

Response to “Notice and Takedown in Everyday Practice: A Review”

Jennifer M. Urban, Joe Karaganis, and Brianna L. Schofield

We are the authors of the research report, *Notice and Takedown in Everyday Practice*.¹ This work combines qualitative and quantitative work on section 512 of the Digital Millennium Copyright Act’s notice and takedown process, a central mechanism for copyright enforcement online, and its “safe harbors,” which protect online intermediaries from liability for their users’ infringement if they cooperate to remove infringing material from their services and take certain other measures.

We submitted *Notice and Takedown in Everyday Practice* to the US Copyright Office in response to its request for empirical research on the effectiveness of section 512. Our report was one of only a handful of empirical research submissions. Though the MPAA and the RIAA did not submit studies of their own, they did commission a critique of our report submitted by Dr. George S. Ford, an economist at Phoenix Center for Advanced Legal and Economic Public Policy Studies.²

As academics, we welcome critical engagement with our work. Empirical research on section 512 is scarce, and there is still much to be learned about how the law is applied and understood by different parties. Unfortunately, Ford’s effort does not advance that conversation. Instead, he repeatedly mischaracterizes our report and its findings and offers alternative conclusions that are unsupported either by our work or by his own evidence. Some of his criticisms appear to stem from careless or motivated misreading of our work. Others reflect poor formal modeling. Still others appear to reflect misunderstanding of the purposes of the statute and the choices Congress made when passing it.

Our research, findings, and recommendations

Our report is comprised of three linked studies:

The first (“Study 1”) is a primarily descriptive and qualitative account of the ways in which the notice and takedown process has been perceived and operationalized by some of its major stakeholders. It is based on interviews and surveys with nearly three dozen U.S. online service providers (“OSPs”) and large rightsholders, documenting the ways in which the notice and takedown process has been perceived and operationalized.

The second (“Study 2”) is a quantitative analysis of takedown notices, taken from a set of over 108 million requests submitted to the Lumen³ archive over a six-month period (most of which relate to Google Web Search).

The third (“Study 3”) is a quantitative analysis of the subset of those 108 million takedown requests sent to Google Image Search. These requests were sent primarily by individuals and small businesses—an important category for policy makers that tends to get lost in discussions dominated by large rights holder organizations and intermediaries.

Study 1 provides a detailed descriptive account of how section 512 operates in practice for OSPs and large rightsholders today. Findings include:

- Detailed descriptions of how notice and takedown is practiced by respondents, including day-to-day practices, how rightsholders strategically target their enforcement activities in the face of large-scale infringement, how participants make decisions in edge cases, and the like;
- All OSPs depend heavily on the safe harbors for freedom to operate;
- Both OSPs and rightsholders consider notice and takedown to be a central tool for online copyright enforcement;
- Large-scale infringement has increased the scale of notice and takedown exponentially and in a manner that is frustrating and costly for some rightsholders and OSPs;
- The majority of OSPs *do not* deal with large-scale infringement or automated notices;
- This majority of OSPs fear the costs of switching to automated notices or *ex ante* measures such as filtering;

- OSPs consider the cost of looking for infringing materials in response to misidentifications or imprecise identifications in notices to be a significant drain on resources;
- OSPs are concerned about notices that target non-infringing materials, in part because they feel constrained by liability fears to respond even to flawed notices by removing targeted material; and
- Descriptions of how the dominance of some services—such as Google’s Search, ContentID and advertising services, and Audible Magic—can affect the ecosystem.⁴

Studies 2 and 3 offer a variety of descriptive statistics about the notices, such as: characteristics of notice senders, the types of works they are protecting, whom they are targeting with notices, and the like. They note demographic shifts over time in the Lumen population. These studies also report a range of “questionable” characteristics we observed in the notices—a diverse group that covers erroneous and overly vague identifications, improper subject matter, and potential fair uses. Overall, 31% of the takedown requests in the Study 2 sample were “questionable,” and 70% of the requests in the smaller Study 3 sample were “questionable” (this becomes 36.3% when one especially prolific sender is removed).

