

Before the United States Senate Committee on the Judiciary,  
Subcommittee on Intellectual Property,  
Hearing on the Digital Millennium Copyright Act at 22:  
How Other Countries Are Handling Online Piracy

March 10, 2020

Statement of Daphne Keller  
Stanford University Cyber Policy Center

Chairman Tillis, Ranking Member Coons, and Members of the Subcommittee:

Thank you for the opportunity to appear at this hearing to offer testimony about laws in Europe and around the world. I specialize in intermediary liability – the body of laws that define Internet platforms’ responsibility for content shared by their users. So while I will primarily speak to copyright and international analogs of the Digital Millennium Copyright Act (DMCA), I will also draw on closely related laws outside the U.S. governing things like platform liability for defamation or violent extremism. These are laws I have studied and written about at Stanford since 2015. I also have extensive real-world experience with global approaches to platform regulation from my prior role as Associate General Counsel at Google, where I worked from 2004 until 2014.

I will first briefly describe the goals and trade-offs inherent in any intermediary liability law; then describe the DMCA’s place in the global landscape of intermediary liability laws; then discuss concerns about filtering proposals and the recently enacted Article 17 of the EU’s Digital Single Market (DSM) Directive.

## I. Intermediary Liability’s Goals and Mechanisms

Broadly speaking, all intermediary liability laws balance three goals. First, they aim to prevent harm – in the case of copyright, to prevent infringement and piracy. Second, they aim to protect Internet users’ lawful speech and activity online. Users’ rights are threatened when laws incentivize risk-averse platforms to err on the side of taking down lawful speech. Third,

intermediary liability laws aim to support innovation and competition among online services. This concern is particularly acute for small businesses that can't afford to hire armies of moderators or build expensive filtering technologies.

When lawmakers update or adjust intermediary liability laws, they are effectively recalibrating the balance between these three often-competing goals, which I will loosely refer to as (1) harm-prevention, (2) speech, and (3) innovation. Many of the legislative knobs and dials available to adjust that balance are relatively well-understood.<sup>1</sup> For example, to make a law more speech-protective, lawmakers might require more robust *counternotice* mechanisms – systems for Internet users to learn about and dispute a platform's decision to take down potentially infringing content.

Many intermediary liability laws around the world, like the DMCA, lead platforms to adopt “notice and takedown” systems. Qualifying intermediaries are generally immunized for user-generated content, but lose that immunity if they fail to remove the content after learning of its existence – typically through notice from a rightsholder. Notice and takedown systems can be important harm-reduction mechanisms. And systems with clearly defined roles and processes, like the DMCA, provide legal certainty to both rightsholders and platforms – a factor that is particularly important to small platforms who can't hire lawyers to assess every claim. But, as numerous studies document, notices are also widely misused by bad or mistaken actors to target lawful speech for removal.<sup>2</sup>

To be clear, my point here is not that U.S. rightsholders are misbehaving. It is that most notice and takedown systems give *anyone* easy leverage to silence other people by complaining to a private platform. They give even well-meaning notifiers too little incentive to be careful about accusations made in a notice. And they give platforms far too much reason to comply with false or mistaken notices, in order to avoid expense and risk to themselves. Examples of abuse in

---

<sup>1</sup> Daphne Keller, *Build Your Own Intermediary Liability Law: A Kit for Policy Wonks of All Ages*, Balkinization (2019), <https://balkin.blogspot.com/2019/06/build-your-own-intermediary-liability.html?m=1>.

<sup>2</sup> See studies cited in Daphne Keller, *Empirical Evidence of “Over-Removal” by Internet Companies under Intermediary Liability Laws*, Stanford Center for Internet and Society Blog (2015), <http://cyberlaw.stanford.edu/blog/2015/10/empirical-evidence-over-removal-internet-companies-under-intermediary-liability-laws>.

notice and takedown systems are innumerable. They include businesses trying to knock their competitors off of important platforms, and researchers attempting to hide evidence of mistakes in their work.<sup>3</sup> The President of Ecuador has reportedly used copyright notices to remove critical news reporting from platforms including YouTube, Facebook, Vimeo, and Dailymotion.<sup>4</sup> During the 2008 U.S. Presidential election, both the McCain and Obama campaigns had material removed from major platforms based on over-reaching DMCA notices.<sup>5</sup>

