1) whether the defendant copied the plaintiff's work (a query based on "expert analysis and dissection"), and (2) whether that copying "constitute[s] an improper appropriation" (a jury evaluation that "would strictly apply the audience test and preclude dissection and expert analysis").[57]

The Ninth Circuit employs a two-prong substantial similarity test, consisting of (1) an extrinsic objective analysis conducted by the court (limited to "specific criteria which can be listed and analyzed" such as "analytic dissection and expert testimony"), and (2) an intrinsic subjective analysis conducted by the jury (based on "the response of the ordinary reasonable person").[58]

Other circuits have adopted the Second Circuit test, the Ninth Circuit test, a combination of the two approaches, or variations on the themes of each.[59] Therefore, the circuit evaluating a copyright infringement claim dictates which test likely applies. Yet, even when armed with a description of the tests utilized, the outcomes of cases are not consistent, increasing the level of uncertainty.

Recent Cases Challenging Traditional Interpretations of Infringement Claims

One area that may be experiencing significant upheaval is music copyright law. It was once considered very difficult to prove a music copyright infringement claim in court.[60] Several well-known cases illustrate this difficulty.

One example is the Second Circuit case *Selle v. Gibb*.[61] The suit alleged that The Bee Gees' hit song *How Deep Is Your Love* infringed an earlier work written by a relatively unknown artist.

---

[56] Harold Lloyd Corp. v. Witwer, 65 F.2d 1, 18 (9th Cir. 1933); NIMMER & NIMMER, *supra* note 48**,** at § 13.03(E)(3).
[57] NIMMER & NIMMER, *supra* note 48**,** at § 13.03(E)(3).
[58] Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp., 562 F.2d 1157, 1163–64 (9th Cir. 1977).
[59] NIMMER & NIMMER, *supra* note 48**,** at § 13.03(E)(3).
[60] *See, e.g.*, Debra Presti Brent, *The Successful Musical Copyright Infringement Suit: The Impossible Dream*, 7 U. MIAMI ENT. & SPORTS L. REV. 229, 229 (1990) (representing the historical view that considered "[m]usical copyright protection . . . a misnomer," and under the traditional regime, "[a] plaintiff seeking to protect his property interest finds little sympathy from the judiciary."); William R. Coulson, *They're Playing Our Song! The Promise and the Perils of Music Copyright Litigation*, 13 J. MARSHALL REV. INTELL. PROP. L. 555, 575 (2014) ("So, do composers ever 'borrow' from each other? Of course, it happens. Examples abound. Even if it amounts to larceny, is it easy to prove? Decidedly not."); *see also* Francis, *supra* note 41, at 494; Lemley, *supra* note 38, at 741–42.
[61] *See* Selle v. Gibb, 741 F.2d 896 (7th Cir. 1984).

The two songs in question shared thirty identical notes and more than forty identical rhythmic patterns, a level of similarity that an expert witness testified could not have occurred independently.[62] Despite these similarities, an expert witness's testimony, a jury decision in favor of the plaintiff, and one of The Bee Gees mixing up the two songs in court during the trial, the court refused to support a finding of access or infringement.[63]

Another example is the 2003 Ninth Circuit case *Newton v. Diamond*, where an exact copy of six seconds of Newton's sound recording was sampled and looped to repeat more than forty times in a Beastie Boys song.[64] In a case where repeated copying was absolutely proven, the copying of three identical protected notes was found to be too minimal to constitute infringement.[65]

Ninth Circuit case *Swirsky v. Carey* demonstrates the difficulty of surviving summary judgment when bringing a music infringement claim.[66] *Swirsky* was originally dismissed; however, on appeal, the court reversed because the plaintiff's expert witness identified a great number of similarities shared by the songs.[67] Though the songs shared nearly identical structures and highly similar bass lines, chord changes, and tempos, the case was originally dismissed under summary judgment, arguably due to the difficulty of applying the law.

The outcome of the recent *Blurred Lines* case may have upset this and other past legal precedent.[68] In 2013, songwriters for the hit song *Blurred Lines* faced allegations that the song infringed Marvin Gaye's classic *Got to Give It Up*.[69] Despite the absence of any identical notes, chord progressions, lyrics, or melodies, the court found for the Gaye estate.[70] Many have suggested that the *Blurred Lines* verdict may have increased concerns over the risk of liability for music copyright infringement.[71]

---

[62] *See id.* at 899–906.

[63] *See id.* at 899–906; *Selle v. Gibb 741 F.2d 896 (7th Cir. 1984)*, MUSIC COPYRIGHT INFRINGEMENT RESOURCE, http://mcir.usc.edu/cases/1980-1989/Pages/sellegibb.html (last visited Mar. 17, 2017).

[64] *See* Newton v. Diamond, 388 F.3d 1189 (9th Cir. 2003).

[65] *See id.* at 1190, 1196–97.

[66] *See* Swirsky v. Carey, 376 F.3d 841 (9th Cir. 2004), *as amended on denial of reh'g* (Aug. 24, 2004).

[67] *See id.* at 845–46.

[68] For a more detailed discussion, please see Regina Zernay, Comment, *Casting the First Stone: The Future of Music Copyright Infringement Law After* Blurred Lines*,* Stay with Me*, and* Uptown Funk, 20 CHAP. L. REV. 177 (2017), http://www.chapmanlawreview.com/wp-content/uploads/2017/04/20-1_Zernay.pdf.

[69] Complaint for Declaratory Relief at 4, Williams v. Bridgeport Music, Inc., No. CV13-06004-JAK (AGRx) (C.D. Cal. Aug. 15, 2013), 2013 WL 4271752.

[70] Williams v. Bridgeport Music, Inc., No. LACV13-06004JAK(AGRX), 2014 WL 7877773, at *6 (C.D. Cal. Oct. 30, 2014).

[71] *See, e.g.*, Adam R. Bialek, *California Jury Finds "Blurred Lines" Infringed "Got To Give It Up": Society's Mixed Signals On Copying and Intellectual Property Rights,* 27 INTELL. PROP. & TECH. L.J. 14, 15 (2015); Blake Brittain, *Musicians More Careful After 'Blurred Lines' Case*, BLOOMBERG BNA (Sept. 17, 2015), http://www.bna.com/musicians-careful-blurred-n17179936188/; Keith Murphy, *7 Reasons the 'Blurred Lines' Verdict Should Have Everyone Spooked*, BET (last visited Feb. 10, 2016), http://www.bet.com/music/photos/2015/03/7-reasons-the-blurred-lines-verdict-should-have-everyone-spooked.html; Tim Wu, *Why the "Blurred Lines" Copyright Verdict Should Be Thrown Out*, THE NEW YORKER (Mar. 12,

Similarly, it is not always clear what role fair use will play in a copyright infringement controversy. The law states that the "fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright."[72] Thus, if the use of copyrighted material in a new work is considered fair use, the use is not infringing. The courts use four factors to determine if fair use applies:

(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.[73]

Case law suggests it is not always clear when and how fair use will apply. For instance, in *Cariou v. Prince*, Richard Prince, "a well-known appropriation artist,"[74] was sued for infringement when he created and exhibited art that incorporated copyrighted photographs taken by photographer Patrick Cariou without his permission.[75] On appeal, the court held that while Prince did not express any intent to "comment[] on Cariou's work or on culture," such an intent is not required to qualify for fair use.[76] Instead, the court stated that the "critical" inquiry is "how the work in question appears to the reasonable observer, not simply what an artist might say about a particular piece or body of work. Prince's work could be transformative even without commenting on Cariou's work or on culture, and even without Prince's stated intention to do so."[77] The court found that Prince's works were a "transformative use," since "in twenty-five of his artworks, Prince has not presented the same material as Cariou in a different manner, but instead has 'add[ed] something new' and presented images with a fundamentally different aesthetic."[78]

Fair use has also been found to apply where images are "transformatively different from their original expressive purpose," even though the images themselves have been used "in their

---

2015), http://www.newyorker.com/culture/culture-desk/why-the-blurred-lines-copyright-verdict-should-be-thrown-out.
[72] 17 U.S.C. § 107 (2011), https://www.copyright.gov/title17/92chap1.html#107.
[73] *Id.*
[74] Cariou v. Prince, 714 F.3d 694, 699 (2d Cir. 2013) (explaining that "[t]he Tate Gallery has defined appropriation art as 'the more or less direct taking over into a work of art a real object or even an existing work of art.'").
[75] *Id.* at 698.
[76] *Id.* at 707.
[77] *Id.*
[78] *Id.* at 708.

