



Information Society Project
Yale Law School

New Controversies in Intermediary Liability Law
Essay Collection

Curated by the
Wikimedia/Yale Law School Initiative on Intermediaries and Information

Edited by Tiffany Li

Spring 2019

Table of Contents

Introduction.....	2
Faith in Filters and the Fate of Safe Harbors	3
It’s Not About What You Know: An Overview of Hyperlink Law’s Troubles.....	5
To Monitor or Not to Monitor? The Uncertain Future of Article 15 of the E-Commerce Directive .	7
A Facebook Supreme Court?.....	10
Want to Kill Facebook and Google? Preserving Section 230 Is Your Best Hope.....	13
Facebook and the European Elections: Overzealous or Uninformed?.....	15
Why There Is No Due Process Online?.....	17
Build Your Own Intermediary Liability Law: A Kit for Policy Wonks of All Ages	20
Privatizing Censorship	24
Expand Intermediary Liability to Protect Reality Itself	27

Introduction

Tiffany Li

What do we talk about when we talk about intermediary liability? Intellectual property laws likely come to mind first – the DMCA, Section 230, and so on. Fresh information ecosystem challenges like content moderation issues and the spread of “fake news” rise to the forefront as well. There are also growing concerns around privacy and cybersecurity issues, as well as the specter of overreaching tech platform governance.

Depending on whom you ask, all of the above may fall under the larger umbrella of issues related to “intermediary liability,” or only a number of them may be directly relevant. As the internet and intermediaries become more important in modern society, a growing number of topics could qualify as related to the liability (or responsibility) of internet intermediaries. Indeed, it is time to recognize that the field of intermediary liability law has significantly grown since its arguable beginnings in a few discrete intellectual property laws. In other words, when we talk about intermediary liability, we must recognize how far and widespread the field has grown.

This essay collection, “New Controversies in Intermediary Liability Law,” explores the contemporary state of intermediary liability law. It features diverse perspectives on the most pressing and current intermediary liability issues, highlighting the different topic areas that matter most today. Annemarie Bridy writes on filters and the fate of safe harbors. Aleksandra Kuczerawy explores the uncertain future of Article 15 of the E-Commerce Directive. Jacob Rogers explains the problems with hyperlinking laws. Anupam Chander considers different models for a “Facebook Supreme Court.” Eric Goldman offers a pitch for preserving Section 230 to protect market entrants. Amélie Heldt delves into the role of Facebook in spreading and controlling misinformation related to E.U. elections. Martin Husovec questions the lack of online due process. Michael Karanicolas warns about the threat of privatized censorship through platforms. Daphne Keller provides a policy tool kit for crafting new intermediary liability laws. And finally, I conclude the collection with some remarks on future directions for intermediary liability law and legal scholarship.

We are honored to present this series of interesting and informative essays on *Balkanization*, with the gracious support of Professor Jack Balkin. We will also publish these essays as a publicly available collection on the Information Society Project website when we have concluded the series. In addition to highlighting the diversity of intermediary liability issues relevant today, we hope this essay collection will further the important discussions currently being had on how best to protect the open internet, promote access to information, and develop and maintain a healthy online environment.

This essay collection is a project of the [Wikimedia/Yale Law School Initiative on Intermediaries and Information](#), which is funded in large part by a generous grant from the [Wikimedia Foundation](#). We thank the Wikimedia Foundation and the [Information Society Project](#) for their support, as well as Rebecca Crootof and Jack Balkin for their substantial efforts in bringing this project to fruition, and of course, all our essay authors for their contributions to this collection.

See published version here: <https://balkin.blogspot.com/2019/05/introducing-new-controversies-in.html>. Tiffany Li is a Resident Fellow at the Information Society Project, where she leads the Wikimedia/Yale Law School Initiative on Intermediaries and Information.

Faith in Filters and the Fate of Safe Harbors

Annemarie Bridy

The challenge of keeping harmful and illegal content off the Internet is as old as the Internet itself. Meeting that challenge, however, has never felt as urgent as it feels now. And technology companies have never felt so pressured to figure out how to do it quickly, at scale. Facebook CEO Mark Zuckerberg recently assured members of Congress that advances in machine learning over the next few years will improve and more fully automate what he admits has been a deeply flawed process for removing banned content from Facebook. In an [op-ed](#) published in *The Washington Post*, Zuckerberg actually recommended federal legislation requiring online platforms to “build systems” that block unwanted speech.

Whereas Zuckerberg is relatively new to the filtering faith, the music and film industries have long extolled the virtues of enforcing copyrights online through [automated content recognition](#) (ACR) technology. For the better part of the last fifteen years, these industries have been arguing that the Digital Millennium Copyright Act’s reactive framework for removing infringing user-generated content (UGC) from online content-sharing platforms is woefully inadequate, and that such platforms should be required to deploy proactive “technical measures” for preventing copyright infringement. Music industry lobbyists point to YouTube’s voluntarily-implemented [Content ID system](#) as proof that filtering technology is available and affordable. If YouTube is already filtering, they argue, why not just make it a legal requirement for everyone?

The music industry views a statutory filtering mandate as the key to capturing revenue now lost in what they call the “[value gap](#)” between what YouTube pays to license copyrighted music and what on-demand streaming services like Spotify pay. In the [United States](#) and the [European Union](#), YouTube and other UGC-sharing platforms have historically been protected by statutory safe harbors that insulate them from liability for users’ infringements, as long as they comply with rightholders’ takedown requests. Safe harbors give UGC-sharing platforms the legal cover they need to provide open forums for public expression.

Because safe harbors in their original form put the burden of monitoring for infringement on rightholders, YouTube has had no regulatory incentive to assume that burden. It has, however, had a business incentive to offer the music industry’s big players access to Content ID in return for licenses to popular content. The terms of those licenses have been negotiated in the shadow of the safe harbors and include what rightholders believe are unfair ad revenue splits for views of infringing UGC videos that rightholders use Content ID to monetize instead of blocking. At the end of the day, the music industry doesn’t want infringing UGC kept off YouTube. That would mean giving up a prime market that has so far been worth more than [six billion dollars to them in ad revenue](#). What the music industry wants is a bigger share of YouTube’s pie, and it aims to get that by convincing policy makers to alter the regulatory incentives around monitoring for all services that allow users to publicly share content.

Now that [streaming has supplanted paid downloads](#) as the dominant format for digital music delivery, the music industry wants all platforms that stream copyrighted content to be treated equally under copyright law—regardless of the fact that dedicated music platforms like Spotify don’t host UGC at all and therefore don’t face the open-ended legal risk that safe harbors are designed to limit. Nor do closed platforms like Spotify offer the general public open-ended opportunities for self-expression and creative production. Because UGC platforms are open to all comers, they cannot possibly

proactively license the entire universe of copyrighted content their users might ever upload. That's why safe harbors exist, and why they have historically placed the burden of monitoring for infringements *ex post* on rightholders.

Wealthy tech giants like YouTube and Facebook can likely afford to bear the legal risk associated with narrowed safe harbors. And they can afford to bear the high cost of operating sophisticated ACR systems in terms of both technological and human resources. (In a 2018 [report](#) on the company's anti-piracy efforts, Google said it has invested \$100M in building and operating Content ID.) Emerging and smaller online businesses lack such resources, however. Constricting safe harbors through *de facto* or *de jure* monitoring obligations for platforms could therefore substantially limit dynamism at the Internet's now highly concentrated edge, where consumers find themselves locked in to mega-platforms with few competitors. Copyright policy adjustments aimed at redistributing wealth from Big Tech to Big Music risk the unintended consequence of further entrenching the few that can "pay to play" under a tightened liability regime.

As the U.S. Copyright Office [mulls recommending changes](#) to the scope of the DMCA safe harbors, and EU member states prepare to transpose Article 17 (formerly Article 13) of the controversial [Digital Single Market \(DSM\) Copyright Directive](#) into domestic law, now is a good time to take a hard look at whether it makes sense to hardwire ACR systems like Content ID into copyright law through "notice-and-staydown" requirements.