Some laws attempt to correct for these problems through procedural rules. In principle, these can work somewhat like the rules of procedure in court, erecting barriers to improper claims and letting the legitimate ones through. These include mechanisms like the DMCA's counternotice provision – which encourages but does not require platforms to permit user appeals of takedown decisions – and the DMCA's 512(f) penalties for claimants whose notices knowingly target lawful speech. Broadly speaking, procedural rules serve intermediary liability law's speech goals, but may impair the harm-prevention goal if they deter legitimate notices, or the innovation goal if they are too burdensome for smaller companies. In practice, it is hard to measure the value of procedural rules. What data we have suggests that once improper removals take place, after-the-fact measures like counternotice are not enough – the removals frequently go uncorrected.<sup>6</sup> A few countries around the world have experimented with more stringent process

---

<sup>3</sup> Jennifer Urban and Laura Quilter, *Efficient Process or 'Chilling Effects'? Takedown Notices Under Section 512 of the Digital Millennium Copyright Act*, 22 SANTA CLARA HIGH TECH. L.J. 621 (2005); Ivan Oransky, *WordPress Removes Anil Potti Posts from Retraction Watch in Error After False DMCA Copyright Claim*, Retraction Watch (2013), <http://retractionwatch.com/2013/02/05/wordpress-removes-anil-potti-posts-from-retraction-watch-in-error-after-false-dmca-copyright-claim/>.

<sup>4</sup> James Ball and Paul Hamilos, *Ecuador's President Used Millions Of Dollars Of Public Funds To Censor Critical Online Videos*, BuzzFeed (2015), <https://www.buzzfeednews.com/article/jamesball/ecuadors-president-used-millions-of-dollars-of-public-funds>.

<sup>5</sup> Center for Democracy and Technology, *Campaign Takedown Troubles* (2010), [https://cdt.org/wp-content/uploads/pdfs/copyright\\_takedowns.pdf](https://cdt.org/wp-content/uploads/pdfs/copyright_takedowns.pdf).

<sup>6</sup> Jonathon Penney, *Privacy and Legal Automation: The DMCA as a Case Study* 22 STAN. TECH. L. REV. 412, 477 (2019) (finding no counternotices in sample set of 1000 DMCA takedowns and reporting that “only 34% of respondents in the survey indicated they would take steps to legally challenge a notice even when they believed it was incorrect or wrong”); Jennifer Urban, et al., *Notice and Takedown in Everyday Practice* (2016) at 95 (noting that both Urban et al's study and “and other research on this issue consistently shows that counter notices are rarely used”), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2755628](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2755628); Annemarie Bridy and Daphne Keller, *U.S. Copyright Office Section 512 Study: Comments in Response to Notice of Inquiry* (2016) (reviewing transparency report data and discussing possible reasons for infrequency of counternotice) at 27-30, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2757197&download=yes](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2757197&download=yes).

rules and transparency measures, for example under Germany’s NetzDG law.<sup>7</sup> As a result, we may have a natural experiment and basis for comparison as time passes. One outcome of Germany’s experiment has emerged already: the German government recently charged that Facebook was avoiding its obligations under NetzDG by simply using the broader prohibitions of its Terms of Service (TOS) instead of the law as grounds to take down content, thus avoiding the law’s transparency requirements.<sup>8</sup>

## II. The DMCA in Global Context

U.S. copyright law under the DMCA, broadly speaking, protects more speech and innovation than most of its European analogs, but less than many other laws around the world. Globally, there are intermediary liability laws that incentivize both more and less content removal than the DMCA. (Intermediary liability rules are not alone in shaping platforms’ removal incentives, of course. Other considerations, such as availability of statutory damages, can be very important.)

At one extreme on this spectrum would be laws that impose strict liability – meaning a platform is automatically liable, even for content it does not know about – or that require platforms to proactively monitor, police, or filter content posted by users. Until recently, laws like this were often hypothesized by academics, but were very rare in the real world. Article 17 of the EU’s new Copyright Directive, which I will discuss more below, has changed that by requiring platforms to proactively monitor their users. But pre-existing European laws under the eCommerce Directive, an EU legislative instrument dating to 2000, already tilted the scales toward removal in subtler ways.<sup>9</sup> One was through relatively vague legal standards in the implementing laws of EU Member States. It is generally harder for platforms in Europe, compared to platforms in the U.S., to know whether they qualify for immunity, because of widespread uncertainty about which services are sufficiently “passive” to qualify for safe

---

<sup>7</sup> Act to Improve Enforcement of the Law in Social Networks (2017) (English translation), [https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG\\_engl.pdf?\\_\\_blob=publicationFile&v=2](https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG_engl.pdf?__blob=publicationFile&v=2).

<sup>8</sup> Thomas Escritt, *Germany fines Facebook for under-reporting complaints*, Reuters (July 22, 2019) <https://www.reuters.com/article/us-facebook-germany-fine/germany-fines-facebook-for-under-reporting-complaints-idUSKCN1TX11C>.