entirety."[79] In *Bill Graham Archives v. Dorling Kindersley Ltd.*, Dorling Kindersley ("DK") published *Grateful Dead: The Illustrated Trip* ("*Illustrated Trip*")," a coffee table book "intended as a cultural history of the Grateful Dead."[80] Plaintiff Bill Graham Archives ("BGA"), "claim[ed] to own the copyright to seven images displayed in *Illustrated Trip*, which DK reproduced without BGA's permission . . . . in significantly reduced form."[81] The images were pictures of concert posters advertising Grateful Dead shows, greatly reduced in size, and were included in the book next to several other related pictures. The court found that "even though the copyrighted images are copied in their entirety, the visual impact of their artistic expression is significantly limited because of their reduced size."[82] Because the remaining fair use factors "weigh[ed] in favor of DK's use," the use of BGA's images in *Illustrated Trip* qualified as fair use.[83]

It is likely that the role of fair use in online infringement cases will also change following *Lenz v. Universal Music Corp.*[84] In *Lenz*, Plaintiff Stephanie Lenz posted a 29-second video on YouTube featuring a baby dancing while a song by the artist formerly known as Prince played in the background.[85] Universal Music, which represents Prince, sent YouTube a DMCA takedown notification for the video, claiming it infringed the song's copyright. In response, Lenz sued Universal, alleging that Universal failed to "first evaluate whether the content qualifies as fair use," in violation of the DMCA's takedown notice procedure.[86] The court held "that the [DMCA] statute requires copyright holders to consider fair use before sending a takedown notification."[87] Based on this decision, it is likely that the fair use assessment requirement will be more strictly enforced by the courts, and the absence of a fair use assessment may serve as a defense for alleged infringers challenging DMCA takedowns.

It is also worth noting that the defense of accidental or "innocent" infringement is not recognized by courts. Accidental or "innocent" infringement "occurs when someone engages in infringing activity not knowing that her conduct constitutes infringement—perhaps most commonly when she knowingly copies from another's work but reasonably believes that her copying is not infringing."[88] Importantly, "since 1931, a defendant's mental state has clearly not been relevant . . . to the question of liability for direct copyright infringement. As the Supreme Court stated that year, '[i]ntention to infringe is not essential under the Act.' So innocent infringers are just as liable as those who infringe knowingly or recklessly."[89] Thus, innocent infringement is still actionable, and once infringement has been shown, the infringer is liable.

---

[79] Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 613–614 (2d Cir. 2006).
[80] *Id.* at 607.
[81] *Id.*
[82] *Id.* at 613, 615.
[83] *Id.*
[84] *See* Lenz v. Universal Music Corp., 815 F.3d 1145 (9th Cir. 2015), *cert. denied*, 137 S. Ct. 416 (2016).
[85] *See id.* at 1148.
[86] *Id.*
[87] *Id.*
[88] R. Anthony Reese, *Innocent Infringement in U.S. Copyright Law: A History*, 30 COLUM. J.L & ARTS 133 (2007).
[89] *Id.* (internal citations omitted).

The outcomes in these and other copyright infringement cases suggest that in some situations, the courts' standard of evaluation may lead to inconsistent results. Offering a reliable and relatively predictable automated system of evaluation may help minimize such inconsistencies.

B. Concerns Regarding Regulation of Online Distribution

With the development of technology and its role in both creating and distributing copyrighted content, regulation has been a longstanding issue.[90] One form of regulation that is a constant concern is government regulation of the Internet. In a 2015 report issued by the Federal Communications Commission ("FCC"), the importance of maintaining an open Internet was mentioned:

> *The open Internet drives the American economy and serves, every day, as a critical tool for America's citizens to conduct commerce, communicate, educate, entertain, and engage in the world around them. The benefits of an open Internet are undisputed. But it must remain open: open for commerce, innovation, and speech; open for consumers and for the innovation created by applications developers and content companies; and open for expansion and investment by America's broadband providers. For over a decade, the Commission has been committed to protecting and promoting an open Internet.*[91]

Many of the recent discussions surrounding government regulation of the Internet focus on the DMCA's safe harbor provisions and its notice and takedown procedure. The entertainment industry in particular has been quite vocal about reforming the DMCA so that OSPs are held more accountable for hosting infringing content.[92] Not surprisingly, content hosts and related parties have protested such changes, citing concerns that it would chill creativity and the continued development of the Internet as an information-sharing tool.

Content owners and their representatives claim that stronger rules are necessary to protect the rights of copyright owners.[93] Some have suggested that the DMCA notice and takedown procedure has become an antiquated and ineffective means of policing infringement.[94] Others

---

[90] *See, e.g.*, Lawrence Lessig, *Code Is Law: On Liberty in Cyberspace*, HARVARD MAGAZINE (June 1, 2000), http://harvardmagazine.com/2000/01/code-is-law-html.

[91] FCC, REPORT AND ORDER ON REMAND, DECLARATORY RULING, AND ORDER FCC 15-24 (FCC Rcd.) (2015), http://transition.fcc.gov/Daily_Releases/Daily_Business/2015/db0312/FCC-15-24A1.pdf.

[92] *See, e.g.*, Micah Singleton & Ben Popper, *The Music Industry Cranks Up the Volume in Its Fight Against YouTube*, THE VERGE (June 3, 2016, 12:48 PM), http://www.theverge.com/2016/6/3/11852146/music-industry-fighting-youtube-dmca.

[93] *See, e.g.*, Irving Azoff, *Dear YouTube: An Open Letter from Irving Azoff*, RECODE (May 9, 2016, 6:00 AM), http://www.recode.net/2016/5/9/11609494/irving-azoff-youtube-artists-streaming-music.

[94] *See, e.g.*, BRUCE BOYDEN, THE FAILURE OF THE DMCA NOTICE AND TAKEDOWN SYSTEM: A TWENTIETH CENTURY SOLUTION TO A TWENTY-FIRST CENTURY PROBLEM (2013), http://cpip.gmu.edu/wp-content/uploads/2013/08/Bruce-Boyden-The-Failure-of-the-DMCA-Notice-and-Takedown-System1.pdf.

have stated that the current system "unfairly favor[s] technology companies and rogue pirate sites" and "forces creators to police the entire Internet for instances of theft, placing an undue burden on these artists."[95]

In contrast, content hosts and related parties discourage increased government involvement. Opponents of increased regulation argue that greater restrictions will interfere with the continued development of the Internet as a platform for innovation, free expression, the broad distribution of content, and creativity.[96] Some have stated that the DMCA takedown notice procedure has, in fact, succeeded in "strik[ing] a balance between facilitating free speech and creativity while protecting the interests of copyright holders" and "fuel[ing] the creation of a booming domestic Internet economy that was worth nearly $1 trillion or 6 percent of GDP in 2014."[97] Others have even expressed concern about the motivation behind the media's efforts to increase regulation, with at least one commentator claiming: "The Internet would have been destroyed . . . but the media did not care, just as long as they got their royalties."[98]

What neither side can ignore are the effects of existing overregulation caused by the current system of self-policing. While the DMCA notice and takedown process has helped content distribution and creation thrive online, there may be some overregulation in the DMCA's self-policing regime. Evidence indicates that the self-policing methods employed by content owners and OSPs interpret copyright law so broadly that an alarming number of parties have been improperly accused of infringement.[99] One study found that "in one in twenty-five cases, targeted content did not match the identified infringed work, suggesting that 4.5 million requests in the entire six-month data set were fundamentally flawed." In addition, "another 15% of the requests raised questions about whether they had sufficiently identified the allegedly infringed work of the allegedly infringing material."[100]

While this system has not prevented the explosive growth of online content creation since the DMCA came into effect, it has had an impact—one that may become more of a pressing

---

[95] Ashley Cullins, *Music Industry A-Listers Call on Congress to Reform Copyright Act*, THE HOLLYWOOD REPORTER (Mar. 31, 2016, 1:48 PM), http://www.hollywoodreporter.com/thr-esq/music-industry-a-listers-call-879718.

[96] Kerry Sheehan, *This Year in U.S. Copyright Policy: 2016 in Review*, ELECTRONIC FRONTIER FOUNDATION (Dec. 25, 2016), https://www.eff.org/deeplinks/2016/12/year-us-copyright-policy-2016-review.

[97] *DMCA Infringement Takedowns Work*, INTERNET ASSOCIATION, http://blog.internetassociation.org/post/141914748038/dmca-infringement-takedowns-work (last visited Jan. 16, 2017).

[98] Mike Konrad, *Copyright Strangulation*, AMERICAN THINKER (Sept. 14, 2013), http://www.americanthinker.com/articles/2013/09/copyright_strangulation.html#ixzz2euJwfxHT.