Many urgent questions arise: Should the wide universe of UGC services that have flourished for two decades under the protection of safe harbors lose that protection so that Big Music can secure bigger payouts from Big Tech? Is the public's interest served by changes to copyright law that could exponentially elevate operating risk for all online services that allow users to share content? Should copyright safe harbors be conditioned implicitly or explicitly on platforms' implementing ACR systems? Are such systems accessible and sustainable for services that lack YouTube's resources? Are ACR systems fit for purpose when it comes to protecting *lawful* expression, including fair use of copyrighted material? If not—and we do have [ample evidence of their limits](#)—how strongly should that militate against public policies requiring or encouraging broader deployment?

The public needs and deserves evidence-based answers to these questions before new laws favoring or requiring deployment of ACR systems are enacted. In Europe, these questions are now largely moot, given [parliamentary approval](#) of the DSM Copyright Directive. In the United States, however, the conversation about potential modifications to the DMCA is only getting started. It is seductive to look for technological solutions to content-related problems on massive platforms like Facebook and YouTube. Given the urgency and the scale of some of those problems, it is in the interest of both the platforms themselves and policy makers to put their faith in a quick technological fix. The public, however, should be skeptical, because the competitive and expressive costs of making UGC platforms filter everyone's speech before it can be shared will be profound.

See published version here: <https://balkin.blogspot.com/2019/05/faith-in-filters-and-fate-of-safe.html>. Annemarie Bridy is the Allan G. Shepard Professor of Law at the University of Idaho College of Law, Affiliate Scholar at the Stanford Law School Center for Internet and Society, and Affiliated Fellow at the Yale Law School Information Society Project. Professor Bridy specializes in intellectual property and information law, with specific attention to the impact of new technologies on existing legal frameworks for the protection of intellectual property and the enforcement of intellectual property rights.

It's Not About What You Know: An Overview of Hyperlink Law's Troubles

Jacob Rogers

The law related to hyperlinks is breaking. Hyperlinks are a system for directing readers around the internet, but the rules vary by the location of whoever makes the link. People use the links once and move on, but the rules seem headed towards active link monitoring. The internet treats hyperlinks as all of a kind, but courts are breaking them down by type, content, and specialty.

In the metaphor of the internet as highway, hyperlinks are the forks, side roads, and driveways into the unknown for the online traveler. Sometimes they are neglected and become overgrown. Other times they lead to a hidden gem. Yet other times still there's a massive pothole the second you turn the corner. Most often, hyperlinks provide valuable information for the explorer, directing them to sources, background, context, or the endless rabbit holes of exploring related topics. Anyone that has come up for air after a few hours on a [deep wiki-walk](#) is familiar with this use of hyperlinks and the way they can make it easy to traverse the internet. A [second type of hyperlink](#) is used for humor, surprising the reader with the result or [the mouseover text](#) previewing where the link is going. Some hyperlinks can be embedded via various technologies to appear on the page to show a preview of what is coming without the need to leave the current URL, and at times the presence or absence of that preview depends on the user's settings (such as whether their browser allows a website to load images). A small handful of hyperlinks are harmful as well, tricking the reader into intentionally visiting a site with malware or spyware that aims to steal information or wreck their computer.

The law, however, does not distinguish effectively between the many ways hyperlinks are used online. Rather, the law on hyperlinking has split by region in different ways. In the United States, the law is currently in flux and has focused mostly around liability for copyrights. [Perfect 10 v. Amazon](#) established what's known as the "server test." In *Perfect 10*, Google thumbnail images that acted as hyperlinks linking back to full size originals were found not to infringe copyright law, in significant part because the full size images were not on Google's servers and therefore had not been copied. However, in a recent case, [Goldman v. Breitbart](#), the court held that the fact that a picture of Tom Brady was visible on a Breitbart web page to the reader was enough for copyright infringement. (The link in this case was an "embedded" image, meaning that the reader would see the image on Breitbart without Breitbart making a copy of it. Technologically, that occurred by instructing the user's browser to access it from Twitter directly.) It's currently unclear where U.S. law is headed and what the legal basis might be for linking to images.

In Europe, the law attempts to distinguish the lawfulness of a hyperlink based on the knowledge of the person who created the link, following the recent [GS Media](#) case. However, *GS Media* overwhelms its knowledge standard with an additional rule that website creators who attempt to make money must carry out an investigation of what they link to and will be presumed to have knowledge of the content behind a hyperlink. While this does not affect every site, such a large number of sites either have minor advertising or encourage users to make some kind of purchase that it impacts a tremendous part of the internet. Europe also has a further complication for copyrighted works in which they consider whether and where the work was already accessible online. If something was already freely available and a link does not meaningfully change who can access it, it may not violate copyright because there was not a publication to a "[new public](#)."

These various standards are problematic because they change hyperlinks from something easy to use into something complicated. While judicial attempts to force changes in behavior can sometimes be effective, hyperlinks are widely used by the general public in a way that is not consistent with either U.S. standards for copyrighted works or European standards for presumed knowledge. The likely result of these cases is either that many people will violate the law unknowingly or that large companies will be forced to implement various blunt technological measures to limit use of hyperlinks, harming online discourse and greatly adding to the difficulty of finding smaller websites that are not already well-known. U.S. law, in particular, is making it increasingly more difficult to share media online without paying licensing fees of some sort, despite the fact that a substantial majority of the population does share all types of media constantly from site to site.

GS Media's knowledge standard and *Svensson's* “new public” ideas might be on the right track if they were made to more closely fit existing industry practices and consumer expectations, rather than try to change the behavior of the public. We *do* trust that someone posting a link is not attempting to harm our computers or luring us into committing crimes. Therefore it can be reasonable to hold someone liable for their links if the link and its context show that they knew (or clearly should have known) that they were leading people to something harmful. But we do not and courts should not require someone making a hyperlink to investigate broadly or continually monitor their hyperlinks to ensure that nothing changes in a way that becomes illegal. Nor should site owners be held strictly liable because they make money when context and content do not make it clear that the owner meant actual harm. As the roads of the internet, hyperlinks should be treated akin to road construction: a construction crew might be liable for building a faulty road, but they are not liable years later when the owner of the property allows it to become overgrown.

See published version here: <https://balkin.blogspot.com/2019/05/its-not-about-what-you-know-overview-of.html>. Jacob Rogers is Senior Legal Counsel at the Wikimedia Foundation. His work includes international litigation, government requests and Trust & Safety work for the Wikimedia Foundation across multiple jurisdictions.

To Monitor or Not to Monitor? The Uncertain Future of Article 15 of the E-Commerce Directive

Aleksandra Kuczerawy

Current policy discourse in the European Union is steadily shifting from intermediary liability to intermediary responsibility. The long-established principle prohibiting general monitoring obligations is currently being challenged by two initiatives in particular, namely the Copyright in the Digital Single Market Directive and the proposal on the Regulation preventing the dissemination of terrorist content online. By holding service providers liable and requiring broad implementation of censoring measures, this growing trend toward imposing monitoring obligations may have significant ramifications for the ability of individuals to freely share and access content online.

No General Monitoring Obligation

The limited liability regime for Internet intermediary services in the European Union is currently going through major changes. To date, the liability exemptions for Internet intermediary service providers have been governed by [E-Commerce Directive](#) (ECD). The Directive applies horizontally to various domains and to any kind of illegal or infringing content.

Under Article 15 of the ECD, E.U. Member States may not impose on intermediary service providers a general obligation to monitor information that they transmit or store. Further, Member States cannot introduce a general obligation to actively look for facts or circumstances indicating illegal activity. The prohibition of monitoring obligations (Recital 47) refers solely to monitoring of a general nature. It does not concern monitoring obligations in a specific case; nor does it affect orders issued by national authorities in line with national legislation.

The ECD, moreover, does allow Member States to require hosting providers to apply [duties of care](#). Such duties of care, however, should only be introduced to detect and prevent certain types of illegal activities, foreseen by national law. The Directive does not specify what exactly such duties of care entail. As a result, the [boundary](#) between duties of care and general monitoring is not clear.

Per Article 15, Member States are not allowed to introduce obligations that would require intermediary service providers to systematically monitor the information they store or transmit. This does not mean that service providers cannot take up such activities on their own initiative. Most of the service providers in the European Union do perform certain voluntary monitoring activities in order to maintain a “civilized” environment on their platforms. Voluntary monitoring, however, can prove detrimental. Exercising too much control could compromise the neutral status of the intermediary service provider and, in consequence, deprive them of the safe harbour protection provided by Article 14 of the ECD. The E.U. intermediary regime does not contain a provision which protects service providers from liability should their voluntary monitoring prove imperfect (such as the one offered by the Section 230(c)(2) of the Communications Decency Act in the United States). The lack of a Good Samaritan-style protection results in a certain level of prudence on the side of the E.U. service providers when administering the monitoring activities on their platforms.