Drawing on all three studies, we offer a number of recommendations. The majority of these include calls to implement “best practices” in notice and takedown for OSPs and rightsholders,⁵ to expand educational efforts for online users and the less-professionalized senders who appeared most likely to send notices targeting non-infringing materials,⁶ and to introduce greater transparency into the system.⁷ We also offer statutory reform suggestions. These are limited in scope and aimed at evening out responsibility for “bad” notices and counter notices—notably by requiring that notices and counter notices not “recklessly misrepresent” claims (a threshold somewhat lower than the current one for sender liability), and by holding notice senders to the same standard of truth as counter notice respondents.⁸

Ford’s critique

By all appearances, Ford’s main goal is to cast doubt on the finding that 31% of the takedown notices in Study 2 were “questionable,” because it suggests that notice accuracy should be improved. To accomplish this goal, he mischaracterizes some of our intentions and findings and obscures others—mostly ignoring Study 1 and dismissing Study 3 in the process. We highlight only a few illustrative flaws in his analysis.

Ford mischaracterizes the purpose of our research.

Ford characterizes our work as “motivated by speech concerns” and a desire “to find mistakes in the notices”—generally in distinction to what he presents as the more important task of quantifying the costs of policy change.⁹ This is inaccurate. Our primary purpose was to provide a descriptive account of how notice and takedown operates in the areas we were able to observe.¹⁰ Such accounts have been missing from the literature, though they are essential to meaningful debates about policy reform. In the course of this work, we considered notice accuracy and potential effects on expression, topics that are also under examined. While these are important aspects of our work, mischaracterizing our motives and reducing our broad, descriptive account to *only* these aspects is false and misleading. Indeed, when convenient, Ford criticizes our work as too “descriptive.”¹¹

Ford inaccurately re-defines “questionable” notices and ignores important costs created by inaccuracy.

In addition to other descriptive statistics, Study 2 and Study 3 report on “questionable” notices we observed in the Lumen data. A notice can be questionable for a number of reasons, including through the misidentification of the targeted material, overly vague identification, improper subject matter, or potential fair use of the targeted material. All are categories that correspond to section 512’s legal requirements.

Ford prefers—though he never clearly articulates—a narrower definition.¹² He ignores the statute to replace our categories with only “requests leading to the takedown of non-infringing works in a manner that impinges expression.”¹³ This is both a bald attempt at limiting the percentage of “questionable” notices—since by his reckoning it leaves out all misidentifications—and unquantifiable.

It is also inaccurate. Besides ignoring section 512's statutory requirements, Ford emphasizes our characterization of some of these mistakes as "relatively minor sins of misidentification," but—unlike Ford—we do not suggest that such notices do *no* harm. That clearly would be incorrect.

First, Ford ignores costs to OSPs completely. This is puzzling in a review of our work. As we document at length in Study 1, misidentifications and imprecise identifications can impose considerable costs on OSPs, which must search fruitlessly through their systems for nonexistent or impossible-to-find material.¹⁴ It is also puzzling in any attempt to model the costs related to section 512: OSP costs are fundamental to the equation of whether section 512 is "working." OSP competition policy was an animating factor behind section 512, which was intended to limit OSPs' liability for their users' infringements, giving them the certainty to invest and grow (indeed, the law's formal name is the On-Line Copyright Infringement Liability Limitation Act). The costs borne by OSPs for alleged infringements go to the heart of this issue. We devote several sections in Study 1 to exploring the effects of enforcement costs on competition, describing how higher costs may reinforce the positions of dominant OSP incumbents.¹⁵ Any true attempt to formally size "harm" must include OSP costs (as well as costs to rightsholders, targets, and the public).