[99] *See, e.g.*, URBAN, KARAGANIS & SCHOFIELD, *supra* note 8, at 2 ("Study 2's quantitative analysis revealed deficiencies in notice and takedown procedures, especially automated requests . . . ."); Jon Gosnell, Comment, *Keeping the Internet Free: Why the DMCA's Safe Harbor Provision Should Be Expanded to Help Curb Overregulation of Content by Removing the Financial Benefit with Right and Ability to Control Exclusion* 1, 3 (2015), https://ssrn.com/abstract=2559455 ("For content hosting sites, there is a tension between going too far and not going far enough in protecting copyright holders' rights. . . . do too much, and information is needlessly stifled. This leads to absurd outcomes like presidential campaign ads getting flagged days before an election, or footage from a NASA rover landing on Mars getting taken down due to a claim from a news channel.").

[100] *Id.*

concern in the future. It is worth examining the source of this overregulation and the effect automated systems have in misapplying the law. Such an outcome merits attention, since "[t]he control over the shifting of costs from one actor to an innocent other is often stated as a justification for legal intervention."[101] A revised automated system for general use that accurately reflects the proper legal standard may one day offer a promising solution to this misapplication of the law and other related issues.

C. Overregulation Through Self-Policing

The DMCA outlines a procedure for copyright owners to enforce their rights, while still offering OSPs protection from secondary liability for hosting infringing content, so long as the service provider acts expeditiously to remove infringing content upon notification (among other requirements).[102]

The DMCA notice and takedown procedure consists of several steps. First, as noted in *Lenz*, the copyright owner must determine if the alleged infringement qualifies as fair use.[103] If, after conducting this analysis, the copyright owner has a good faith belief that the alleged infringement does not qualify as fair use, the copyright owner may send a takedown notice to an agent designated by the OSP to receive takedown notifications on the OSP's behalf.[104] Upon receiving this notification, the OSP must act expeditiously to remove the allegedly infringing content and promptly notify the user who posted the content that the content has been removed.[105] The user who posted the allegedly infringing content may reply to the OSP by sending a counter notification explaining why the content is not infringing.[106] If the counter notification is properly sent and received, the OSP must put the content back up and promptly notify the copyright owner that the content is being reposted within 10 business days, unless the copyright owner notifies the OSP that he or she has filed a court action against the alleged infringer.[107]

Scouring the Internet for potentially infringing material is an arduous task, one for which copyright owners and their representatives (sometimes collectively referred to as "rightsholders") may not have the resources to achieve. In response to this difficulty, automated systems or "software bots" were created to "crawl the web for infringing material."[108]

---

[101] Donald J. Kochan, *Bubbles (or, Some Reflections on the Basic Laws of Human Relations)*, FORDHAM ENVTL. L. REV. 133, 159 (2015).
[102] 17 U.S.C. 512(c) (2011), https://www.copyright.gov/title17/92chap5.html#512.
[103] Lenz v. Universal Music Corp., 815 F.3d 1145 (9th Cir. 2015), *cert. denied*, 137 S. Ct. 416 (2016).
[104] 17 U.S.C. 512, https://www.copyright.gov/title17/92chap5.html#512.
[105] *Id.*
[106] *Id*.
[107] *Id*.
[108] URBAN, KARAGANIS & SCHOFIELD, *supra* note 8, at 31.

The effect of this was a flood of DMCA takedown notices.[109] For example, "[i]n 2009, Google received 4,275 notices. In 2012, it received 441,370 notices."[110] By early 2016, "Google was receiving between 17 and 21 million requests a week."[111] To combat the dramatic increase in DMCA takedown notices, several of the larger OSPs turned to automated systems to process the overwhelming number of requests.[112]

OSPs are also very protective of their DMCA safe harbor protections.[113] Secondary liability for copyright infringement would expose OSPs to incalculable potential losses, as millions of users utilize OSPs, and each individual infringement constitutes a separate action.[114] In a proactive attempt to avoid hosting infringing material in the first place, sites like YouTube, Vimeo, and others have adopted AIRSs to detect infringement upon user upload.[115] In the case of YouTube, qualified users may select the action YouTube takes if its Content ID system detects potential infringement—block the video, monetize the video, or track the video's views.[116] For instances where the rightsholder and the alleged infringer both wish to monetize the video, the dispute may be remedied through a number of options, including an escrow-styled arbitration and dispute mechanism.[117] These AIRSs have, on net, been a positive contribution to the online digital landscape, ameliorating the need for disputes to take on a legal character, largely through monetization and voluntary remediation.

These sophisticated self-policing systems, however, appear to have incorporated parameters that go beyond what the law requires.[118] As a result, a great deal of content is now being improperly flagged as infringing.[119] In some cases, parties may have willfully abused the DMCA notice and takedown process.[120]

---

[109] *Id.* at 32.
[110] *Id.*
[111] *Id.*
[112] *Id.* at 53.
[113] *Id.* at 8.
[114] *Id.* at 130 ("Currently, the statute directs courts to calculate statutory damages by applying a set amount to each work infringed. This 'per infringed work' provision did not anticipate the numbers of works that users can place onto large online platforms or the absurd results—requests of billions or even trillions of dollars in damages—that can ensue. Fear of such staggering liability costs can cause OSPs to make overly conservative decisions regarding notice and takedown compliance. This is especially true when OSPs make decisions that could affect their eligibility for the safe harbor in the first place, because the loss of the safe harbor could result in liability for a potentially large number of infringed works.").
[115] *See id.* at 53; Andrew Flanagan, *Vimeo to Launch Music Copyright ID System*, BILLBOARD (May 21, 2014, 10:00 AM), http://www.billboard.com/biz/articles/news/digital-and-mobile/6092241/vimeo-to-launch-music-copyright-id-system-exclusive.
[116] *How Content ID Works*, *supra* note 15.
[117] *Monetization During Content ID Disputes*, YOUTUBE, https://support.google.com/youtube/answer/7000961?hl=en (last visited Apr. 23, 2017).
[118] *See, e.g.*, Gosnell, *supra* note 99.
[119] URBAN, KARAGANIS & SCHOFIELD, *supra* note 8, at 2–3.
[120] *See, e.g.*, Kevin Collier, *Meet RightsCorp, the Anti-Piracy Company That Can Kick You Off the Internet*, THE DAILY DOT (July 30, 2014, 5:00 AM), http://www.dailydot.com/layer8/rightscorp-copyright-disconnect-internet-20-fine/; Michelle Poe, *Wireless Internet Router Exposes Provo Woman To Threats, Lawsuit*, KUTV (Apr. 23, 2014, 9:23 AM), https://web.archive.org/web/20140706070103/http:/www.kutv.com/news/features/gephardt/stories/vid_62.sht

As already mentioned, more famous examples include Paramount's mismatched takedown notice claiming *Anchorman: The Legend of Ron Burgundy* infringed *An Officer and a Gentleman*,[121] NBC's takedown of a political campaign video from President Obama "days before a critical voter registration deadline,"[122] Scripps News Service's removal of a public domain video posted by NASA on NASA's own YouTube page,[123] and the use of the takedown process by a political candidate to shut down another political organization's YouTube account twice.[124] Lesser-known examples can be found throughout the Internet. One website even features "DMCA Horror Stories."[125] Incidents include Sony's removal of a political commentary on YouTube that featured excerpts from Martin Luther King's "I Have a Dream" speech because a Sony-copyrighted song also included excerpts from the speech,[126] and an unknown artist using YouTube's notification process to falsely claim ownership of material in order to receive ad revenue from another party's video.[127]

D. Automated Copyright Enforcement As a Potential Solution

An ACES may reduce the above-described problems of uncertain outcomes, potential government overregulation, and overregulation in the private sector.

For an automated system to assist with the largest number of cases, the parts of the courts' infringements tests currently applied inconsistently must be adjusted for consistent application. The act of establishing more consistent tests is likely to create more certainty in and of itself. In some cases, the automated system could offer complete resolution. In more difficult cases, the system's evaluation may play a lesser role by serving as additional. In either situation, the introduction of an objective evaluation arguably increases the likelihood of more predictable results.

The risk of government overregulation may also be minimized. A properly designed system should regulate strictly within the confines of the law. If this is achieved, and the system is recognized as consistent and reliable, it may help limit how much, if any, additional government regulation is necessary.

---

ml; *Takedown Hall of Shame*, ELECTRONIC FRONTIER FOUNDATION, https://www.eff.org/takedowns (last visited Jan. 16, 2017).

[121] URBAN, KARAGANIS & SCHOFIELD, *supra* note 8, at 91.

[122] *NBC Issues Takedown on Viral Obama Ad*, *supra* note 33.

[123] Higgins, *supra* note 34.

[124] *Attempt to Silence the Political Speech at Right Wing Watch*, *supra* note 35.