Current policy discourse in the European Union is steadily shifting from intermediary liability to intermediary responsibility. The amended Audiovisual Media Services Directive, the Directive on

copyright in the Digital Single Market, and the proposal for a Regulation on preventing dissemination of terrorist content online, as well as recent soft law initiatives, such as the Code of Conduct on Countering Illegal Hate Speech Online and the most recent Code of Conduct on Disinformation, are all examples of that trend.

The former is understood as a negligence-based approach, while the latter emphasizes the need for proactive measures. The long-established principle prohibiting general monitoring obligations is currently being challenged by two initiatives in particular, namely the Copyright in the Digital Single Market Directive and the proposal on the Regulation preventing the dissemination of terrorist content online.

Copyright in the DSM Directive

In mid-February 2019, at the end of the so called “trilogue” procedure, the negotiating E.U. institutions reached a [compromise](#) on the text of the Copyright in the DSM Directive. The [text](#) passed the vote by the European Parliament on 26 March 2019 and was [ultimately approved](#) by the Council on 15 April 2019. Article 17 (formerly Article 13) of the Copyright in DSM Directive includes a [substantial change](#) to the established intermediary liability regime, in that it makes service providers directly liable for the content uploaded by their users. To avoid liability, service providers must enter into licensing agreements with any owners of any content they may possibly host. Per Article 17.4b, if the service provider does not want to (or is not able to) pay licensing fees, they have to demonstrate they made best efforts to ensure the works are unavailable “in accordance with high industry standards of professional diligence.” Only then may they be able to escape direct liability for the content of their users. The Copyright in DSM Directive stipulates that the application of Article 17 “shall not lead to any general monitoring obligation.” Interestingly, the final text of the provision no longer refers specifically to Article 15 of the ECD. The provision clearly aims to pacify the [numerous critics](#) of the Directive. Despite the insistence that Article 17 will not lead to general monitoring obligation, it is hard to imagine in what other way service providers can ensure copyrighted works are not made available without proper licensing. To [effectively recognize](#) infringing content, a technological tool must be used to examine all newly uploaded content on the platform and comparing it with an existing database. This [amounts](#) to installing upload filters by the service providers and systematic monitoring of the entirety of the users’ content. Despite repeated attempts to convince the broad public that the Copyright in DSM Directive was not meant to introduce upload filters, several [officials](#) admitted, soon after the vote, that the filters are [unavoidable](#).

Proposal for a Regulation on Preventing Dissemination of Terrorist Content Online

In September 2018, the European Commission issued a [proposal](#) for a Regulation on preventing dissemination of terrorist content online. Apart from one-hour content removal orders (Article 4) and content referrals (Article 5), the proposed regulation provides that service providers should take “proactive measures to protect their services against the dissemination of terrorist content” (Article 6). According to the European Commission’s proposal in Article 6, service providers are expected to check against publicly or privately-held tools containing known terrorist content. They may also use “reliable technical tools to identify new terrorist content,” either using those available on the market or those developed by the service provider. If the competent authority considers the taken measures insufficient, it may request that the provider takes “specific additional proactive measures.” If no agreement can be reached, the competent authority has the power to impose “specific additional (...) proactive measures.”

Recital (19) of the European Commission’s proposal states that imposing “specific proactive measures should not, in principle, lead to the imposition of a general obligation to monitor,” as prohibited by Article 15 of the ECD. After this optimistic note, the proposal explains that in light of the “particularly grave risks associated with the dissemination of terrorist content,” the decisions adopted on the basis of the Regulation could, in fact, derogate from Article 15’s prohibition. The derogation would apply to “certain specific, targeted measures, the adoption of which is necessary for overriding public security reasons.” It would seem, therefore, that a general monitoring obligation may very well be the intended outcome of the Regulation, in [contradiction](#) with Article 15.

On 17 April 2019, the European Parliament adopted its [Report](#) on the proposed Regulation. The Parliament Report made significant changes to the whole proposal and to Article 6, which no longer mentions “proactive” measures. Instead, Article 6 now states that the service providers “may take specific measures” that should be “effective, targeted and proportionate.” The amended Article 6.4 provides that competent authority “may send a request for necessary, proportionate and effective additional specific measures” to hosting providers that have received a substantial number of removal orders. However, the competent authority “shall not impose a general monitoring obligation, nor the use of automated tools.” The Parliament Report also amended Recital (19) of the proposed Regulation. The new version of the recital does not contain the worrying statement that Article 15’s prohibition could be derogated from due to overriding public security reasons.

The changes introduced by the Parliament Report [eliminate](#) the most significant threats to the Internet freedoms. The issue, however, is not yet resolved. The proposed Regulation will continue to proceed through the “trilogue” procedure, where proactive measures may still be reintroduced by the Council of the European Union.

Outlook

According to the Court of Justice of the E.U., as described in [Scarlet v. SABAM](#) and [SABAM v. Netlog](#), a requirement to install a filtering system capable of identifying specific types of content, for almost all information stored by the users, applied indiscriminately to all of them, as a preventive measure, and for unlimited period of time, amounts to a general monitoring obligation. It is difficult to imagine how the “best efforts” and “proactive measures,” as envisaged by the Copyright in DSM Directive and the Terrorist Content Regulation respectively, would not constitute a general monitoring obligation. Despite multiple warnings that the proposed measures would undermine the E.U. acquis, policy makers are still moving forward. One therefore cannot help but wonder if the times of no general monitoring obligations are coming to an end. By holding service providers liable and requiring broad implementation of censoring measures, this growing trend toward imposing monitoring obligations may have significant ramifications for the ability of individuals to freely share and access content online.

See published version here: <https://balkin.blogspot.com/2019/05/to-monitor-or-not-to-monitor-uncertain.html>. Dr. Aleksandra Kuczerawy is a postdoctoral researcher at the Centre for IT and IP Law at the KU Leuven, Belgium. She is the author of INTERMEDIARY LIABILITY AND FREEDOM OF EXPRESSION IN THE EU: FROM CONCEPTS TO SAFEGUARDS (Intersentia, 2018).

A Facebook Supreme Court?

Anupam Chander

To borrow Churchill's line about democracy, a Facebook Supreme Court is the worst idea, except for all the others.

In 2018, Mark Zuckerberg introduced the idea of establishing an independent board that would make the most difficult decisions with respect to content. He compared the new body to a [Supreme Court](#). In January 2019, Nick Clegg, Facebook's Vice-President of Global Affairs and Communications, announced the charter of this independent oversight board. Facebook may have sought to reduce apprehensions of its growing global power by ceding some control to an outside body. However, it was clear that Facebook was borrowing the apparatus, and even the personnel, of government: not only was Facebook implementing a pseudo-judicial body, but Nick Clegg had once served as the Deputy Prime Minister of the United Kingdom.

As [Dawn Nunziato](#) has observed, the internet represents the “most powerful forum for expression ever created.” How decisions over content are made on one of the internet's principal platforms—a platform that connects literally billions of people—is of great importance. Scholars such as [Jack Balkin](#), [Tarleton Gillespie](#), [Daphne Keller](#), [Kate Klonick](#), [Thomas Kadri](#), and [Sarah Roberts](#) have powerfully analyzed how internet platforms make decisions on the content that is carried on their sites and the role of intermediaries in free expression today. In my scholarship, I have sought to demonstrate a “[First Amendment/?Cyberlaw Dialectic](#)” in which “the First Amendment constituted cyberlaw, and cyberlaw in turn constituted free speech.”

Facebook's many critics argued that such an oversight board would serve principally as window dressing, or that the introduction of the outside mechanism was merely rearranging the deck chairs on the Titanic. Yet, the Facebook Oversight Board marks a major new experiment. This essay compares the Oversight Board with its alternatives. After all, the Oversight Board must be considered not only on the basis of its flaws—of which there will likely prove to be many—but rather in comparison to its alternatives.

I will consider five alternatives, which I will dub: Mark Decides; Democracy by Likes; Feudal Lords; Judge Robot; and Official Censorship.