<sup>9</sup> Directive 2000/31, 2000 O.J. (L 178) 1 (EU).

harbors. Similarly, European laws are often less clear than the DMCA about when a platform might be considered to have enough “knowledge” of infringement to trigger a takedown obligation. This uncertainty means that, any time a platform receives a notice alleging that a user’s speech is illegal, they have a greater incentive to simply honor the notice and take down the speech.

While Europe’s laws generally incentivize more removal than the DMCA, laws in some other countries incentivize less – effectively prioritizing speech and innovation goals at the expense of potentially tolerating more harms, in the form of infringement. Canada’s “notice and notice” copyright system, for example, requires platforms to facilitate certain communications between rightsholders and users accused of infringement, but not to take down content.<sup>10</sup> Another example is Chile’s “judicial notice” system, in which copyright removal requests must first be reviewed by a court before they are delivered to a platform.<sup>11</sup>

Outside the copyright context, Supreme Courts in Argentina and India have said that such judicial or agency review is *constitutionally required* – that users’ speech rights are violated if the law simply leaves platforms to resolve potentially complex legal questions based on private notices.<sup>12</sup> A Mexican court held that a publisher affected by a takedown request must, as a matter of due process rights, be notified and given an opportunity to participate in a hearing.<sup>13</sup> These rulings track recommendations from human rights officials, including the UN Free Expression Rapporteur Frank La Rue, who said in an official report that “censorship measures should never be delegated to a private entity[,]” that in general “no one should be held liable for content on the Internet of which they are not the author,” and that “provisions for the imposition of liability on intermediaries should have sufficient judicial safeguards so as not to cause or encourage private censorship[.]”<sup>14</sup> Judicial review as a predicate for takedown is also included in intermediary

---

<sup>10</sup> Copyright Act Sect. 41.25 (Canada).

<sup>11</sup> Law No. 20.435 Art. 85 (Chile).

<sup>12</sup> *M. Belen Rodriguez c/Google y Otro s/ daños y perjuicios*, Civil R.522.XLIX (2014) (Argentina) (noting exceptions where unlawful nature of content is “evident and can be seen directly” without judicial assessment); *Singhal v. Union of India*, 12 SCC 73, Para. 100, 117 (2015) (India).

<sup>13</sup> *La Fortuna v. INAI*, 355/2016 (Mexico), <https://perma.cc/5WQC-JJMV>.

<sup>14</sup> UN General Assembly, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression* (Frank La Rue) A/HRC/17/27 (2011) Para. 43 and 47, <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G11/132/01/PDF/G1113201.pdf?OpenElement>; see also *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression* (David Kaye)

liability legislation in Brazil<sup>15</sup> and in recommendations from global civil society organizations.<sup>16</sup> While Brazil does permit takedown based on private notice in copyright cases, Brazilian courts recently penalized both a platform and a rightsholder for taking down the wrong thing. They ordered YouTube to reinstate musical parody videos, and held both YouTube and the rightsholder that had requested the removal liable for damages.<sup>17</sup>

As these examples illustrate, international approaches to intermediary liability fall along a broad spectrum. In some cases, the variation reflects legislative policy choices balancing the laws' priorities. In others, courts have held that laws *must* prioritize speech, based on national equivalents of the First Amendment. U.S. courts, too, have struck down laws that went too far in encouraging private censorship. Key U.S. Supreme Court cases in this area involve “analog intermediaries” such as bookstores. As the Supreme Court said in one such case, strict liability would incentivize the store owner to “restrict the books he sells to those he has inspected[,]” effectively “impos[ing] a restriction upon the distribution of constitutionally protected” speech. Such “self-censorship, compelled by the State, would be a censorship affecting the whole public, hardly less virulent for being privately administered.”<sup>18</sup> Concern about speech suppression that is *motivated* by laws but *carried out* by private companies is highly relevant for lawmakers weighing the competing goals of platform liability laws today.<sup>19</sup>

In the next portion of my remarks, I will discuss rights-based concerns about the EU's new filtering requirement in Article 17 of the DSM Directive.

---

(2016), <https://perma.cc/44AY-ZX9G>; *Freedom of Expression and the Internet*, OAS Report (Catalina Botero) (2013) Para. 105 and 106.

<sup>15</sup> Marco Civil da Internet, Federal Law no. 12.965 (Brazil) (for non-copyright removals, “in order to ensure freedom of expression and to prevent censorship, an Internet application provider shall only be subject to civil liability for damages caused by virtue of content generated by third parties if, after specific court order, it does not take action, according to the framework and technical limits of its services and within the time-frame ordered, to make the infringing content unavailable.”)