Second, we did not say, and did not mean, that such mistakes will have *no* effect on expression or other values supported by section 512. Study 1 also documents that the direct costs from misidentification can indeed extend to expression. Because of the potentially high liability costs of missing infringing content, OSPs described responding to imprecise notices by removing non-infringing material—at the limit, entire pages or sites—in order to be sure they removed poorly targeted alleged infringements.¹⁶ Less direct costs can also affect expression: as OSPs must increase resource expenditure on searching in response to badly targeted automated notices, they may have to decrease the resources they dedicate to finding and addressing notices that target fair uses or other non-infringing expression (or resources they dedicate to addressing notices from the smaller rightsholders not privileged to use automated systems).¹⁷

Ford ignores all of these costs: the direct financial costs to OSPs; the impact they can have on OSP practices and knock-on effects on targets and other rightsholders; and the negative effects they may have on online expression. His true goal is to reduce the percentage of "questionable" notices in Study 2. By excluding these costs, Ford claims to whittle down the category of problematic notices from 31% to 2%, then .2% and .02%—eventually suggesting that it may be zero. For unspecified reasons, he sticks with the unsupported .2% number for the rest of his argument. This is not a serious engagement with our work.

Ford attempts to dismiss notices that are "questionable" even under his definition.

Study 3 presents a different problem for Ford's analysis. Overall, the study provides the largest analysis to date of notices sent by smaller senders—an important group in discussions about the purposes and effectiveness of section 512.¹⁸ Rather than raising problems with identification, most of the problematic requests in Study 3 involved problems with the underlying claim, such as the targeting of potential fair uses or complaints that didn't involve copyright claims at all. Moreover, these requests were most commonly directed not to file-sharing sites, but to material posted on social networks, blogs and personal websites, and news sites. Most of the questionable notices in Study 3, in other words, fit Ford's narrowest definition.

Since he cannot define these notices away, Ford focuses instead on the fact that they are a proportionally small part of the whole Lumen dataset,¹⁹ ignoring the absolute numbers (around 30,000 requests, of which over 70% were questionable). But Ford offers no explanation of why our sample size or technique should be considered inadequate.²⁰ In practice, our findings align closely with publicly available information about rejection rates for notices at OSPs that are likely to receive "small sender" requests—as well as with the OSP reports offered in Study 1.²¹ Still, we do not claim (or expect) that all smaller senders are necessarily like those in Study 3, and call for more research to understand their experiences and needs. And on the expectation that smaller senders to other services may differ in various ways, we limited relevant recommendations to educational resources that can be helpful in a broad array of contexts and relatively limited statutory changes that even out incentives in a wide variety of situations.

Instead of actual cost-benefit accounting, Ford creates inaccurate models that lead to absurd results.

Ford repeatedly criticizes us for “fail[ing] to quantify” costs and benefits. But he provides no such accounting or new insights himself. His models include no actual costs, for any stakeholder. His analysis instead rests on unsupported assumptions that (major) rightsholders’ (unquantified) costs are too high, that any harm to expression or other public costs are negligible, and that OSP costs are either irrelevant or non-existent, either of which is clearly wrong in light of the statutory goals and our empirical work.

In lieu of numbers, Ford provides two formalized “models.” The first, discussed above, attempts to describe “questionable” notices and “harm” but provides no actual numbers and leaves out costs and benefits for many stakeholders.

Ford’s second model purports to describe the overall cost to the system of increased accuracy in notices (a goal to which some of our recommendations are directed).²² Though it is sold as providing an overall picture, the model includes only the costs to rightsholders of searching for infringements and avoiding noticing mistakes (oddly, it doesn’t even include rightsholders’ costs of generating, sending, or tracking notices). Characteristically, it leaves out all costs to OSPs of searching for infringement in response to notices, costs to OSP of processing notices, costs to targets of responding to notices, costs to society of mistakes, and so on. The model doesn’t even take into account the costs to *rightsholders* of inaccurate notices—*notices cost something to generate, even if they are too poorly targeted to be actionable—or the benefits rightsholders may gain from more efficient search methods.*²³