[125] *DMCA Horror Stories*, FIGHT FOR THE FUTURE: TAKEDOWN ABUSE, https://www.takedownabuse.org/stories/toc/ (last visited Mar. 30, 2017).

[126] *Sony Takedown over Martin Luther King Speech*, FIGHT FOR THE FUTURE: TAKEDOWN ABUSE, https://www.takedownabuse.org/stories/sony_vs_fight_for_the_future/ (last visited Mar. 30, 2017).

[127] BoyInABargainBin, FIGHT FOR THE FUTURE: TAKEDOWN ABUSE, https://www.takedownabuse.org/stories/royalty_free_claim/ (last visited Mar. 30, 2017).

Overregulation in the private sector appears to be a consequence of systems designed to protect the interests of the users commissioning the software, rather than to reflect the actual legal standard. Though widening the net may help rightsholders catch more infringing content and prevent service providers from hosting as much of it, as described above, this approach has already led to the occasional flagging of non-infringing content. Offering an objective alternative may help clarify the line between what is permitted by the law and what constitutes extra precautionary measures. In the meantime, voluntary private sector solutions like AIRSs and other CRSs seem to be working effectively.

II. Automated CRSs Currently in Use by the Private Sector

One proposed method of creating the automated system described in this paper is through the government's adoption of a modified version of existing systems, since creating such a system from the ground up could take years. The discussion that follows provides a brief explanation of how such systems function and how these systems could be modified to function for the purposes described herein.

A. CRSs Currently In Use

Several examples demonstrate how automated systems are currently being used to identify infringement.

In addition to the electronic systems used by schools to detect student plagiarism,[128] writers have access to a number of free software applications that search for online plagiarism of their work.[129] Online file-sharing website Dropbox uses similar software to digitally fingerprint files and block potentially infringing files before upload.[130] YouTube utilizes its own automated Content ID system to evaluate content for infringement.[131] Many other well-known entertainment companies like Warner Bros, Sony, Disney, and even Facebook license automated content recognition software from a company called Audible Magic.[132]

Note also that YouTube's monetization feature introduced a new method of handling infringing content. Rather than automatically removing infringing videos, YouTube's Content ID system

[128] *Professors Use Technology to Fight Student Cheating*, U.S. NEWS & WORLD REPORT (Oct. 3, 2008, 5:26 PM), https://www.usnews.com/education/articles/2008/10/03/professors-use-technology-to-fight-student-cheating.
[129] *See* Oleksiy Synelnychenko, *How to Protect the Copyright of My Web Content*, IPWatchdog (June 5, 2013), http://www.ipwatchdog.com/2013/06/05/how-to-protect-the-copyright-of-my-web-content/id=40655/ (providing a list of tools to detect online plagiarism).
[130] Sebastian Anthony, *How Dropbox Knows You're a Dirty Pirate, and Why You Shouldn't Use Cloud Storage to Share Copyrighted Files*, EXTREMETECH (Mar. 31, 2014, 2:07 PM), https://www.extremetech.com/computing/179495-how-dropbox-knows-youre-a-dirty-pirate-and-why-you-shouldnt-use-cloud-storage-to-share-copyrighted-files.
[131] *How Content ID Works*, *supra* note 15.
[132] *Audible Magic Accuses YouTube of Fraud Over Content ID Trademark*, TORRENT FREAK (Jan. 11, 2017), https://torrentfreak.com/audible-magic-accuses-youtube-of-fraud-over-content-id-trademark-170111/.

allows qualified users to instruct YouTube to run ads on infringing videos so rightsholders may automatically receive ad revenue.[133]

This type of monetization may have spurred an increase in content creation. Since first launching its Content ID CRS, YouTube "has paid out $2 billion to rightsholders who have chosen to monetize claims" (as of July 2016).[134] A corresponding increase in rightsholders profiting from the monetization of their works and content generation overall has been observed. Over 8,000 YouTube partners, including "major network broadcasters, movie studios and record labels" have claimed "over 400 million videos."[135] In addition, "[a]s of March 2015, creators filming in YouTube Spaces have produced over 10,000 videos which have generated over 1 billion views and 70+ million hours of watchtime."[136] These numbers suggest that YouTube's monetization CRS, a private sector AIRS, may be facilitating the progress of the useful arts.

As popular as content recognition software has become, the companies selling or using it are tight-lipped about exactly how the software works.[137] Due to the limited information available, what follows is a general description of how existing content recognition software works and how it could be modified to serve the purposes recommended in this paper.

B. How Automated CRSs Function

Based on the descriptions that are available, CRSs appear to have the following steps in common:

(1) The submission of copyrighted material for inclusion in the software's database;
(2) The use of a software algorithm to analyze the material for distinctive features and create a "digital fingerprint" of the material based on this analysis;
(3) The addition of the digitally fingerprinted material to the software's database;
(4) The evaluation of potentially infringing material by scanning the new material with the software and comparing it to the material in the software's database; and
(5) The flagging of any matching material based on the parameters programmed into the software.[138]

Slightly restated, an automated system functions by analyzing the work in question and creating a digital "signature" or "fingerprint" of the work.[139] The system then compares this

[133] *How Content ID Works*, *supra* note 15.
[134] *Id.*
[135] *Id.*
[136] *Statistics*, YOUTUBE, https://www.youtube.com/yt/press/statistics.html (last visited Apr. 25, 2017).
[137] URBAN, KARAGANIS & SCHOFIELD, *supra* note 8 at 3 ("Rightsholders expressed concern that publicizing details of their enforcement practices might enable pirates to circumvent them. OSPs worried that revealing their practices could subject them to negative attention from rightsholders, targets, or other OSPs, forcing them to change their practices.").
[138] Strickland, *supra* note 27.

digital fingerprint to an existing database of previously fingerprinted works to see if a match can be found.[140] If the work in question matches an existing work within the database, the new work is flagged as infringing.[141]

Thus, in order to function optimally, an automated system must have an existing database of digitally fingerprinted material against which it can compare potentially infringing works.[142] There are a number of ways in which automated systems can digitally fingerprint works. For instance, the digital fingerprinting of music may be accomplished through several means:

> *Some programs analyze the tempo and beat of a song. Others measure the song's amplitude and frequency. Fingerprinting software usually takes several samples that last just a few seconds each from a single recording. A few companies offer software that analyzes entire audio clips in order to get as complete a fingerprint as possible. At least one current product analyzes a song for landmarks -- distinctive acoustic moments in the clip -- then analyzes the sound around the landmarks. Ideally, the landmarks will be readily identifiable when scanning other music.*[143]

Once the database has been created, the automated system uses an algorithm to determine if a new work matches an existing work in the database "within a certain range of probability."[144] The comparison process has been likened to "the way forensics experts once matched a suspect's fingerprints to those found at a crime scene" in that "[b]efore sophisticated computer software and advanced methods for examining fingerprints became available, experts would look for points of similarity between different fingerprints. In most cases, the specialist would need to demonstrate at least 16 points of similarity for a print to be considered a match."[145] Likewise, content recognition software often allows users to indicate how much of a match must be present in order for a work to be flagged as infringing.[146]

What this process suggests is that a rich database of existing material is crucial to the proper functioning of an automated system. Additionally, the accuracy of the digital fingerprint can be increased based on the depth of the automated system's analysis of a work. Finally, there appears to be some flexibility in what features an automated system will recognize as a match, and the threshold at which a system will flag a work for infringement.

As a result, while automated systems may currently be programmed to suit the needs of the private sector clients commissioning their use, it is possible that the systems could be modified

---

[139] *Id.*
[140] *See id.*
[141] *See id.*
[142] *See id.*
[143] *Id.*
[144] *Id.*
[145] *Id.*
[146] *See id.*

to match the legal standard for infringement more closely. Such a modification could create a tool for use by the courts and the general public to identify actionable infringement. Even in cases where the system's evaluation is not conclusive and human evaluation is required, an automated system can provide useful evidence that encourages a more fair and predictable outcome.

C. The Limitations and Drawbacks of Automated CRSs

The Limits of Filtering responds to calls by many "policymakers and copyright industry lobbyists" advocating a requirement that online service providers implement CRSs in order to receive DMCA safe harbor protections.[147] The report engages in a detailed discussion of the current limits and drawbacks of CRSs, ultimately concluding that "functionality and inherent limitations of the most common filtering technologies [] demonstrate why a mandatory filtering regime would pose grave dangers to the viability of the internet ecosystem in exchange for [] minimal effects on online infringement."[148]

Although the report correctly notes "such technologies are not sufficient to consistently identify *infringements* with accuracy," it is important to recognize that at some point—perhaps sooner rather than later—it may be possible to more closely align objective legal standards of copyright infringement evaluation with more robust and accurate CRSs.[149] Thus, the next sections will focus on what such a system, if it were to materialize, might look like and how it might come about.