<sup>16</sup> Manila Principles on Intermediary Liability, <https://www.manilaprinciples.org> (“Intermediaries must not be required to restrict content unless an order has been issued by an independent and impartial judicial authority”).

<sup>17</sup> Civil Appeal No. 0000412-86.2016.8.24.0175 (Meleiro); Civil Appeal No. 0000447-46.2016.8.24.0175 (Meleiro).

<sup>18</sup> *Smith v. California*, 361 U.S. 147, 153-54 (1959); *Bantam Books, Inc. v. Sullivan*, 372 U.S. 58 (1963) (rejecting bookseller liability based on notice from government agency); *see also CTD v. Pappert*, 337 F. Supp. 2d 606, 610 (E.D. PA 2004) (striking down statute requiring ISPs to block child pornography based on notice from government agency).

<sup>19</sup> *See generally* Daphne Keller, *Who Do You Sue? State and Platform Hybrid Power over Online Speech* (2019), Hoover Institute, [https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech\\_0.pdf](https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech_0.pdf).

### III. Filtering Proposals and Article 17 of the EU DSM Directive

Article 17 of the EU DSM Directive instructs covered platforms to make “best efforts” to prevent uploads of specified works, but to simultaneously protect users’ rights to quote, criticize, parody, and otherwise lawfully use those same works. While it does not use the word “filter,” Article 17 creates a de facto filtering mandate. The European Commission has convened “stakeholder dialogues” to discuss, among other things, whether any technology is actually capable of doing what Article 17 requires: distinguishing between infringing and non-infringing uses of the same work. There is much more to Article 17, but I understand that other testimony will explain the new law in more detail, so I will not expand on it here. Instead, I will focus on three things. First, the law’s implications for speech rights. Second, the risks it creates for other rights, including equal protection – risks that researchers are only beginning to appreciate, and that have received far too little attention in the public discussion to date. Third, the problems with relying on interested companies, and in particular on major platforms like Facebook and Google, as sources of information about filters’ capabilities.

#### A. Speech and Censorship Issues

Article 17 was enacted amidst considerable controversy. Poland’s government is currently challenging it before the EU’s highest court, the Court of Justice of the European Union (CJEU), arguing that it violates the EU’s Constitutional safeguards for free expression.<sup>20</sup> The outcome of that case is hard to predict, in part because Poland’s arguments are not publicly known. But the likely objections are easy to surmise, based on the public debate that preceded both the Copyright Directive and another proposed filtering measure in the EU’s draft Terrorist Content Regulation.<sup>21</sup>

---

<sup>20</sup> Andrew Liptak, *Poland has filed a complaint against the European Union’s copyright directive*, The Verge (May 25 2019), <https://www.theverge.com/2019/5/25/18639963/poland-european-union-copyright-directivefiled-complaint-court-of-justice>.

<sup>21</sup> Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online, COM(2018) 640. The Regulation is close to being finalized in “trilogue” proceedings as of March 2020, according to leaked compromise amendments. [https://www.politico.eu/wp-content/uploads/2020/03/SKM\\_C45820030612100.pdf](https://www.politico.eu/wp-content/uploads/2020/03/SKM_C45820030612100.pdf).

The most widely raised objections to Article 17 involve Internet users’ rights to speak and access information online. As numerous commenters pointed out, while the best filters are good at spotting duplicate content, they are not able to discern when content is re-used in a new context. This context-blindness is a key reason why filters are relatively well-accepted as tools for platforms to weed out content that is illegal in every context, in particular for horrific material like child sexual abuse images. Filters remain very controversial, though, in areas where context matters, including copyright. In the EU’s most recent filtering debate, regarding terrorism, UN human rights officials,<sup>22</sup> academics (including myself), and civil society groups including the ACLU raised concerns about, in the words of the ACLU’s letter, “enacting laws that will drive Internet platforms to adopt untested and poorly understood technologies to restrict online expression.”<sup>23</sup>

European courts have created an important body of case law regarding potential filtering or monitoring requirements for platforms. For many years, the CJEU rejected proposed filtering mandates, citing both rights concerns and legislative barriers under the eCommerce Directive.<sup>24</sup> In 2019, however, the Court approved a monitoring injunction for the first time, in a case involving platform liability for defamation. In that case, *Glawischnig-Piesczek v. Facebook Ireland*, an Austrian court ruled that a Facebook user broke the law by calling a politician a “lousy traitor,” a “corrupt oaf,” and a member of a “fascist party.”<sup>25</sup> It ordered Facebook to

---

<sup>22</sup> David Kaye, Joseph Cannataci and Fionnuala Ni Aoláin, Mandates of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression; the Special Rapporteur on the right to privacy and the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism (2018),

<https://spcommreports.ohchr.org/TMResultsBase/DownloadPublicCommunicationFile?gId=24234>.