A good test of an economic model is whether it avoids producing absurd results under simple assumptions. Ford’s model fails this test. The optimal level of infringement search (for rightsholders—OSP and others being completely elided) is defined as a function of the care taken to avoid false positives, with *any* increase in the level of care (from 0) reducing the optimal level of search.²⁴ This is an extraordinary assertion. It implies that *any* care (Ford has now set aside entirely any idea that costs also arise from bad notices) is too much. Viewed another way, it appears from Ford’s argument that *no* level of mistake or abuse would be sufficient to require reform, because any incentive to increase accuracy would reduce searches by rightsholders (and thus increase infringements—a causal assumption that is not proven and indeed is unlikely). Under Ford’s model, the most efficient outcome is for rightsholders to abandon accuracy entirely and send notices targeting every page on the Internet. This is not a serious attempt to describe the costs in the system.

Ford dramatically misrepresents our recommendations for human review.

Ford does provide numbers at one point. For some reason, he reads our observation that the level of potential mistakes in the notices shows that they “would benefit from human review” as a recommendation for human review of *all* automated notices. He extrapolates from the number of notices to conclude that applying this to our Lumen dataset would take “27 million man-hours” and cost somewhere between \$4 billion and \$8 billion to implement.²⁵ This is nonsense. While entertaining, it is based on a fundamental mischaracterization of our work. Perhaps Ford misread our actual recommendations—largely based on *actual, existing practice* documented in Study 1—that both rightsholders and OSPs use human spot checks to review some automated notices, and in some cases use sampling to do so.²⁶ It is hard to see how, since “spot check,” “cross-check,” or “sample” is clearly stated in each recommendation. We are left thinking that motivated reasoning makes for careless reading.

Ford makes other careless errors.

Similar carelessness infects other critiques.

We called for reforms to statutory damages. We based our suggested reforms, in large part, on a report from the Internet Policy Task Force (“IPTF”), part of the United States Department of Commerce, whose report drew from a rigorous multistakeholder proceeding. Ford misidentifies the IPTF as the “Intellectual Property Task Force” (sic) and misclassifies it as a “special interest [group] generally opposed to strong copyright enforcement.”²⁷

We stated that notice senders “are in the best position to improve accuracy and avoid the worst mistakes.” Ford claims there is no support for this “debatable” proposition. It may be debatable, but it is the considered position of the U.S. Congress, which in section 512 placed the burden of identifying infringing material unequivocally on copyright holders because it considered that they know what material they own, and “are thus better able to efficiently identify infringing copies than service providers [...] who cannot readily ascertain what material is copyrighted and what is not.”²⁸ Courts have repeatedly considered, and repeatedly affirmed, this requirement.²⁹

Concluding thoughts.

It is possible that our suggested best practices and reforms³⁰ could raise the cost of enforcement for Ford’s large rightsholder clients too high for good policy. We think this is quite unlikely: we learned in Study 1 from large rightsholders and OSPs what they currently do and used this to inform our recommendations in the absence of actual cost numbers (which no respondent was willing to disclose).

Still, Ford’s criticisms of our paper, and his approach in pursuing them, appear to be rooted in a concern that some of our findings and recommendations could raise costs for his clients. This is fair enough, as far as it goes, and we would be glad to reconsider our recommendations if he were to provide persuasive evidence that we should do so. But after careful consideration, we cannot find it in his review. Ford offers no assessment of actual costs, no sense of the trade-off between rightsholder costs and others’ costs, and no attempt to show the magnitude of difference in cost caused by our recommendations (with the exceptions of the misfire regarding human review noted above and the formal model that counts *any* accuracy as too costly). We do agree that requiring rightsholders to avoid “reckless” misrepresentations in their notices requires more of them than the current standard that requires avoiding only “knowing, material misrepresentations,” but without more, we do not see how avoiding recklessness—which, after all, is less burdensome than even negligence—is an unreasonable cost. Allowing for limited statutory damages for truly bad notices, similarly, hardly seems to threaten good faith, reasonably responsible senders. Ford characterizes this as a “tax” on sending notices, but it is a tax that applies only to the reckless. The claims of due diligence by major rightsholder groups, if reflected in their actual practice, should give them no reason to worry.