III. Potential Uses for an ACES

The use of automated systems for infringement evaluation could solve a number of problems with the current copyright enforcement system. Owners of existing copyrights would be better equipped to protect their ownership rights, as automated systems could affordably identify infringement and encourage out-of-court settlements without costly legal battles. New content creators could avoid infringing before releasing material, leading to more certainty and protection for their copyrights. The identification of potential infringement may even facilitate the licensing of existing copyrighted material and the creation of innovative new works.

While the development of ACESs may be a task best suited for private organizations, the government can play a significant role by encouraging development of these systems and court recognition of their validity. We may already be heading in this direction, as at least one court has suggested automation may serve as a legitimate means of fair use evaluation for DMCA

[147] ENGSTROM & FEAMSTER, *supra* note 18, at i.
[148] *Id.*
[149] *Id.* at 21 (emphasis in original).

takedown notices.[150] Expanding the role of automation to include infringement evaluations designed for use in courts could be the next step.

One possible use of an ACES is for private individuals to determine if infringement has occurred. For content creators, this may provide a more reliable method to determine if one of their new works has accidentally infringed on an existing work. If actionable infringement is found, the content creator may either contact the copyright owner to negotiate a license, or modify the new work so that it is no longer infringing. Both solutions help reduce potential liability for accidental infringement.

For copyright owners, an ACES that reflects the current legal standard could provide a more accurate assessment of whether another content creator has committed actionable infringement. If infringement has been found by the system, the copyright owner may be in a better position to contact the alleged infringer and negotiate a license without filing suit. If an agreement cannot be reached, the automated system's findings may place the copyright owner in a stronger legal position.

For both copyright owners and creators of new content, a properly functioning ACES offers a more certain and protectable set of property rights. A finding of actionable infringement generated by the automated system gives a copyright owner stronger evidence with which to bring suit. Conversely, the absence of an infringement finding by the automated system may provide a strong defense against an infringement claim.

An ACES may also facilitate the licensing of existing material by new content creators. Making it easier to find copyright owners and negotiate licenses could encourage progress and innovation, and allow copyright owners to maximize the value of their intellectual property rights.[151]

A well-designed ACES could also serve as a useful tool for the courts. If programmed correctly, such a system may provide a more objective standard of evaluation. To function correctly, the automated system must be programmed to meet the current standard utilized by the relevant court for a finding of infringement. It may also be necessary to identify inconsistencies in the courts' applications of its tests and adjust these to achieve a more consistent application. Once the system has been programmed, in theory, it should operate in the same manner each time it

---

[150] *See* Lenz v. Universal Music Corp., 801 F.3d 1126, 1135 (9th Cir. 2015), *opinion amended and superseded on denial of reh'g*, 815 F.3d 1145 (9th Cir. 2016), *cert. denied,* 137 S. Ct. 416, 196 L. Ed. 2d 293 (2016) ("consideration of fair use may be sufficient if copyright holders utilize computer programs that automatically identify for takedown notifications content . . . ."). Note, however, that on review, the appellate court removed references to the use of automated systems in its amended opinion. Ashley Cullins, *"Dancing Baby" Appeals Court Decision Stands Minus the "Fair Use" Algorithms*, THE HOLLYWOOD REPORTER (Mar. 17, 2016 2:20 PM), http://www.hollywoodreporter.com/thr-esq/dancing-baby-appeals-court-decision-876557.

[151] THE DEPARTMENT OF COMMERCE'S INTERNET POLICY TASK FORCE, GREEN PAPER ON COPYRIGHT POLICY, CREATIVITY, AND INNOVATION IN THE DIGITAL ECONOMY, 87–89 (2013) [hereinafter GREEN PAPER ON COPYRIGHT POLICY], https://www.uspto.gov/sites/default/files/news/publications/copyrightgreenpaper.pdf.

makes a determination, leading to more consistent and predictable results. As previously mentioned, an objective evaluation of whether infringement has occurred is one of the prongs of both the Second and Ninth Circuit substantial similarity tests.

It is important to remember that some copyright controversies may not be solved without human evaluation. In these cases, an automated system cannot and should not be viewed as the sole means of evaluation. But neither should it be entirely discounted, as it may serve as useful, objective evidence.

Keeping all this in mind, an ACES, if programmed to function consistently and reliably, could assist courts in making its objective determinations. This may diminish the costs associated with long, drawn-out suits, and could be utilized to more fairly dismiss cases under summary judgment, or, when appropriate, to move a case forward to a jury.

IV. Pathways to Potential Implementation

A number of hurdles will need to be overcome before an effective automated copyright enforcement system could be actualized. Even if such a system were within reach, there are a number of issues that may make its implementation not only difficult, but arguably unwise. Nonetheless, this is a technology that will continue to mature; and as it does, policymakers will increasingly be faced with calls to reform the current copyright enforcement system. Under the right political circumstances, an ACES may end up being a policy option worth considering, with some important caveats.

In the interest of considering a hypothetical future scenario, the discussion that follows describes the prerequisites for the emergence of such a system.

A. Creating the Software

Several significant challenges stand in the way of using existing software for the automated system described here.

While the private sector has developed automated systems to detect copyright infringement, as described above, many of the systems do not use an entire digital fingerprint to conduct comparisons, preventing a complete comparison of songs. Aggravating the problem, there are several Internet sites that provide "workarounds" to avoid getting flagged for infringement by automated systems. One site recommends simply changing the pitch of an audio sample.[152] Another suggests various alterations to an audio file, such as pitch, timing, and length, which

[152] Ted, *How to Get Around That Pesky Copyrighted-Audio Filter on YouTube and Facebook*, TURKEY MONKEY (July 19, 2009), http://www.turkeymonkey.com/2009/07/19/how-to-get-around-that-pesky-copyrighted-audio-filter-on-youtube-and-facebook/.

can "thwart" YouTube's Content ID system.[153] At least one YouTube user has posted video instructions suggesting that adding a different song to the beginning and end of a song you wish to post may also bypass YouTube's Content ID system.[154]

Thus, while it may be possible to reprogram existing software to require the use of complete digital fingerprints for comparisons, this approach may be technically infeasible at the present time.

In addition to the ease of bypassing existing systems, processing speed and storage requirements may be limiting factors. In 2011, the International Society for Music Information Retrieval released *Echoprint – An Open Music Identification Service*, a white paper describing the function of Echoprint, the open-source CRS software currently used by Spotify.[155] Echoprint's database at the time consisted of over 30 million tracks.[156] In terms of storage space, "[t]he index takes roughly 5 gigabytes of disk space per 100,000 tracks, while the storage requires 15 gigabytes."[157] Regarding processing time "[q]ueries can be executed in practice in 100 milliseconds after computing the code string for a signal, which on current hardware takes under a tenth of a second."[158] Keeping in mind that Echoprint is limited to audio files and only evaluates a portion of information about each audio file, it is unknown how much more storage space and processing speed would be required to create and store more complete digital fingerprints. It is also unclear where and how the system would be hosted.

Finally, while adjusting the system so it more closely reflects the relevant legal standard may one day be within the reach of software developers, for the time being, the many varied types of content that would need to be accurately fingerprinted limits this technology's efficacy in providing accurate matches for content beyond audio and video. As *The Limits of Filtering* notes, no fingerprinting tool currently exists to evaluate software, encrypted files, architectural designs, or handmade items sold online.[159] Even for audio and video, matching is notoriously difficult.

Until these challenges have been overcome, the available software cannot be considered suitable for the ACES this paper envisions.

---

[153] Scott Smitelli, *Fun with YouTube's Audio Content ID System*, SCOTT SMITELLI (Apr. 19, 2009), http://www.scottsmitelli.com/articles/youtube-audio-content-id.

[154] 4 Cracks n' Tuts, *How to Bypass WMG YouTube Copyright Audio MP3 Music Files Easy*, YOUTUBE (Aug. 3, 2012), https://www.youtube.com/watch?v=Pm_cpeJyVX0.

[155] *See* DANIEL P.W. ELLIS ET AL., ECHOPRINT - AN OPEN MUSIC IDENTIFICATION SERVICE (2011), http://ismir2011.ismir.net/latebreaking/LB-7.pdf. The Echoprint white paper is discussed at length by *The Limits of Filtering. See generally* ENGSTROM & FEAMSTER, *supra* note 18.

[156] ELLIS ET AL., *supra* note 155, at 2.

[157] *Id.*

[158] *Id.*

[159] ENGSTROM & FEAMSTER, *supra* note 18, at 2, 14.