Mark Decides

In December 2015, some employees inside Facebook were troubled. A candidate for U.S. president was calling for a ban on immigration for people of a particular religion. Users had flagged the content as hate speech, triggering a review by Facebook's community-operations team, with its many employees in several offices across the world. In [internal messages](#), some Facebook employees declared that posts constituted hate speech. On its face, a statement targeting a particular religious group for a ban on immigration would seem to violate Facebook's community guidelines.

Facebook's head of global policy management, Monika Bickert, [explained internally](#) “that the company wouldn't take down any of Mr. Trump's posts because it strives to be impartial in the election season.” Facebook [explained](#) its decision to the public as follows: “In the weeks ahead, we're going to

begin allowing more items that people find newsworthy, significant, or important to the public interest—even if they might otherwise violate our standards.”

According to the [Wall Street Journal](#), “The decision to allow Mr. Trump’s posts went all the way to Facebook Inc. Chief Executive Mark Zuckerberg, who ruled in December that it would be inappropriate to censor the candidate.” Zuckerberg is the ultimate arbiter of what stayed up or what came down on Facebook.

A central difficulty of Facebook’s current model is that it places enormous power in the hands of Facebook’s leadership and to those employees to whom the company leaders choose to delegate this power. When Facebook deleted a Norwegian government minister’s posting of the famous photograph of the naked Vietnamese girl fleeing a napalm attack, the Norwegian government complained of censorship. Facebook [reversed](#) its decision, even though its community guidelines banned nudity. (“While we recognize that this photo is iconic, it’s difficult to create a distinction between allowing a photograph of a nude child in one instance and not others,” a spokesman for Facebook said in response to queries from the [Guardian](#).) In the wake of the censorship of the photograph, Espen Egil Hansen, the editor-in-chief and CEO of Norway’s largest paper, would [declared](#) that Zuckerberg was “the world’s most powerful editor.”

Democracy by Likes

What if Facebook put governance decisions to a vote, asking people to like or dislike a particular post? We do not typically solve controversies over a particular statement through popular vote. This may be because such a mechanism might often degenerate into a contest focused on the popularity of the controversial content, rather than a reasoned assessment of whether the content violated the community guidelines. This would have the effect of reinforcing popular views at the cost of minority viewpoints.

Feudal Lords

What of Reddit-style moderators, granted the authority to regulate particular discussions, charged with the authority to remove posts they found to be a violation of that group’s guidelines? This essentially becomes a kind of dispersed version of Mark Decides—instead of a single king, multiple lords. If only a few “lords” have power, this approach raises the same concentrated power issues of Mark Decides. If there are many “lords,” however, there might be enough alternatives that controversial content might find some home somewhere. Thus, such an approach might only shift the material to different corners of Facebook, rather than removing material that genuinely violates Facebook’s community guidelines.

Judge Robot

Perhaps we could rely upon a computer to make content decisions. In fact, of course, if Facebook is making millions of content decisions each day, it is likely relying in significant part on AI. Natasha Duarte, Emma Llansó, and Anna Loup [have argued](#), however, that “large-scale automated filtering or social media surveillance [...] results in overbroad censorship, chilling of speech and association, and disparate impacts for minority communities and non-English speakers.” As [Duarte et al.](#) describe, decision-making by AI would not result in unbiased decisions, but rather would potentially amplify bias. I have described this in my scholarship as a kind of “[viral discrimination](#).”

Official Censorship

There is reason to worry about the great private power concentrated in companies like Facebook, which reserves the right to delete information that violate its community guidelines. An alternative is to vest such decision-making in traditional governments—either through courts or administrative bodies. Controversies over content would then be determined by an official process, rather than by corporate employees following often secret processes and secret rules. Of course, it is unclear whether individuals would have the resources to bring or defend claims that some material should be removed. Not only may the process be expensive, it may also be slow. Furthermore, governments may use their content management powers to target negative or opposition information as “fake news.” [This concern](#) will be heightened in illiberal states. Finally, states with a strong commitment to free expression, like the United States, would find it difficult to censor content that was not illegal except under time, place, and manner restrictions that are [difficult to translate](#) to the internet.

Afterword: Facebook’s Last Experiment with Outside Decision-Making

This is not Facebook’s first experiment with ceding decision-making to outsiders. In 2009, Facebook permitted its users to vote on proposed changes to the terms of use—though on terms that made practically all such votes advisory rather than mandatory. That Facebook democracy proved short-lived.

Under the [now-defunct system](#), a site-wide vote would be triggered if proposals to modify Facebook’s terms of service received comments from at least 7,000 users. Then if at least 30 percent of active users voted on the proposal, Facebook would treat the vote as binding. Otherwise the vote would be merely advisory. Given that Facebook already had a billion users, there was little likelihood that any vote would be binding.

In December 2012, Facebook put two policy changes to a [vote](#): whether Facebook should be able to share data with Instagram and whether Facebook should end users’ rights to vote on further governance questions. 88 percent of the 668,872 voters resoundingly rejected the changes. But the vote fell far short of the 300 million or so required for a binding vote, and thus could be safely ignored. After merely three votes, Facebook ended its chimerical experiment with democracy. Perhaps Facebook’s new experiment with external governance might prove longer-lived.

See published version here: <https://balkin.blogspot.com/2019/05/a-facebook-supreme-court.html>. Anupam Chander is Professor of Law, Georgetown University; A.B. Harvard, J.D. Yale. The author is grateful to Delia Brennan and Ryan Whittington for superb research assistance, and also thankful for a Google Research Award to support related research.

Want to Kill Facebook and Google? Preserving Section 230 Is Your Best Hope

Eric Goldman

You probably feel antipathy towards Facebook and Google. Most people do. Yet, as incumbents with extraordinary amounts of wealth, it sometimes feels like Facebook and Google are impervious both to competition and regulation.

Indeed, regulators have limited tools to corral Facebook and Google in part due to 47 U.S.C. §230 (Section 230), which immunizes online services from many types of liability for third-party content they republish. This immunity supplements the First Amendment's speech and press protections, but Section 230 does more than promote free speech. The immunity also plays a major role in economic and competition policy. Section 230 implicitly provides what amounts to a financial subsidy to online republishers. That subsidy, counterintuitively, may represent our best hope for dethroning the Internet giants.

Subsidy

Section 230's immunity has established the legal foundation for the modern Internet. Virtually every major Internet service depends on Section 230 heavily to acquire, curate, and disseminate third-party content; and many of us engage with those services many times an hour. As a result, Section 230-protected services have generated an extraordinary amount of private and social economic benefits.

Section 230 is a flagship example of Internet exceptionalism, in effect, legally regulating the Internet differently than other media. Traditional legal doctrines of offline publishing usually hold publishers liable for any third-party content they choose to republish. Section 230 says the opposite; when third-party content is involved, the online republisher isn't liable for it. Thus, the legal outcome depends on the medium: the exact same content, from the same author, can create liability for offline republishers and not for online republishers.

By immunizing online republishers from liability for third-party content, Section 230 frees republishers from the costs associated with protecting against such liability. In effect, Section 230 provides an implicit financial subsidy to all Internet republishers—including Internet giants like Google and Facebook—compared to offline republishers. Superficially, that makes no sense at all. The Internet giants are among the most valuable companies to ever exist. They seem like undeserving candidates for government-mandated privileges.

Yet, the wisdom of this policy becomes clearer when realizing that, even as Section 230 privileges the Internet giants, it also plants the seeds of their future destruction. Section 230's subsidy reduces barriers to enter the online republishing marketplace, which, in turn, keeps the marketplace open for the next generation of startups that hope to usurp the current Internet giants.

Enhanced Competition

If the rules of offline publishing applied to the Internet, online republishers would implement effective measures to reduce their exposure for third-party content. Instead, due to Section 230's immunity,

online republishers of third-party content do not have to deploy industrial-grade content filtering or moderation systems, or hire lots of content moderation employees, before launching new startups. This lowers startup costs generally; in particular, it helps these new market entrants avoid making potentially wasted investments in content moderation before they understand their audience's needs. Accordingly, startups do not need to replicate Google's or Facebook's extensive and expensive content moderation operations, nor do they need to raise additional pre-launch capital to defend themselves from business-crippling lawsuits over third-party content.