<sup>23</sup> Access Now et al., Letter to Members of European Parliament (2019),

<https://cdt.org/files/2019/02/Civil-Society-Letter-to-European-Parliament-on-Terrorism-Database.pdf>; see also

Article 19 et al., letter on European Commission regulation on online terrorist content (2019);

<https://www.article19.org/resources/joint-letter-on-european-commission-regulation-on-online-terrorist-content/>;

WITNESS et al., Letter to the Committee on Civil Liberties, Justice and Home Affairs (2019),

[https://drive.google.com/file/d/1WTgl5hjJ\\_cAE1U0OjqaQ9AucU6HNIhoi/view](https://drive.google.com/file/d/1WTgl5hjJ_cAE1U0OjqaQ9AucU6HNIhoi/view).

<sup>24</sup> Case C-70/10, *Scarlet Extended SA v. SABAM*, (2011); Case C-360/10, *SABAM v. Netlog NV*, (2012); Case C-324/09, *L’Oréal SA v. eBay* (2011); Case C-484/14 *McFadden v. Sony Music* (2016); see also Case C-494/15, *Tommy Hilfiger Licensing LLC and others v. Delta Center A.S.* (2016) (discussing physical marketplace).

<sup>25</sup> Case C-18/18, *Eva Glawischnig-Piesczek v. Facebook Ireland Limited* (2019); see also Daphne Keller, *Facebook Filters, Fundamental Rights, and the CJEU’s Glawischnig-Piesczek Ruling* (forthcoming in GRUR International Journal of European and International IP Law); *Dolphins in the Net: Internet Content Filters and the Advocate General’s Glawischnig-Piesczek v. Facebook Ireland Opinion*, Stanford Center for Internet and Society White Paper (2019), <https://cyberlaw.stanford.edu/files/Dolphins-in-the-Net-AG-Analysis.pdf>.



proactively filter all posts on the platform to ensure that no user could make the prohibited statements again. The CJEU held (1) that this monitoring requirement was permissible, and (2) that Austria could enforce the order *globally*, requiring Facebook to silence users everywhere in the world.

A filtering requirement like the one imposed in Facebook's Austrian case would be unconstitutional in the U.S. It raises serious questions about Internet users' rights under EU law as well. The CJEU did not examine the rights question in Facebook's case, but is expected to do so when it reviews Poland's challenge to Article 17. According to other CJEU precedent, filters interfere with two user rights: free expression and privacy. The European Court of Human Rights (ECtHR) has, in one case, ruled that holding a platform strictly liable for defamation posted by users – effectively forcing it to monitor or filter user speech in order to avoid liability – violated the European Convention's equivalent of the First Amendment.<sup>26</sup> In another case, however, the ECtHR accepted a strict platform liability rule for hate speech.<sup>27</sup> Courts in other parts of the world, including the supreme courts of India and Argentina in the cases I mentioned earlier, have also indicated that strict liability rules for platforms violated their constitutions.

It remains to be seen how the CJEU will assess expression rights issues in evaluating Poland's challenge to Article 17. Presumably the law's defenders will argue first that the threat of over-removal by filters is exaggerated, and second that any errors can be addressed by platforms carrying out human review of filters' decisions and providing users with complaint and redress mechanisms, as required by Article 17.9. These are factual propositions: that filters will work, or else that platforms' human employees will fix the filters' mistakes. As I will discuss later in my testimony, there is substantial cause for concern about both of these claims, and a dearth of reliable information to help assess them.

---

<sup>26</sup> *Magyar Tartalomszolgáltatók Egyesülete (MTE) v. Hungary*, App. No. 22947/13, ECtHR (2016).

<sup>27</sup> *Delfi AS v. Estonia*, App. No. 64569/09, ECtHR (2015); see also Daphne Keller, *New Intermediary Liability Cases from the European Court of Human Rights: What Will They Mean in the Real World?*, Stanford Center for Internet and Society Blog (2016), <http://cyberlaw.stanford.edu/blog/2016/04/new-intermediary-liability-cases-european-court-human-rights-what-will-they-mean-real>.