B. Developing a Digital Library

As discussed above, an ACES cannot function optimally without a rich database of accurately fingerprinted digital works. Fortunately, such a database may be within reach. A merger of works that have already been digitally submitted to the Copyright Office and the works that have been submitted to private sector databases such as YouTube's Content ID system could serve as an excellent starting point.

There may also be government support for the creation of this database. In July 2013, the Department of Commerce's Internet Policy Task Force released its *Green Paper on Copyright Policy, Creativity, and Innovation in the Digital Economy*.[160] In its report, the Task Force emphasized the importance of maintaining an accurate database of ownership information, specifically for the purposes of licensing:

> *The most basic prerequisite for obtaining licenses is reliable, up-to-date information about who owns what rights in what territories. Users need to find the right holders from whom to obtain permission, and right holders or their representatives need to be contacted to determine terms of use. As online businesses seek licenses for large repertoires of works to be offered in multiple countries in a variety of formats, and as multimedia uses become more common, the need for comprehensive globally-linked databases is growing.*[161]

The House Judiciary Committee appears to share this sentiment. House Judiciary Committee Chairman Bob Goodlatte (R-Va.) and Ranking Member John Conyers, Jr. (D-Mich.) recently issued a press release describing the "first policy proposal to come out of the Committee's review of U.S. Copyright law," which stated:

> *Today, we are releasing our first policy proposal, which identifies reforms to modernize the Copyright Office so that it can meet the challenges of the 21st Century. Among the reforms in this document are granting the Copyright Office autonomy with respect to the Library of Congress, requiring the Copyright Office to maintain an up-to-date digital, searchable database of all copyrighted works and associated copyright ownership information, and many others reforms.*[162]

These reports suggest that the government has recognized the value of a searchable digital database of copyrighted works and intends to build this database to meet the needs of digital users.

---

[160] GREEN PAPER ON COPYRIGHT POLICY *supra* note 151.
[161] *Id.* at 89.
[162] Press Release, House Judiciary Committee Chairman Bob Goodlatte, Goodlatte & Conyers Release First Policy Proposal of Copyright Review (Dec. 8, 2016) (on file with author).

C. Clarifying the Standard Utilized by Courts for Infringement

It is important to note that the inconsistent standards used by courts to evaluate infringement may make it difficult to program consistent parameters into an ACES. However, this should not operate as a complete bar to the use of an automated system. Rather, the most consistent elements of the courts' standards of evaluation should be identified, the automated system should be used with the expectation that it will not completely replace all human evaluation, and any clearly inconsistent standards should either be modified or, in the alternative, acknowledged and treated as known exclusions to the automated system.

In cases where a 100% match of infringement has been found, the reliability of an ACES is unlikely to be challenged. Rather, the burden may shift to the alleged infringer to demonstrate that the use was authorized, the plaintiff does not have ownership rights, or the use was permitted under a different defense.

It is in cases where a lesser percentage of potential infringement has been detected that the validity of an automated system may be questioned. This may occur in varying degrees and for a variety of reasons. Thus, it seems prudent to examine some of the reasons why minor dissimilarities may occur and adjust the role of the automated system's evaluation accordingly.

Dissimilarities Due to Purposeful Alteration to Avoid Detection

As previously mentioned, content may be modified for the specific purpose of bypassing detection by automated systems. Examples include: altering the pitch of a song;[163] changing the tempo and speed of a song;[164] or limiting snippets of copyrighted video content to less than a few seconds to avoid getting "red-flagged."[165]

In such cases, the content is likely infringing despite these minor modifications. A reasonable solution may be to program the automated system so that it does not judge whether the content is infringing, but, rather, it reports what percentage of similarity between the two works exists.

[163] Ted, *supra,* note 152.
[164] Smitelli, *supra* note 153.
[165] Stephen McArthur, *How to Beat a YouTube Content ID Copyright Claim – What Every Gamer and MCN Should Know*, GAMASUTRA (June 24, 2014, 4:41 PM), http://www.gamasutra.com/blogs/StephenMcArthur/20140624/219589/How_to_Beat_a_YouTube_ContentID_Copyright_Claim__What_every_Gamer_and_MCN_Should_Know.php.

Dissimilarities Due to Partial Appropriation

Some forms of appropriation include the utilization of small snippets of copyrighted content (e.g., sampling),[166] or the copying and incorporation of copyrighted elements or themes to create new works (e.g., derivative works).[167] Such appropriation may be difficult for an automated system to identify, or, conversely, too frequently misidentified if the threshold for identification is set too low.

Some of this may be addressed through programming. Setting the automated system to identify similar or identical elements that compose a significant portion of the work could be considered enough for a court to take notice and for parties to consider action. For example, perhaps the automated system could be set to flag three seconds of the same chord progression or melody between two songs, or 5 percent of a work of art that matches 5 percent of another work. In such cases, the similarities would be identified and noted by the software, but users should anticipate the additional requirement of human evaluation.

Similarities in Effect Without Identical or Similar Elements

The recent *Blurred Lines* case introduced an arguably new theory of copyright infringement.[168] Under the district court's holding, it has been said that music copyright infringement may be found when the style and feel of a song is similar to a previous work, even in the absence of identical chord progressions, melodies, lyrics, or other musical elements.[169] Such a standard is unlikely to be programmed easily into an ACES, or, if programmable, may result in a number of false positives.

Two solutions are available for this situation. The first is to disregard the use of an automated system of evaluation in such cases, as it is likely to be outside the capabilities of the software. The second is to alter the court's standard of evaluation, since, in its current state, this theory of infringement seems to be more subjective than objective. This is contrary to the substantial similarity tests used by the courts, which require both an objective and subjective finding of infringement.

---

[166] "Sampling" is defined as "the act of using a small part of a recording (such as a song) as part of another recording." *Sampling*, MERRIAM-WEBSTER, http://www.merriam-webster.com/dictionary/sampling (last visited Jan. 14, 2017).

[167] 17 U.S.C. § 101 (2011), https://www.copyright.gov/title17/92chap1.html#101 ("A 'derivative work' is a work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted. A work consisting of editorial revisions, annotations, elaborations, or other modifications, which, as a whole, represent an original work of authorship, is a 'derivative work.'").

[168] *See generally* Zernay, *supra* note 68.

[169] *Id.*

It may be too much to ask the courts to adjust their standard of evaluation. Such a course of action may also take years to achieve. Rather, it seems wiser to exclude the use of an automated system in cases such as these, and limit its application to controversies involving more easily proven forms of infringement.

D. The Role of the Courts

At this time, courts have not fully embraced automation as a tool for infringement evaluation. Recent rulings suggest courts may be acknowledging the potential role of automation, but there still appears to be some hesitancy. The district court's holding for *Lenz v. Universal Music Corp*. indicated automation could be used to perform the fair use analysis required prior to sending a DMCA takedown notice.[170] However, on review, the appellate court "cut out several paragraphs concerning how automation might be leaned upon by companies like Universal that have to evaluate fair use on a mass scale."[171]

Thus, automated infringement analysis may still be considered new to the courts. This could be an obstacle to admissibility. Under the Federal Rules of Evidence:

> *The United States Supreme Court has declared that the trial court has a gatekeeping obligation to determine whether the explanative theory underlying every expert witness' testimony, regardless of whether based on scientific, technical or other specialized knowledge, is "reliable." . . . Federal Rule of Evidence 702 was amended effective December 1, 2000 . . . requiring that the trial judge before permitting an expert to testify determine that "(1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case." As restyled, Rule 702 now provides that the expert witnesses testimony must be shown to be "the product of reliable principles and methods", Rule 702(c), and that it must also [have] been shown that "the expert has reliably applied the principles and methods for the facts of the case," Rule 702(d).*[172]

Until analyses by ACESs are recognized as commonly accepted and reliable forms of evidence, admissibility of such evidence may require close review by courts.[173]

---

[170] Lenz v. Universal Music Corp., 801 F.3d 1126, 1135 (9th Cir. 2015), *opinion amended and superseded on denial of reh'g*, 815 F.3d 1145 (9th Cir. 2016), *cert. denied,* 137 S. Ct. 416, 196 L. Ed. 2d 293 (2016); Cullins, *supra* note 150.

[171] *Id.*

[172] *Reliability "Gatekeeping" Under Daubert/Kumho/Rule 702: Historical Development and Assessment*, 5 HANDBOOK OF FED. EVID. § 702:5 (7th ed. 2016).