In a counterfactual world without Section 230's financial subsidy to online republishers and the competition enabled by that subsidy, the Internet giants would have even more secure marketplace dominance, increased leverage to charge supra-competitive rates, and less incentive to keep innovating. In other words, without Section 230, the marketplace would ossify, and existing legal regulations would help lock in the incumbents.

Admittedly, it feels strange to tout Section 230's pro-competitive effect in light of the dominant marketplace positions of the current Internet giants, who acquired their dominant position in part due to Section 230 immunity. At the same time, it's likely short-sighted to assume that the Internet industry has reached an immutable configuration of incumbents. Internet history is filled with dominant players—Microsoft, Netscape, Yahoo, AOL, MySpace, and others—who were displaced by upstarts, often in unexpected ways by unanticipated competitors.

Similarly, Google and Facebook probably will not be dislodged by head-on competitors launching a comprehensive keyword-driven search engine or a mass-market general-purpose social networking service. Instead, they are likely to be dislodged by indirect competitors who address consumers' needs through radically different technological or operational approaches.

Those disruptive innovators absolutely require legal immunity to grow big and popular enough to change consumer practices and gain consumer loyalty, without being swamped by lawsuits and the high costs of content moderation obligations. Section 230 is an essential piece to ensure that future Google- and Facebook-killers have a chance of emerging.

Conclusion

Focusing on the financial subsidy to the Internet giants fundamentally misunderstands Section 230. If you really want to stick it to Google and Facebook, you should fight to preserve Section 230's competition-enhancing benefits. Otherwise, you are implicitly rooting to squelch the future competitive threats they should face, which only strengthens the Internet giants' marketplace dominance.

See published version here: <https://balkin.blogspot.com/2019/06/want-to-kill-facebook-and-google.html>. Eric Goldman is a Professor of Law, and Co-Director of the High Tech Law Institute, at Santa Clara University School of Law. Before he became a full-time academic in 2002, he practiced Internet law for eight years in the Silicon Valley. His research and teaching focuses on Internet, IP and advertising law topics, and he blogs on these topics at the Technology & Marketing Law Blog [<http://blog.ericgoldman.org>].

Facebook and the European Elections: Overzealous or Uninformed?

Amélie P. Heldt

In late March 2019, Facebook announced it would tighten up its rules for political advertisement and is now – once again – under attack for getting it wrong. The attempt of the world’s largest social media platform to protect the E.U. elections from foreign interference is effectively restricting European candidates to their country of residence.

The network infrastructure is particularly suited to spread (mis)information, which makes the problem of so-called “fake news” on social media platforms hard to solve. Recent work shows how the attention economy relies on content that provokes strong emotions, and that it benefits from the data generated through user engagement to optimize behavioral advertisement. (See, for example, recent publications from [Zuboff, Crockett, Benkler et al., Ghosh and Scott.](#)) It is more likely that controversial content will be [algorithmically prioritized](#) in a user’s newsfeed, including both viral content and [false information](#) that was designed and spread with the intention to mislead the recipient. Plus, the distribution of misinformation can be amplified by buying [targeted advertisement space](#), just like with any other digital marketing campaign. This mix can become quite explosive, even more when it comes to disinformation in periods of [election campaigns](#). The accompanying effects can be threatening for democracy if fake news posts are realistic enough [to be believed](#) and if the [exposure](#) to them is high. Since Brexit and the U.S. presidential campaigns in 2016, the use of [social media marketing](#) for political purposes has become more common but the way it operates remain opaque. The dimension of this issue was revealed by the [Cambridge Analytica](#) scandal in March 2018, which got Facebook into a pretty pickle.

In an effort to fix mistakes of the past and to counter allegations it would not live up to its responsibilities, Facebook published its [new rules for political advertisement](#) in the European Union two months ahead of the election days. The two main novelties are that advertisers need to be “authorized in their country to run ads related to the European Parliamentary elections,” and that “all ads related to politics and issues on Facebook and Instagram in the EU must be clearly labeled.” With the latter, Facebook is implementing a transparency-enhancing tool that experts have been [demanding](#) for a long time. At the same time it makes Facebook the *de facto* arbitrator over what is political speech and what is not. It also does not solve the problem of adapting political communication to potential voter profiles. More urgent, however, is the problem with the first rule, which is designed to be “a real barrier for anyone thinking of using Facebook to interfere in an election from outside of a country.” However, these elections are pan-European, and so are the campaigns.

This shows that Facebook struggles with the E.U. elections, not fully considering the actual circumstances. The problem with authorizing advertisers only in their countries is that it restricts the candidates considerably in their sphere of action. The parliamentary groups within the European Parliament are composed of delegations from Member States but each belonging to the same political family. Thus, their campaigns run both on a national and a supra-national level. For example, Manfred Weber, top candidate for the European People’s Party (the largest political group in the European Parliament) and the potential next president of the European Commission, can only campaign in Germany as a member of his German party. It also means that European parties and candidates are treated as “foreign interference” by Facebook, clearly missing on the whole rationale of the European Union and its parliamentary elections (and the E.U. single market).

As [reported](#), the most senior E.U. civil servants believe that failing to recognize the role of pan-European political parties and institutions “would encroach upon fundamental EU rights and freedoms, such as free movement and political participation.” In [this letter](#) they also complained about Facebook preventing the E.U. institutions from calling on citizens to vote. (Ironically, Nick Clegg, Facebook’s head of global policy and communication, is a former British MP who should be familiar with the functioning of E.U. elections.) The representatives of eight political groups followed suit, with this [open letter](#) to Mark Zuckerberg. So far, it remains unclear if Facebook will grant only single exceptions or thoroughly change its rules. Both politicians and institutions depend on social media to reach voters that are spread out in different countries, hence not reachable by classic means like election posters. The whole electoral process is based on a free flow of information, enabling the formation of opinions. An inherent part of citizens’ electoral rights is their freedom of information, which is currently hindered by Facebook’s new policy. Just a few weeks ahead of the elections, it seems unreal that the campaigns would be seriously disrupted and it shows, once again, the immense power of Facebook over not only individual communication but also the public discourse.

See published version here: <https://balkin.blogspot.com/2019/06/facebook-and-european-elections.html>. Amélie P. Heldt is a junior researcher and doctoral candidate with the Leibniz Institute for Media Research, Hamburg, and currently a Visiting Fellow with the Information Society Project at Yale Law School.

Why There Is No Due Process Online?

Martin Husovec

Online information gatekeepers are in the spotlight. Their roles are being questioned and societal expectations reformulated daily – not only in Europe, but around the globe. However, much of the attention of regulators is biased only towards achieving removal of the objectionable content. Owing to a never-ending stream of controversies, the regulators fail to see (or worse, decide to ignore) that, as much as societies risk under-removal of illegitimate content, they also risk over-removal of legitimate speech of their citizens.

No other regulator better illustrates this mindset than parts of the European Commission. As a direct offspring of the European refugee crisis, the European Commission set up an informal agreement with technology companies to quickly remove hate-speech in May 2016. Since then, the Commission [publicly communicates](#) that the less notified content is rejected by platforms (and therefore removed), the better for all of us. It does not take an expert to recognize that this thinking assumes that underlying notifications are flawless—something that the European Commission does not evaluate in its monitoring exercise. Despite the criticism, the Commission continues to celebrate increasing removal rates as some form of ‘evidence’ of the fact that we are improving. In reality, we are far from knowing what the net positive value of this exercise is.

Academics have long argued that even the baseline system of intermediary liability, which allocates responsibilities with several stakeholders under a notice and takedown regime, is prone to over-removal of legitimate speech. Faced with potential liability, providers have a rational bias towards over-removal; they err on the side of caution. These arguments have been [proven right](#) by daily news and rigorous empirical and experimental studies.

Although some regulators have started recognizing this as an issue, many still do not think that magnitude of the problem is too severe, in particular when compared to social problems associated with failing to enforce the laws. To be fair, even academics cannot yet properly tell what the aggregate magnitude of this problem is. We can point to the gap between false positives in removals and extremely low user complaint rates at the service level, but not too much more than that. The individual stories that make up this graveyard of erroneously blocked content are mostly unknown.