## B. Important Additional Issues

Policy discussions of potential platform filtering requirements have long focused on speech. I've emphasized it here both because of its paramount importance under the U.S. Constitution, and because it is central to the international case law and literature to date. But speech is by no means the only issue. Other pressing concerns about filters include threats to user privacy and a very troubling new body of research regarding bias and disparate impact. These issues warrant considerably more attention, and should shape the next stage of our thinking about filtering mandates like Article 17.

- **Disparate Impact:** A small but growing body of research suggests that automated content control systems have disparate harmful impact on minority communities. One study, for example, found that algorithms to filter out “toxic” speech unfairly penalized speakers of African American vernacular English.<sup>28</sup> In the copyright realm, critics have raised concerns about filtering errors hurting hip hop artists more than musicians in other genres.<sup>29</sup> The empirical studies on point are very new, mostly from 2019. So far, I know of no cases or laws addressing this issue.
- **Privacy:** Ordering a platform to monitor users’ communications raises serious questions about surveillance and privacy. The CJEU has noted but never closely examined this concern, and I am not aware of other international precedent on point. However, in one potentially relevant case, the CJEU struck down a law that similarly leveraged the capability of private communication channel owners, requiring carriers to retain vast

---

<sup>28</sup> Maarten Sap et al., *The Risk of Racial Bias in Hate Speech Detection*, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics 1668 (2019), <https://homes.cs.washington.edu/~msap/pdfs/sap2019risk.pdf>; Anna Chung, *The Suppression of Black Language In the Age of Machine Learning*, Medium (2019), <https://onezero.medium.com/howautomated-tools-discriminate-against-black-language-2ac8eab8d6db>; Thomas Davidson et al., *Racial Bias in Hate Speech and Abusive Language Detection Datasets*, arXiv:1905.12516 (2019); Center for Democracy and Technology, *Mixed Messages? The Limits of Automated Social Media Content Analysis* (2017), <https://cdt.org/insight/mixed-messages-the-limits-of-automated-social-media-content-analysis/>; see also Jillian York and Karen Gullo, *Offline/Online Project Highlights How the Oppression Marginalized Communities Face in the Real World Follows Them Online*, EFF (2019), <https://www.eff.org/deeplinks/2018/03/offlineonline-project-highlights-how-oppression-marginalized-communities-face-real>.

<sup>29</sup> Toni Lester & Dessimlava Pachamanova, *The Dilemma of False Positives: Making Content ID Algorithms More Conducive to Fostering Innovative Fair Use in Music Creation*, 24 UCLA ENT. L. REV. 51, 53 (2017).

amounts of user data for law enforcement use.<sup>30</sup>

- **Competition:** As representatives of small businesses in both the US<sup>31</sup> and EU<sup>32</sup> have attested, expensive legal requirements that may be tolerable for a Facebook or a Google (which has spent over \$100 million on ContentID<sup>33</sup>) can be crippling for their smaller competitors. Such requirements can also deter investment in new market entrants entirely.<sup>34</sup> For Article 17, likely costs include not only licensing or developing filters, but also reconfiguring existing systems; ongoing maintenance and user support; paying employees to review and correct filters' errors or hear appeals; and technical lock-in resulting from adoption of filtering technology. In Europe, this concern about burden on businesses has a legal dimension, and was raised as a factor in several of the CJEU filtering cases. The DSM Directive attempts to mitigate this problem by reducing obligations of smaller companies in their first three years of operation. This, too, will be an important legal experiment to observe as it plays out in Europe.
- **Due Process:** European scholars and NGOs have charged that filtering mandates and other laws requiring platforms to “adjudicate” legal claims about users’ speech may violate the equivalent of due process rights – effectively outsourcing speech governance from courts and legislators to private companies.<sup>35</sup> A handful of rulings and legal instruments around the world, including the DSM Directive itself, have begun to wrangle

---

<sup>30</sup> *Digital Rights Ireland Ltd v. Minister for Communications, Marine and Natural Resources*, Cases C-293/12 and C-594/12 (2014).

<sup>31</sup> Evan Engstrom, executive director of the startup advocacy organization Engine, co-authored an op ed on point with me. *Only giant internet firms may be able to comply with one-size-fits-all rules*, San Francisco Chronicle (2018), <https://www.sfchronicle.com/opinion/openforum/article/Only-giant-internet-firms-may-be-able-to-comply-12905469.php>.

<sup>32</sup> Jos Poortvleig, *240 EU Businesses Sign Open Letter Against Copyright Directive Art. 11 & 13*, NextCloud (2019), <https://nextcloud.com/blog/130-eu-businesses-sign-open-letter-against-copyright-directive-art-11-13/>.