[173] The process of establishing automated systems as a generally accepted method of infringement evaluation by courts requires a detailed and involved discussion of its own. For the purposes of this paper, it may be enough simply to state that general acceptance of any new method is not automatic, and until general acceptance is

Fortunately, these are not impossible obstacles. So long as the software is properly designed, refined to ensure accurate performance, and frequently updated to encompass the largest amount of copyrighted material, over time, automation is likely to become a generally accepted form of evaluation and admissibility will become less of an issue.

V. Potential Obstacles

In addition to the concerns raised above, a number of issues may need to be addressed before implementing an ACES becomes politically and economically feasible. In this final section, we will discuss some of the potential difficulties that may interfere with immediate implementation.

A. The Risk of Misuse

Much has been said about the misuse of the current DMCA notice and takedown process. It is wise to acknowledge that a similar version of this behavior may appear if an ACES is implemented.

Rather than allowing this possibility to exclude the use of automated systems altogether, solutions have been proposed by others for current systems that may be equally applicable to the ACES proposed here. A report on a recent set of studies commented that:

> *The increased use of automated systems by large rightsholders, as well as [OSPs], raised questions of accuracy and due process. Though rightsholders and OSPs generally use some accuracy checks today, we identified a clear need for better mechanisms to check the accuracy of algorithms, more consistent human review, and a willingness by both rightsholders and OSPs to develop the capacity to identify and reject inappropriate takedown requests.*[174]

These studies do not seem to indicate that automation should be ruled out altogether. Instead, proposed solutions include ensuring the accuracy of algorithms, incorporating more human involvement in the evaluation process, and providing a better system of responding to false positives.

These solutions seem equally applicable to the ACES discussed here. During the initial implementation period, it will be crucial to evaluate the system's performance closely and refine its programming in response to issues that will undoubtedly arise. As copyright law fluctuates, it will be necessary to adjust the automated system accordingly. There are situations

---

achieved, the reliability of the method will likely require a case-by-case review when courts consider its admissibility.

[174] URBAN, KARAGANIS & SCHOFIELD, *supra* note 8, at 3.

where human evaluation cannot be replaced by an automated system, and in such cases, the automated system must be treated as a secondary tool, or considered inapplicable. Remaining mindful of the limitations of an ACES will allow courts and the general public to maximize its usefulness and accuracy.

B. Support from the Senate and House Judiciary Committees

Actually implementing such a system would likely be too expensive for the government to justify building its own database and automated system from the ground up. It may be preferable to license the use of existing CRSs and the richest database(s) of digitally fingerprinted content, then modify the licensed components to match the needs of the courts and the general public. This would likely be more cost-effective and efficient than the alternative of building an entirely new automated system and corresponding digital database.

The Department of Commerce's Internet Policy Task Force *Green Paper on Copyright Policy, Creativity, and Innovation in the Digital Economy* expressed the Task Force's support of funding for a U.S. public database, stating "it is important to ensure that the Copyright Office has sufficient resources" to develop the database.[175]

YouTube's Content ID system appears to be a strong candidate for licensing. Its database has existed since 2006 and may very well be updated frequently, since YouTube is under constant internal and external scrutiny. YouTube describes its Content ID database as "the most comprehensive in the world," with "more than 50 million active reference files."[176] As one of the largest repositories of online content, YouTube likely has a rich database of digitally fingerprinted material.

Audible Magic may also be an excellent option, as it is the company behind the automated content recognition software utilized by Shazam, Warner Bros, Sony, Disney, and Facebook.[177] Additionally, from 2006 to 2009, YouTube licensed software from Audible Magic for its Content ID system.[178] Other equivalent options may exist, and are worth exploring.

While licensing an automated system and modifying it to suit the purposes described here may be less expensive than building the system from the ground up, it may still be costly. These expenses are likely to require approval from the Senate and House Judiciary Committees, thus, the support of these Committees is critical.

The expense could be offset by the reduced burden on courts, speedier resolution of infringement cases through the use of the ACES, and perhaps even increased copyright

[175] GREEN PAPER ON COPYRIGHT POLICY, *supra* note, at 151.
[176] *Statistics*, YOUTUBE, https://www.youtube.com/yt/press/statistics.html (last visited Mar. 18, 2017).
[177] *Audible Magic Accuses YouTube of Fraud over Content ID Trademark*, *supra* note 132.
[178] *Id.*

registration income, as more content creators will realize they have a valuable and more readily enforceable set of rights.

C. Technical Limitations

The Limits of Filtering rallies against recent calls for DMCA reform that would require OSPs to implement CRSs in order to qualify for DMCA safe harbor protections:

> *Ultimately, conditioning access to the safe harbor on an OSP implementing a filtering technology would undermine the certainty that the DMCA has created without a commensurate impact on infringement. Policymakers should resist calls to upend the DMCA on the illusory hopes that filtering technologies can replace the commonsense copyright regime that has allowed the internet to flourish.*[179]

While many of the concerns raised in *The Limits of Filtering* primarily speak to this issue, there are several limitations that also apply to the automated system proposed here.

The Limits of Filtering warns that "all filtering tools can be evaded through basic manipulation of the file."[180] Though fingerprinting systems may be less susceptible to such manipulation, they are not impervious, as the encryption of files "stymies even these more sophisticated fingerprinting techniques."[181]

It is also important to note that copying content is not always considered infringement:

> *[E]ven if it were possible to consistently identify content accurately with filtering technologies, such technologies are not sufficient to consistently identify infringements with accuracy, as they can only indicate whether a file's contents match protected content, not whether a particular use of an identified file is an infringement in light of the context within which the media was being used.*[182]

This touches on the fair use argument raised in *Lenz v. Universal Music Corp.*,[183] and lends support to the suggestion that an automated system cannot and should not completely replace all human evaluation of copyright controversies.

The Limits of Filtering also notes that OSPs may not have access to a complete database of existing material.[184] This is a significant problem under the mandatory filtering regime *The*

[179] ENGSTROM & FEAMSTER, *supra* note 18, at ii.
[180] *Id.* at 17–18.
[181] *Id.* at 18.
[182] *Id.*
[183] *See* Lenz v. Universal Music Corp., 815 F.3d 1145 (9th Cir. 2015), *cert. denied*, 137 S. Ct. 416 (2016).
[184] ENGSTROM & FEAMSTER, *supra* note 18, at 18–20.

Limits of Filtering advocates against; an incomplete database decreases the accuracy of an OSP's filtering system, subjecting the OSP to the potential loss of its DMCA safe harbor protections due to circumstances beyond its control. Theoretically, the ACES described here would be created using well-populated, pre-existing databases, making this problem less pronounced. However, it is an important issue, as any CRS requires a rich and complete database of accurately fingerprinted material to function optimally.

Finally, *The Limits of Filtering* highlights the fact that "[a]s encoding and compression standards continue to evolve, the fingerprinting technologies themselves must also develop to keep pace with standards."[185] The paper correctly observes that "standards must continue to evolve to support the proliferation of digital content on the Internet."[186] In order to maintain accuracy, CRSs must be frequently updated in response. Failure to do so creates a high risk that CRSs will fall behind technology and cease to produce reliable evaluations.

Unfortunately, overcoming these and other limitations shifts the likelihood of implementing a government-adopted ACES further into the future. However, by recognizing the benefits of such a system and identifying it as a goal, the government and its related agencies may begin to take actions that encourage more rapid evolution of CRSs in the direction of achieving an effective version of an ACES for general use.

VI. Policy Recommendations

As discussed in previous sections, the current environment seems to be functioning well.

What follows are policy recommendations for legislators and regulators looking at potentially using this technology as a copyright management tool, now or in the future.

A. Refrain from Conditioning Safe Harbor Protections on the Use of CRSs

At this time, the current copyright regime, warts and all, has functioned in a reasonably effective manner. Content creation is at an all-time high and intermediary liability protections afforded by the DMCA are robust enough to balance the interests of OSPs and rightsholders. While there are certainly issues that still need to be ironed out, such as the legal uncertainties surrounding the various approaches to determining what constitutes infringement, the status quo is enabling and spurring creative creations and innovation. Though there may be means of improving upon the status quo, the largest single concern at this time is a deterioration of the current environment. As a result, it is imperative that policymakers refrain from advocating for additional conditionality set on the safe harbor provisions of the DMCA. In particular, safe harbor provisions for OSPs should not be conditioned on the use of CRSs, automated or otherwise.

[185] *Id.* at 21.
[186] *Id.*

Because so many of the fingerprinting tools currently in use for automated copyright detection purposes are designed to operate on single media types, no one overarching solution for automated copyright management is currently feasible. The technical limitations of filtering technologies, the easily circumvented nature of many CRSs, and the "content bloom" engendered by the proliferation of digital communications technologies are all significant reasons to refrain from considering rules that would mandate automated content filtering as a precondition for safe harbor protections under the DMCA.