To their credit, the stakeholders have successfully voiced the problem recently. Several upcoming pieces of the Union law—such as the Digital Single Market (DSM) Directive, the Terrorist Content Regulation, and the Platform to Business Regulation—now include some commitment towards safeguards against over-removal of legitimate speech. However, these are still baby steps. We are lacking a vision of how to *effectively* achieve high-quality delegated enforcement that minimizes under-removal *and* over-removal at the same time.

Article 17(9) of the DSM Directive mandates that E.U. Member States require some online platforms dealing with copyrighted content to “put in place an effective and expeditious complaint and redress mechanism that is available to users of their services.” The right holders who issue requests for removal have to justify their requests, and the platforms must use humans to review these user complaints. The Member States have to facilitate alternative dispute resolution (ADR) systems and should ensure respect for some types of copyright exceptions and limitations. The Terrorist Content

Regulation aims to prescribe such mechanisms to the hosting platforms directly. Although the Commission proposed a full reinstatement obligation for wrongly removed content, the European Parliament recently suggested to soften it towards a mere obligation to hear a complaint and explain its decision (as seen in Article 10(2) of the proposal). Article 4 of the Platform to Business Regulation prescribes that complaint processes are available for cases of restriction, suspension or termination of services of business users.

All of these initiatives, even though well-intended, show a great deal of misbalance between two sides. While the regulators are increasingly ramping up the effort to increase the volume and speed of removals, by finding more wrongful content online and blocking it more quickly, their approach is almost surgical when it comes to over-removal. They suddenly want the platforms to weigh all the interests on a case-by-case basis. While the regulators apply all pressure possible on the detection and removal side by prescribing automation, filters and other preventive tools which ought to be scalable, they limit themselves to entirely ex-post individual complaint mechanisms that can be overruled by platforms in cases of over-removal errors. When fishing for bad speech, regulators incentivize providers to use the most inclusive nets, but when good speech gets stuck in the same nets, they provide the speakers only with a chance to talk to providers one-on-one, thus giving them a small prospect of change.

We fail to create equally strong incentives for providers to avoid over-removal at scale. Without parity in incentives, delegated enforcement by providers is no equal game; and without equality of weapons, there is no due process. Even with policies like the ones currently baked in the European Union, the users (whether private or business ones) have to invest to counter false allegations. They bear the cost, although they cannot scale up or speed up their defense. Without strong ex-ante incentives for higher quality review, the cost of mistakes is always borne by the users of those platforms since the correction takes place ex-post after a lengthy process. Even if somehow legitimate speakers prevail after all, the system, by definition, defies the legal maxim that justice delayed is justice denied.

The solutions that we need might not always be that complicated. The first [experimental evidence](#) suggests that exposing platforms to counter-incentives in a form of external ADR, which also punishes their over-removal mistakes by small fees in exchange for legal certainty, can in fact reduce the over-removal bias and thereby lower the social costs of over-blocking. The logic here is simple: if platforms bear the costs of their mistakes because over-removal suddenly also has a price tag, they have more incentive to improve by investing resources into the resolution of false positives too. Moreover, since platforms can learn at scale, each mistake is an opportunity for the benefit of everyone else, thereby improving the technology and associated governance processes in the long-run.

However, to complicate things further, regulators need to find a way to strike a balance between user's expectations to share their lawful content and platform's interest to pick and choose what to carry. Treating all platforms as states by imposing must carry claims to all legal content [overshoots the target](#) to the detriment of speech. However, treating platforms as purely private players underappreciates their existing social function. We need to find a mechanism that preserves the contractual autonomy, and ability to shape communities along some values or preferences, which at the same time safeguards due process of speakers. However, due process has to mean something more than mere explanation from a human. It has to amount to credible and timely contestability of decisions, which platforms cannot simply override without too much effort.

See published version here: <https://balkin.blogspot.com/2019/06/why-there-is-no-due-process-online.html>. Martin Husovec is Assistant Professor at Tilburg University (appointed jointly by Tilburg Institute for Law, Technology and Society & Tilburg Law and Economics Center) and Affiliate Scholar at Stanford Law School's Center for Internet & Society (CIS). He researches innovation and digital liberties, in particular, regulation of intellectual property and freedom of expression.

Build Your Own Intermediary Liability Law: A Kit for Policy Wonks of All Ages

Daphne Keller

In recent years, lawmakers around the world have proposed a lot of new intermediary liability (IL) laws. Many have been miscalibrated – risking serious [collateral damage](#) without necessarily using the best means to advance lawmakers’ goals. That shouldn’t be a surprise. IL isn’t like tax law or farm subsidies. Lawmakers, particularly in the United States, haven’t thought much about IL in decades. They have little institutional knowledge about which legal dials and knobs can be adjusted, and what consequences to expect.

This post will lay out a brief menu, framed for a U.S. audience, of IL legal mechanisms. Most are relatively well-understood from laws and literature [around the world](#); a few are newly emerging ideas. It foregrounds legislative choices that affect free expression, but does not try to identify hard limits created by the First Amendment or other free expression laws.

Of course, crafting laws isn’t really like ordering off a menu. It’s more like cooking: the ingredients intermingle and affect one another. A law holding platforms liable for defamatory speech they “know” about, for example, may mean something different depending whether the law lets accused speakers explain and defend their posts. But isolating the options in modular form can, I hope, help in identifying options for pragmatic and well-tailored laws.

IL laws generally try to balance three goals. The first is preventing harm. It’s no accident that intermediary immunities are typically weakest for content that poses the greatest threats, including material criminalized by U.S. federal law. The second is protecting speech and public participation. For this goal, one concern is to avoid over-removal – the [well-documented](#) phenomenon of platforms cautiously deferring to bogus legal accusations and taking down users’ lawful speech. Another is to encourage new market entrants to build, and investors to fund, open speech platforms in the first place. The third, related goal is encouraging technical innovation and economic growth. A rule that creates great legal uncertainty, or that can only be enforced by hiring armies of moderators, raises formidable barriers to entry for potential competitors with today’s mega-platforms. Lawmakers use the doctrinal dials and knobs listed in the remainder of this post to adjust policy trade-offs between these goals.

Major Free Expression Considerations

Who decides what speech is illegal?

Outside the United States, blanket immunities like [CDA 230](#) are rare. But it’s not uncommon for courts or legislatures to keep platforms out of the business of deciding what speech violates the law. One IL model widely endorsed by free expression advocates holds platforms immune unless a court or other government authority rules content illegal. In practice, this highly speech-protective standard typically has exceptions, requiring platforms to act of their own volition against highly recognizable and dangerous content such as child sex abuse images. Lawmakers who want to move the dial more toward harm prevention without having platforms adjudicate questions of speech law can also create

accelerated administrative or TRO processes, or give platforms other responsibilities such as educating users, developing streamlined tools, or providing information to authorities.

Must platforms proactively monitor, filter, or police users’ speech?

Human rights literature includes strong warnings against making platforms monitor their users. Many IL laws expressly bar such requirements, though they have gained traction in recent European legislation. One concern is that technical filters are likely to [over-remove](#), given their inability to recognize [contexts](#) like news reporting or parody. (However, filtering is relatively accepted for child sexual abuse images, which are unlawful in every context.) Another is that, when platforms have to review and face over-removal incentives for every word users post, the volume and invasiveness of unnecessary takedowns can be expected to rise. Legal exposure and enforcement costs under this model may also give platforms reason to allow only approved, pre-screened speakers – and deter new market entrants from challenging incumbents.

Must platforms provide “private due process” in takedown operations?

Improving platforms’ internal notice-and-takedown processes can protect against over-removal. A widely supported civil society document, the [Manila Principles](#), provides a list of procedural rules for this purpose. For example, a platform can be required or incentivized to notify speakers and let them defend their speech – which may help deter bad-faith notices in the first place. Accusers can also be required to include adequate information in notices, and face penalties for bad-faith takedown demands. And platforms can be required to disclose raw or aggregate data about takedowns, in order to facilitate public review and correction.

Can platforms’ use of private Terms of Service prohibit lawful expression?

Platforms often prohibit disfavored but legal speech under their Terms of Service (TOS). To maximize users’ free expression rights, a law might limit or ban this restriction on speech. In the United States, though, such a law might violate platforms’ own speech and property rights. Platforms’ value for ordinary users would also decline if users were constantly faced with bullying, racial epithets, pornography, and other legal but offensive matter. (I address relevant law in depth [here](#) and explore possible regulatory models in that paper’s final section.)