We did not address every mistake, much less Ford’s many misunderstandings regarding copyright, fair use, and constitutional law. Instead, we offer some general concluding observations. Our work is a broad-based, descriptive piece that attempts to provide a picture of as much of the system as possible, with recommendations for further research to cover portions of the system we could not observe. Similarly, our “best practice” and policy recommendations seek to take into account what we learned to improve the system across the board. Ford’s self-described “review” of this broad work only presses large rightsholders’ interests. In the end, we barely recognized our work in it. Ford also misunderstands or ignores the actual law and policy as developed by Congress. He appears unaware that any stakeholder beyond his clients has an interest in notice and takedown, let alone interests affirmatively recognized in the law. The considerable complexity of the system escapes him, and thus his models fail to reflect even a respectable portion of it. While we are certainly sympathetic to major rightsholders’ difficulties managing large-scale infringement—something we take account of throughout our report—this is not the only interest at stake, and it is not the only interest Congress sought to take into account in crafting the DMCA. OSPs, rightsholders, targets, and the public all garnered Congressional consideration and affordances. Our work follows that guide. There is more to do to understand this complex system. We wish that Ford’s effort were more useful to this important work.

Notes

¹ Jennifer M. Urban, Joe Karaganis, & Brianna L. Schofield, NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE (Mar. 2017), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2755628. The study received support from Google Inc. and the Sloan Foundation. Neither funder directed our approach in any way, and neither funder reviewed any methods, data, results, or reporting before public release. The study also received an in-kind donation of the OpusData database and expert database customization advice from Nash Information Services.

² George S. Ford, *Notice and Takedown in Everyday Practice: A Review* (2017).

³ LUMEN, <https://lumendatabase.org/> (last visited May 26, 2017).

⁴ NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1 at 64-75. We note that these are limited findings. More research is needed in this area.

⁵ NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1 at 134-138.

⁶ *Id.* at 137-140.

⁷ *Id.* at 131-132.

⁸ *Id.* at 127-133.

⁹ Ford, *supra* note 2 at 2, 9.

¹⁰ As we describe in the introduction, “the research provides a detailed account of notice and takedown, first from the perspective of stakeholders who interact with the system, and then through independent reviews of takedown requests.” *Notice and Takedown in Everyday Practice*, *supra* note 1 at 9.

¹¹ Ford, *supra* note 2 at 9.

¹² Ford, *supra* note 2 at 5-6, and then rather confusingly embedded as assumption throughout. Ford never actually defines the term “questionable notice,” *per se*, but uses his value N as its equivalent. He further attempts to formalize with the expression hN , representing the “average level of harm” (h , otherwise undefined) multiplied by the “number of requests leading to the takedown of non-infringing works in a manner that impinges expression” (N). This is meant to “size” harm. Ford, *supra* note 2 at 5. Already we see that N , which does not reflect our methods or analysis, is merely an attempt to limit the definition of “questionable” notices, that it disregards any number of statutory flaws and potential issues that create costs within the system, and that “average harm” is not quantified. Ford then adds the concept of “error rate,” but also defines it only in terms of N , thus merely adding to the inaccuracy. He proceeds to conflate all of these concepts in his discussion, an unsurprising set of mistakes in light of his actual goal: to define away as many “questionable” notices as possible. *Id.* at 5-9 and *passim*.

¹³ Ford, *supra* note 2 at 5.

¹⁴ NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1 at 38-40.

¹⁵ NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1 at 64-75. We note that these are limited findings. Competition policy is an area that needs more research.

¹⁶ See NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1 at 38-43.

¹⁷ As Ford’s later “optimal search” model would have it (had he included OSPs in it), it is indeed *necessary* that increased OSP search costs would decrease their ability to address substantive problems, respond to other notices, and the like. Ford, *supra* note 2 at 27.

¹⁸ See NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1 at 83.