<sup>33</sup> Paul Sawers, *YouTube: We've invested \$100 million in Content ID and paid over \$3 billion to rightsholders*, VentureBeat (2018), <https://venturebeat.com/2018/11/07/youtube-weve-invested-100-million-in-content-id-and-paid-over-3-billion-to-rightsholders/>.

<sup>34</sup> Matthew Le Merle et al., *The Impact of U.S. Internet Copyright Regulations on Early-Stage Investment: A Quantitative Study*, Booz & Company; Oxera, *The Economic Impact of Safe Harbours on Internet Intermediary Start-Ups* (2015).

<sup>35</sup> Christina Angelopoulos et al., *Study of Fundamental Rights Limitations for Online Enforcement Through Self-Regulation* 50–51 (2016), <https://perma.cc/8QAW-79QT>; Aleksandra Kuczerawy, *The Power of Positive Thinking: Intermediary Liability and the Effective Enjoyment of the Right to Freedom of Expression* (2017), J. INTELL. PROP. INFO. TECH. & E-COM. L., [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3033799](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3033799).

with this issue by including rights of appeal to courts or regulators.<sup>36</sup> This shift, which effectively exposes platforms to litigation for both over- and under-policing, forfeits some of the laws' benefits for innovators. In exchange, it should in theory more effectively protect users' rights – although the efficacy and practicality of Article 17's after-the-fact corrections remain to be seen.

### C. Inadequate Public Information about Filter-Based Removal Systems

Understanding Article 17's real benefit for harm-prevention, and real costs for speech, innovation, and other important values, requires answering some very concrete questions about filtering systems' real-world performance. Policymakers need to understand both filters' technical capability to discern lawful and unlawful content, and the realistic chances that platform employees will meaningfully correct for errors by reviewing filters' output or users' appeals. As the ACLU and other groups wrote in their letter to the European Parliament, though, “lawmakers and the public have no meaningful information” about how well such systems work now, or could work in the future.<sup>37</sup> As I will emphasize in this final portion of my testimony, no commercial actor in this space – including major platforms like Facebook and Google – has adequate incentives to provide that information.

What little we know from studies of filters made available for independent review is not encouraging.<sup>38</sup> Even representatives of Audible Magic, which is perhaps the highest profile filtering vendor, have said in the EU stakeholder dialogue that determinations about parody, criticism, and other contextually-legal uses “must be handled by human judgment[.]” The

---

<sup>36</sup> Other examples include Mexico's *La Fortuna* case, *supra* note 13, and the EU's revised Audio Visual Media Services Directive, which allows users of video hosting platforms like YouTube to appeal certain takedown decisions – not including those based on copyright – to media regulators. Directive (EU) 2018/1808, amending Directive 2010/13/EU, Art. 30.

<sup>37</sup> *Supra* note 23.

<sup>38</sup> Center for Democracy and Technology, *supra* note 28 (finding accuracy rates of 70-80% for commercially available natural language processing filters); Evan Engstrom and Nick Feamster, *The Limits of Filtering* (2017) (examining filters including the open-source audio fingerprinting tool used by Spotify), <http://www.engine.is/the-limits-of-filtering>.

company disclaimed any representation that “technology can solve this problem in an automated fashion.” Facebook, similarly, stated that its filters are “not able to take context into account.”<sup>39</sup>

That shortcoming puts all of the responsibility for applying real judgment about users’ online speech onto platform employees. This has real costs for innovation: it is unclear which smaller platforms can even afford the additional hiring to take on this task. And it has real costs for speech and user rights. Decisions about whether to follow filters’ recommendations will presumably be made by the same employees who, understandably but routinely, take down the wrong thing in existing notice and takedown systems. They will still have incentives to err on the side of removal. The difference is that these workers will now be the first and only source of human judgment assessing a firehose of machine-generated matches. For a filter assessing the half-million posts users make on Facebook each minute, with a hypothetical (and remarkably low) error rate of 0.1%, that firehose would include 720,000 misjudgments each day.<sup>40</sup> Some experts have predicted that human review will provide little more than a “rubber stamp” for machines’ decisions.<sup>41</sup> While appeals by users could in theory offset the harm to lawful speech from a “take down first, ask questions later” approach, experience with existing counternotice systems gives little reason for optimism.<sup>42</sup>

The truth is that no one in the public discussion really knows how common filtering failures are, how often human review corrects them, how effective counternotice systems are, or what patterns of bias or disparate impact may infect these systems. European lawmakers enacting Article 17 got much of their information from private companies, including rightsholders, platforms, and vendors of filtering software. This body should not repeat that mistake. Major commercial actors in this space do not have adequate motivation to be frank about filters’ shortcomings. Importantly, this includes the platforms themselves. Facebook and Google, while framed as the “anti-filtering” voices in the debate, are in the business of building algorithms to assess, “understand,” and categorize human communications, including for the purpose of

---

<sup>39</sup> Paul Keller, *Article 17 stakeholder dialogue: What have we learned so far?*, Communia (Jan 6 2020), <https://www.communia-association.org/2020/01/21/article-17-stakeholder-dialogue-day-5-depends/>.