The previously cited report from Engine describes how such a mandate would undermine the certainty of the DMCA's safe harbor provisions:

> *A safe harbor is only useful for mitigating uncertainty if the prerequisites for obtaining its protections are clear, but it is decidedly unclear how policymakers could precisely define what sorts of filtering technologies would be deemed adequate to claim the safe harbor. Any filtering requirement would either have to endorse a particular technology (and quickly become outdated) or establish a "reasonableness" standard that would require clarification from courts before [OSPs] could have any confidence in their protection under safe harbor.*[187]

That diminished certainty, in turn, would have imposed significant costs on the health and vitality of the Internet ecosystem.

The report goes on to note that, "[t]he negative impact a proactive filtering mandate would have on [OSPs]—and the concomitant decrease in creative output from those who rely on the [I]nternet as a medium of production and distribution—significantly outweighs its benefits."[188] And it concludes:

> *[I]f the DMCA's safe harbor protections were predicated on an [OSP] implementing a "reasonable" content filtering system, the uncertainty surrounding any [OSP's] eligibility for the safe harbor would undermine one of the central purposes of the DMCA: providing "more certainty . . . in order to [allow OSPs to] attract the substantial investments necessary to continue the expansion and upgrading of the Internet."*[189]

In short, voluntarily implemented AIRSs, as previously discussed, can be a valuable tool for OSPs seeking to address concerns over copyright infringement. Mandating their use, or the use of other CRSs is an unnecessary and potentially unwise approach to remedying online copyright infringement concerns. Given the fragile order that has emerged online in the post-DMCA

[187] ENGSTROM & FEAMSTER, *supra* note 18, at 28.
[188] *Id.* at 25.
[189] *Id.* at 30.

Internet ecosystem, it would be better for policymakers to refrain from actions that could potentially undermine the current status quo.

B. Avoid One-Size-Fits-All Solutions

The Limits of Filtering highlights a troublesome shortcoming in current CRSs—most only evaluate one or two forms of content:

> *Because fingerprinting technologies rely on algorithms that process the underlying media content of a given file, they are naturally constrained to a small subset of copyrightable content. For example, because an audio fingerprinting tool's central algorithm examines, say, the frequency values in a song file, it cannot be used to identify copyrighted photographs or software programs, which contain no audio frequency values. As such, to filter all files on a given site, an [OSP] would need to obtain a different fingerprinting tool for each type of media that is (or could be) hosted across the entire site. Considering the incredibly wide scope of copyrightable content, there are many types of content for which no fingerprinting tool exists, such as architectural designs or handmade items sold on Etsy.*[190]

Thus, Echoprint and Shazam only evaluate audio. Anti-plagiarism software used by schools is likely limited to the written word. YouTube's Content ID system reviews audio and video. And, as *The Limits of Filtering* has stated, CRSs may not yet exist for other forms of copyrightable content.

Therefore, a one-size-fits-all solution is not only inappropriate, it is currently non-existent. Whether the development of a government-adopted system should be stalled until an all-encompassing system is created is debatable. If an extremely stable and reliable CRS emerged that specializes in one form of copyrightable content, such as music, it may be worth considering its adoption for public use while we wait for CRSs to develop for other forms of content. Perhaps this approach is more effective, as it allows for the adoption of several highly specialized CRSs over time, rather than settling for a single CRS that may operate as a jack-of-all trades, but a master of none.

C. Facilitate Open Access to Databases of Copyrighted Works

Where legislators can have a positive impact in the application of automated CRSs, their eyes should focus on promoting tools that can positively contribute to growing the ecosystem of online content production.

[190] *Id.* at 16.

Incentivizing rightsholders to contribute accurate and up to date content identifiers enriches an ACES's database, and a rich database is a key component to ensuring the efficiency and accuracy of matching content. Additionally, mechanisms for preventing uploads of false fingerprints will be necessary to ensure the database is able to provide accuracy and certainty among users who depend on its smooth operation.

D. Broader Copyright Office Reform

Over time, a number of broad reforms to the Copyright Office may be necessary to optimize the performance of the copyright enforcement system in general.

As noted at the start of this paper, the long term of copyright duration has created a much higher risk of infringement. This may be due in part to the limited ability of content creators to come up with new and innovative content. Some analysts attribute the lessened ability of modern musicians to come up with more sophisticated compositions to a "decline in musical literacy."[191] The limited set of artistic content that the general public finds appealing may also be a contributing factor. As noted in *Darrell v. Joe Morris Music Co.*, "[i]t must be remembered that, while there are an enormous number of possible permutations of the musical notes of the scale, only a few are pleasing; and much fewer still suit the infantile demands of the popular ear."[192] A shorter term of copyright duration may help address these issues, as more works will be converted to the public domain and become available for content creators to utilize.

The Copyright Office appears to have reform plans. In a recent interview with *The Hollywood Reporter,* Michelle Lee, then-Under Secretary of Commerce for Intellectual Property and Director of the United States Patent and Trademark Office (USPTO) discussed copyright system reform:

> *Our copyright system is really designed to play a critical role to promote and incentivize the creators of artistic works, while also promoting their dissemination and use. The Internet has significantly transformed how copyright policy and copyright works are being distributed nowadays. Computers can make an identical copy instantly. There is a tension between promoting and incentivizing that innovation and permitting free dissemination and fair dissemination. The two have to work in tandem in order for us to have a system that works.*
>
> *The Copyright Act hadn't been changed for a really long time. No one had studied it since the Clinton Administration. This was the first comprehensive look at copyright issues related to the digital economy. The USPTO took the lead, and the National Telecommunications and Information Administration both worked on it,*

[191] Charles Cronin, *I Hear America Suing: Music Copyright Infringement in the Era of Electronic Sound*, 66 HASTINGS L.J. 1187, 1206, 1224 (2015).
[192] Darrell v. Joe Morris Music Co., 113 F.2d 80, 80 (2d Cir. 1940).

> *and we came up with a white paper. It's on remixes, first sale doctrine and statutory damages. We looked at everything and got a lot of stakeholder input. We then produced the white paper that focused on those three topics. . . .*
>
> *It's just the beginning because, as with any piece of legislation, there's still lots of work to be done and a lot of stakeholder discussions to be had in terms of best practices and so forth. Of course, in the background, Chairman Goodlatte and others in the House are looking at pretty comprehensive reviews of copyright issues and they've been having hearings.*[193]

There has even been discussion of creating a separate court exclusively to handle copyright infringement controversies. Ms. Lee stated:

> *We also noted the benefits of a new forum to adjudicate small copyright infringement claims below a certain dollar amount not as encumbered by district court litigation with all the costs and all the discovery to provide an alternative avenue to adjudicate small [dollar] amount copyright infringement cases. So those were concrete legislative recommendations that we made to Congress.*[194]

Finally, as content is more widely created and distributed, the application of the fair use doctrine may be refined. The *Cariou v. Prince* case highlights the current uncertainty about the role of intent in a fair use analysis. Both *Cariou* and *Bill Graham Archives v. Dorling Kindersley Ltd.* raise questions about what constitutes sufficient transformation to qualify as fair use. *Lenz v. Universal Music Corp* will likely increase the role of the fair use analysis in the DMCA notice and takedown process.

Conclusion

Updating the copyright enforcement system has become an important priority, as the current system does not address the exponential growth of digital content creation and online distribution tools. Furthermore, the system of self-policing that has resulted from the DMCA notice and takedown procedure has evolved into a form of overregulation that fails to apply the law correctly. In light of these issues, a government-adopted ACES may offer an ideal solution. By licensing an existing automated system, modifying it to match the current legal standards, and factoring in human supervision for the more difficult infringement controversies, an ACES may help modernize the current copyright system and more fairly and accurately reflect the law.

---

[193] Ashley Cullins, *President Obama's IP Advisor Talks Reform Efforts and a New Administration*, THE HOLLYWOOD REPORTER (Dec. 23, 2016 7:30 AM), http://www.hollywoodreporter.com/thr-esq/president-obamas-ip-advisor-talks-reform-efforts-a-new-administration-936404.

[194] *Id.*

In spite of these potential benefits, the limitations and drawbacks of existing CRSs may place the achievement of a government-adopted CRS in the distant future. Many challenges must be overcome before an effective ACES can be implemented. In light of the effectiveness of the DMCA's current safe harbor procedures, augmented by the proliferation of voluntarily-adopted AIRSs, for the time being, maintaining the current DMCA framework and promoting CRSs and voluntary solutions in support of this framework may be the most sound method of ensuring the continued progress of the useful arts.

Despite the present technological confines, the time to start considering the implications of an automated system is now, as this goal may help influence the evolution of existing systems and lead to a system that better balances the interests of all involved.