¹⁹ This is because the vast numbers of automated notices sent by large rightsholders “swamp” everything else.

²⁰ Though he relies on our data when he thinks it favorable to his point, see e.g., Ford, *supra* note 2 at 12-13, 18, 20-21, Ford has a number of other complaints about the fact that the Lumen data do not statistically represent notice and takedown across the entire internet. No such data are available, and we discuss our study limitations at length in our methods sections. We refer readers to the studies for details on our data-gathering, and the lack of data available overall, but we are comfortable that Study 1 provides a broadly representative picture of notice and takedown from the perspectives of OSPs and large rightsholders and that the data in Studies 2 and 3—overwhelmingly involving Google, an important notice and takedown participant according to respondents in Study 1—are also informative. In addition, Study 2 and 3 observations both buttressed key observations from OSPs and rightsholders in Study 1 and vice versa. We read all the results together in considering our recommendations, and called for more research into the areas we could not observe (for example, targets’ and small rightsholders’ direct experiences). Ford simply ignores a large portion of this.

²¹ See NOTICE AND TAKEDOWN IN EVERYDAY PRACTICE, *supra* note 1 at 31 (stating that in 2014, Wikimedia Foundation reported receiving twenty-one takedown requests and rejected 76%, and Reddit reported receiving 176 takedown requests and rejected 62%) .

²² Ford, *supra* note 2 at 26-27.

²³ Ford also does not attempt to quantify the cost of infringement; he instead makes the claim that it is a monotonically decreasing function of search—a doubtful proposition for which he provides no evidence. Ford, *supra* note 2 at 26, 27. In this he follows the major industry groups like the RIAA and MPAA, which gave up on cost-benefit analyses for infringement a decade ago in the face of criticism of their opaque methodologies, arbitrary assumptions, and the general failure of predictions of wider industry and cultural decline. See *Soc. Sci. Research Council*, MEDIA PIRACY IN EMERGING ECONOMIES (Joe Karaganis ed., 2011), <http://piracy.americanassembly.org/the-report/>.

²⁴ Ford has set up the equation so that the overall value will always be negative. But there is no need to parse the equation, as he helpfully writes out that “the optimal level of search will decline with increases in the level of care. A higher level of mandated care to avoid false flags, like those proposed in *Everyday Practice*, may result in an improvement in accuracy but also a reduction of search and, consequently, an increase in infringement, F(s).” Ford *supra* note 2 at 27.

²⁵ Ford, *supra* note 2 at 23-24.

²⁶ *Id.* at 134-136.

²⁷ Ford, *supra* note 2 at note 82.

²⁸ S. REP. No. 105-190 at 48; H.R. REP. No. 105-551, pt.2 at 57-58.

²⁹ *Perfect 10, Inc. v. CCBill, LLC*, 488 F.3d 1102, 1114 (9th Cir. 2007); *UMG Recordings, Inc. v. Veoh Networks Inc.*, 665 F. Supp. 2d 1099, 1108-09 (C.D. Cal. 2009), *aff'd sub nom. UMG Recordings, Inc. v. Shelter Capital Partners LLC*, 667 F.3d 1022 (9th Cir. 2011), opinion withdrawn and superseded on reh'g, 718 F.3d 1006 (9th Cir. 2013), and *aff'd sub nom. UMG Recordings, Inc. v. Shelter Capital Partners LLC*, 718 F.3d 1006 (9th Cir. 2013); *Hendrickson v. eBay, Inc.*, 165 F. Supp. 2d 1082, 1093 (C.D. Cal. 2001).

³⁰ Ford characterizes us as saying that statutory reform is a “high priority” (quoting those two words from a longer sentence). We actually said that addressing mistake and abuse would be a “high priority for reform” with the implicit understanding that this is if Congress chooses to pursue reform. We actually do not think statutory reform is so high a priority: well-implemented best practices and educational efforts may be most effective at helping notice and takedown work well for all parties. But if reform is in the offing, our recommendations are supported by what we were able to study about notice and takedown, and we stand behind them.