<sup>40</sup> Facebook, *Company Information*, <https://newsroom.fb.com/company-info/>.

<sup>41</sup> Ben Wagner, *Liabile, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems*, Policy & Internet (2019), <https://doi.org/10.1002/poi3.198>.

<sup>42</sup> See *supra* note 43.

serving advertisements. Optimistic and buzzword-laden statements from Mark Zuckerberg and others about the future of content moderation using Artificial Intelligence and Machine Learning should be understood in this light. Platforms cannot be considered disinterested sources of information about technologies that lie at the heart of their own future business plans.

Major platforms' motivations in the filtering debate are further complicated by their market position and current policy struggles. Most obviously, legal filtering mandates are widely expected to advantage incumbents at the expense of smaller competitors. Platforms with ongoing licensing relationships with major rightsholders also have reason to reassure their business partners about filters' efficacy – not to emphasize their shortcomings. And for big companies with many regulatory irons in the fire, “coming to the table” to negotiate filtering or other content restriction mandates may buy much-needed political credit as they attempt to avoid or reduce the burden of new regulation in areas ranging from privacy to competition to taxation.

One European NGO, after listening as Google, Facebook, and other industry stakeholders “extolled the virtues of their content matching algorithms, claiming negligible numbers of false positives,” reached the dire conclusion that “both rightsholders and big platforms have no interest” in safeguarding user rights.<sup>43</sup> I don't believe that's true. Many of my former colleagues at Google cared deeply about Internet users' rights. I expect that their equivalents at Facebook and at rightsholder companies do as well. But the bottom line is that these are for-profit companies, and key parts of their business benefit from high expectations about the capabilities of automated content management technologies. The content and technology industries on “both” sides of the filtering debate do not come close to representing all sides of the issue, and should not be the public's primary sources of information about filters and takedown systems. If lawmakers want meaningful information about the real trade-offs involved in laws like Article 17, they must cast a wider net.

---

<sup>43</sup> Paul Keller, *supra* note 39; *Article 17 stakeholder dialogue (day 5): It all depends* (Jan 21 2020), <https://www.communia-association.org/2020/01/21/article-17-stakeholder-dialogue-day-5-depends/>.

#### IV. Conclusion

Thank you again for the invitation to speak. I welcome any questions about ways to design intermediary liability laws to optimize the balance between the legitimate goals of harm prevention, speech protection, and promotion of innovation.

\*\*\*

#### Appendix A: List of Attached Publications

- Daphne Keller, *Intermediary Liability 101: Foundations and Emerging Issues 2020* (presentation deck), <http://cyberlaw.stanford.edu/blog/2020/01/intermediary-liability-101-update-2020>.
- Daphne Keller, *Build Your Own Intermediary Liability Law: A Kit for Policy Wonks of All Ages*, Balkinization (2019), <https://balkin.blogspot.com/2019/06/build-your-own-intermediary-liability.html?m=1>.
- Daphne Keller, *Internet Platforms: Observations on Speech, Danger, and Money*, Hoover Institution Aegis Series Paper No. 1807 (2018), [https://www.hoover.org/sites/default/files/research/docs/keller\\_webreadypdf\\_final.pdf](https://www.hoover.org/sites/default/files/research/docs/keller_webreadypdf_final.pdf).
- Daphne Keller and Paddy Leerssen, *Facts and Where to Find Them: Empirical Research on Internet Platforms and Content Moderation*, forthcoming in Nathaniel Persily and Joshua Tucker, eds, *SOCIAL MEDIA AND DEMOCRACY*, Cambridge U.P., [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3504930](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3504930).
- Daphne Keller, *Dolphins in the Net: Internet Content Filters and the Advocate General's Glawischnig-Piesczek v. Facebook Ireland Opinion*, Stanford Center for Internet and Society White Paper (2019), <https://cyberlaw.stanford.edu/files/Dolphins-in-the-Net-AG-Analysis.pdf>. (A subsequent piece examining the CJEU's ruling and remaining fundamental rights issues to be examined by Member State courts, *Facebook Filters, Fundamental Rights, and the CJEU's Glawischnig-Piesczek Ruling*, is forthcoming in the GRUR International Journal of European and International IP Law.)