Can speakers defend their rights in court?

Platform over-removal incentives come in part from asymmetry between the legal rights of accusers and those of speakers. Victims of speech-based harms can often sue platforms to get content taken down. Speakers can almost never sue to get content reinstated. A few untested new laws in Europe try to remedy this, but it is unclear how well they will work or how speakers’ claims will intersect with platforms’ power to take down speech using their TOS.

Are leaving content up and taking it down the only options?

IL laws occasionally use more tailored remedies, in place of binary take-down/leave-up requirements – like making search engines suppress results for some search queries, but not others. Platforms could also do things like showing users a warning before displaying certain content, or cutting off ad revenue or eligibility for inclusion in recommendations. In principle, IL law could also regulate the algorithms

platforms use to rank, recommend, or otherwise amplify or suppress user content – thought that would raise particularly thorny First Amendment questions and be extremely complex to administer.

Treating Platforms Like Publishers

Making platforms liable for content they control

Most IL laws strip immunity from platforms that are too actively involved in user content. Some version of this rule is necessary to distinguish platforms from content creators. More broadly, putting liability on an entity that exercises editor-like power comports with traditional tort rules and most people’s sense of fairness. But standards like these may play out very differently for Internet platforms than for traditional publishers and distributors, given the comparatively vast amount of speech platforms handle and their weak incentives to defend it. Laws that reward passivity may also deter platforms from trying to weed out illegal content and generate legal uncertainty about features beyond bare-bones hosting and transmission.

Making platforms liable for content they know about

Many legal systems hold platforms liable for continuing to host or transmit illegal content once they “know” or “should know” about it. Laws that rely on these *scienter* standards can protect legal speech somewhat by defining “knowledge” narrowly or adding elements like private due process. Other legal regimes reject *scienter* standards, considering them too likely to incentivize over-removal.

Using “Good Samaritan” rules to encourage content moderation

Platforms may be deterred from moderating content by fear that their efforts will be used against them. Plaintiffs can (and do) argue that by moderating, platforms assume editorial control or gain culpable knowledge. Concern about the resulting perverse incentives led Congress to create CDA 230, which makes knowledge and control largely irrelevant for platform liability. This encouraged today’s moderation efforts but also introduced opportunities for bias or unfairness.

Different Rules for Different Problems

Legal claims

IL laws often tailor platforms’ duties based on the claim at issue. For example, they may require urgent responses for particularly harmful content, like child sex abuse images; deem court review essential for claims that turn on disputed facts and nuanced law, like defamation; or establish private notice-and-takedown processes in high-volume areas, like copyright.

Platform technical function

Many IL laws put the risk of liability on the entities most capable of carrying out targeted removals. Thus, infrastructure providers like ISPs or domain registries generally have stronger legal immunities than consumer-facing platforms like YouTube, which can do things like take down a single comment or video instead of a whole page or website.

Platform size

Recently, experts have raised the possibility of special obligations for mega-platforms like Google or Facebook. Drafting such provisions without distorting market incentives or punishing non-commercial platforms like Wikipedia would be challenging. In principle, though, it might improve protections on the most popular forums for online expression, without imposing such onerous requirements that smaller market entrants couldn't compete.

General Regulatory Approach

Bright-line rules versus fuzzy standards

IL rules can hold platforms to flexible standards like “reasonableness,” or they can prescribe specific steps. Platforms – especially the ones that can't hire a lawyer for every incoming claim – typically favor the latter, because it provides relative certainty and guidance. Free expression advocates also often prefer clear processes, because they reduce the role of platform judgment and allow legislatures to add procedural protections like counter-notice.

Liability for single failures versus liability for systemic failures

Some recent European [laws](#) and proposals accept that takedown errors are inevitable and do not impose serious financial penalties for individual items of content. Instead they penalize platforms if their overall takedown system is deemed inadequate. This approach generally reduces over-removal incentives, but is more viable in legal systems with trusted regulators.

Liability for platforms versus liability for speakers

Internet users may see little reason to avoid disseminating unlawful content when the legal consequences of their actions fall primarily on platforms. Laws could be structured to shift more risk to those individuals. For example, claims against platforms could be limited if claimants do not first seek relief from the responsible user. Or platforms' immunities could be made contingent on preserving or disclosing information about online speakers – though this would raise serious concerns about privacy and anonymity rights.

See published version here: <https://balkin.blogspot.com/2019/06/build-your-own-intermediary-liability.html>. Daphne Keller is Director of Intermediary Liability at the Stanford Center for Internet and Society, and was previously Associate General Counsel at Google.

Privatizing Censorship

Michael Karanicolas

Privatization can be a controversial practice. To its proponents, it is an engine of efficiency, introducing a competitive atmosphere to stodgy and self-perpetuating bureaucracies. But there are also externalities which can come into play when governments abrogate direct responsibility over an area of administration. A private prison may be run at less cost to the taxpayer, but will it respect the rights of inmates and devote sufficient resources to their rehabilitation? Privatizing a water company could turn it profitable, but this might come at the cost of an increase in contaminants or a refusal to service unprofitable areas. Despite the common refrain that [government should be run like a business](#), there is an important distinction between the core functions of these two types of entities. A private company's purpose, its only purpose, is to maximize profit for its shareholders. A government's purpose is to promote and protect the rights of its people.

Regulating speech is among the most important, and most delicate, tasks that a government may undertake. It requires a [careful balancing](#) between removing harmful content while providing space for controversial and challenging ideas to spread, and between deterring dangerous speech while minimizing a broader chilling effect that can impact legitimate areas of debate. The challenges in regulating speech are among the most vibrant and hotly debated areas of law and philosophy, with a voluminous history of jurisprudence and academic theory on how regulations should be crafted.

Today, this entire school of thought is being cast by the wayside, as the practical functions of content regulation are being increasingly handed over to an industry which is not only totally unprepared to handle the subtleties and technical challenges associated with defining the contours of acceptable speech on a global scale, but has, as far as possible, resisted taking responsibility for this function.

How did we get here?

In the early days of the commercial Internet, policymakers realized that the commercial and social potential of this new medium could best be realized if service providers were protected against direct liability for the words of their users. Without it, scalability of the kind achieved by Facebook and Twitter would never have been possible. However, this has turned into a double-edged sword. Having been allowed to grow without an expectation of policing their users, the world's biggest tech firms were built around business models that make it very difficult to control how their products are being used.

Now, governments are demanding that the companies start taking responsibility, and impose content controls that suit their needs. In some cases, these involve fairly well recognized categories of harmful content, such as hate speech or child abuse imagery. Other examples revolve around content which is outlawed locally, but whose prohibition runs counter to global freedom of expression standards, from [risqué photos of the King of Thailand](#) to material [deemed to violate conservative religious standards](#). In some instances, companies have entered into collaborative relationships with governments to remove content that is determined to be objectionable, notably (and controversially) in [Israel](#). Demands for private sector cooperation are backed by a variety of coercive measures, including the imposition of large fines, threats to block a company's website, and even the arrest and imprisonment of company employees.

The end result is a “privatized” system of content control, which is run at the behest of government authorities, but which is operated and enforced by the tech companies. To understand why this is problematic, consider the case of South Korea, where content enforcement decisions are made by the [Korea Communications Standards Commission \(KCSC\)](#), an administrative body whose members are appointed by the President. The KCSC is [notoriously heavy handed](#), and frequently targets sites which criticize politicians or challenge sensitive policy areas. Their decisions are issued to the platforms, rather than to the users who post the material, and come in the form of non-binding requests for removal. Weak intermediary liability protections mean that, in practice, these requests are always followed. However, the fact that the decisions are not formally binding means that, technically, enforcement originates from the platform, rather than the KCSC, which strips users of any procedural safeguards, such as a right of appeal or even notification that their material is subject to removal.

This practice of “laundering” government content restrictions through the private sector allows for mechanisms of control which vastly outstrip what might otherwise be permissible in a democratic context. For example, Germany’s [Network Enforcement Act \(NetzDG\)](#), which came into force in 2018, requires companies to remove “obviously illegal” material within 24 hours of being notified of its existence. More recently, [proposals from European Parliament](#) could push the deadline for responding to “terrorist content” notifications to just one hour. No judicial or administrative process in the world operates this quickly. Similarly, traditional content restrictions were designed on the understanding that their applicability would be limited by the resources available for enforcement. But in the context of private sector platforms, enforcement is expected to be close to 100 percent, creating a vastly more intrusive system.

These issues are compounded by the fact that, due to the size and scale of the major platforms, the only practical avenue to developing moderation solutions that approach what governments are demanding is to lean heavily on automated decision-making systems. But while AI is relatively competent at screening for nudity, content that implicates hate speech or copyright infringement is vastly more difficult since it is inherently contextual. An identical statement made in Myanmar and in Canada [could qualify as hate speech in the former but not in the latter](#), due to the fact that one country has a much higher level of underlying ethnic tension. Not only is AI presently incapable of making this type of determination, but it is questionable whether the technology will ever be able to do so.

Moreover, in a context where the legal framework sets a minimum standard of enforcement, with harsh penalties for dropping below that standard, platforms are incentivized to err on the side of caution and remove anything which even approaches the line. This problem has been [widely documented](#) with regard to the DMCA system of copyright enforcement, including clear instances where it has been [gamed to target political opponents](#). Increasing automation will only exacerbate this tendency.

None of this is to suggest that tech companies should have no responsibilities with regard to the impact of their products on the world. But perspective is important. The resiliency of the Internet to pervasive forms of content control is a feature of the technology, not a bug. Just as we celebrate the inability of Vladimir Putin to remove an [embarrassing image of himself](#) or Xi Jinping’s struggles to stop Internet users from [comparing him to Winnie the Pooh](#), it is these same characteristics that make it so difficult to clamp down on the [viral spreading of video of the Christchurch attack](#).

The new privatized enforcement models, which are being embraced, to some degree, by virtually every developed democracy, threaten many key safeguards that were developed to prevent the abusive application of content restrictions. While there are clearly problems in moderating online speech that need to be addressed, the solution to these challenges must be crafted within well-recognized global norms of freedom of expression, including appropriate checks and balances, and not as a private sector solution to what is fundamentally a matter of public interest.

See published version here: <https://balkin.blogspot.com/2019/06/privatizing-censorship.html>. Michael Karanickolas is a human rights advocate who is based in Halifax, Canada. He is a graduate student in law at the University of Toronto and, as of July 2019, will be the incoming WIII Fellow at the Information Society Project at Yale Law School.

Expand Intermediary Liability to Protect Reality Itself

Tiffany Li

Intermediary liability is not, perhaps, the most exciting phrase in law. It's certainly not as buzzworthy as "impeachment" or "homicide." However, as a legal concept, intermediary liability is interesting and worthy of attention, not only because it is vital to understanding the role of tech platforms in society, but also because the intermediary liability issues of today may one day form the foundations for a new understanding of reality itself. As such, this essay suggests an expansion of the field of "intermediary liability" to encompass the responsibilities and potential risks that will arise as new forms of technological intermediaries change our understanding of online and offline reality.

It may sound like a bit of a stretch to say that intermediary liability law will shake the foundations of reality. However, consider the nature of what an intermediary is, and what intermediaries do. Currently, the internet intermediaries of the present act as venues and hosts, intermediating between people and information. Intermediaries are search engines, social media apps, web hosting providers, and the like. These online intermediaries seem clearly separable from our offline, "real" lives in the physical world.

Yet, as our world grows increasingly digitized, it is all but inevitable that human beings will live more of their lives online than offline. In an increasingly online world, in which we interact with each other through an ever increasing number of new intermediaries, the concept of intermediary liability must be recalibrated to adapt to new technologies. New intermediaries will include the engines for virtual and augmented reality (VR/AR) environments, as well as smart cities and Internet of Things (IoT) environments. Intermediary liability doctrine should expand to include these new technological intermediaries.

There is already potential for technology intermediaries to gain immense power over users, and new technological advances will likely exacerbate this power dynamic. Intermediary liability law is already insufficient to address harms that cross the boundaries of online and offline space. Problems like revenge porn, swatting, extremism, and election manipulation stretch the bounds of what we traditionally have understood to be responsibilities of internet intermediaries. By expanding and updating the concept of intermediary liability to include new technological intermediaries, we may be able to hold powerful actors in check before new technologies become so pervasive that the distinction between online and offline fails entirely.

Today, the phrase "intermediary liability" generally brings to mind issues related to information (data and content) on the internet, as well as the responsibilities tech platforms have over that information. However, Facebook and Google are not the proto-intermediaries. Before the internet, [telecommunications intermediaries](#) faced similar questions. Before then, print publishers (along with re-publishers, sellers, re-sellers, and so on) also wrestled with many of the questions we view as paramount in intermediary liability law today. With each new wave of information technology comes a new form of information intermediary, and along with it, a new line of intermediary liability laws.

Currently, intermediary liability laws consider tech platforms to be intermediaries, entities that act as go-betweens for individuals, providing venues for communication and information access and exchange. However, the "intermediary" nature of information platforms may soon be changing, as

the internet becomes more of an all-encompassing space than a liminal staging ground. In the early days of the internet, people “went online.” Now, for many, “going offline” is becoming the more unusual state. Consider the amount of time the average consumer spends interacting with the internet via various devices, whether they take the form of mobile phones, desktop computers, or touchscreen refrigerators in grocery markets that can [scan consumer faces](#) and offer targeted advertisements using facial recognition technology.

We are not in danger of living in the Matrix yet. However, VR/AR technology is improving. Artificial intelligence and advanced machine learning systems are advancing. The IoT is growing at an incredible pace. The burgeoning [5G industry](#) will only increase this shift, as the [low latency networks](#) will allow for greater proliferation of IoT systems. Smart cities may soon become commonplace. In this new connected world, we will need new laws to protect against technological harms. Current intermediary liability doctrines must change to protect against these new harms.

Expanding the field of intermediary liability law will require exploring new research questions. Here are just a few that come to mind:

What is the “intermediary” nature of a technology that allows for [brain-to-brain](#) direct communication? How should intermediary liability laws change to reflect that?

How does intermediary liability work when the “layers” of intermediaries become an interconnected web?

Which networks and which services count as intermediaries when IoT devices proliferate to an extent that we have truly connected smart cities?

If and when VR/AR technology improves to a point that we can live substantial portions of our lives either in a virtual environment or in an environment augmented by digital technology, which entity will be the VR/AR intermediary?

If Facebook or Google creates the backbone for future VR worlds, they could effectively control the reality of the future. At that point, will we still consider these intermediary companies to be properly regulated by simple intermediary liability laws?

When thinking about these and other somewhat outlandish science-fiction-like future scenarios, I often reflect on one of my favorite quotes from [Jack Balkin](#):

“If we assume that a technological development is important to law only if it creates something utterly new, and we can find analogues in the past—as we always can—we are likely to conclude that because the development is not new, it changes nothing important. That is the wrong way to think about technological change and public policy, and in particular, it is the wrong way to think about the Internet and digital technologies.

“Instead of focusing on novelty, we should focus on salience. What elements of the social world does a new technology make particularly salient that went relatively unnoticed before? What features of human activity or of the human condition does a technological change foreground, emphasize, or problematize? And what are the consequences for human freedom of making this aspect more important, more pervasive, or more central than it was before?”

The new intermediary technologies of IoT, smart cities, cloud computing, artificial intelligence, machine learning, and VR/AR are novel, yes, but what's important to understand is what these new technologies make salient about human society: namely, that the next tech platforms will be more than intermediaries between people and information. Future tech platforms will be intermediaries between people and the world itself.

These technologies underline a growing trend towards a more digitized, online life, where the lines between what is "real" and what is "virtual" may be softly blurring. This increased connectivity is leading to, if not a virtual reality world, then at least a gradual virtualization of reality. The intermediaries that power the internet, the IoT, and connected and virtual systems will only grow in power and influence, and the law must keep pace to protect individuals from new technological harms. The field of intermediary liability law can and should expand to include the new questions posed by these future virtual intermediaries.

See published version here: <https://balkin.blogspot.com/2019/06/expand-intermediary-liability-to.html>. Tiffany Li is a Resident Fellow at Yale Law School's Information Society Project, where she leads the Wikimedia/Yale Law School Initiative on Intermediaries and Information.