MLRC Media Law Resource Center
BULLETIN

2020 Issue No. 1 July 2020

LEGAL FRONTIERS IN DIGITAL MEDIA

Why Tying Section 230 to Political Neutrality Violates the First Amendment

By Berin Szóka[1]

A month ago, President Trump demanded action against social media sites for "censoring" conservatives. His Executive Order[2] called for federal legislation. Last week, the Department of Justice made a more specific proposal[3] almost simultaneously with Sen. Josh Hawley's introduction[4] of the "Limiting Section 230 Immunity to Good Samaritans Act."[5] This bill would put a gun to the head of the largest social media websites, forcing them to give up editorial control over their services if they want to stay in business.

The First Amendment would not allow Congress to *directly* require websites to be politically "neutral" or "fair": the Supreme Court has recognized that the First Amendment protects the editorial discretion of websites no less than newspapers. Both have the same right to decide what content they want to carry; whether that content is created by third parties is immaterial. Hawley attempts to lawyer over the constitutional problem, using an intentionally convoluted process to conceal the bill's coercive nature and to present himself as a champion of "free speech," while actually proposing to empower the government to censor online content as never before.

Instead of directly meddling with how websites moderate content, Hawley's bill relies on two legal sleights of hand. The first involves Section 230 of the Communications Decency Act of 1996. That law made today's Internet possible — not only social media but all websites and services that host user content — by protecting them from most civil liability (and state criminal prosecution) for content created by third parties. Given the scale of user-generated content — with every comment, post, photo and video potentially resulting in a lawsuit — websites simply could not function if Section 230 did not immunize them not just from ultimate liability but from the litigation grindstone itself. Hawley knows that all sites that host user content depend on Section 230, so he's

---

[1] Berin Szóka is a Senior Fellow at TechFreedom, which he founded in 2010. Previously, he was a Senior Fellow and the Director of the Center for Internet Freedom at The Progress & Freedom Foundation. Before joining PFF, he was an Associate in the Communications Practice Group at Latham & Watkins LLP, where he advised clients on regulations affecting the Internet and telecommunications industries. Before joining Latham's Communications Practice Group, Szoka practiced at Lawler Metzger Milkman & Keeney, LLC, a boutique telecommunications law firm in Washington, and clerked for the Hon. H. Dale Cook, Senior U.S. District Judge for the Northern District of Oklahoma. Szoka received his Bachelor's degree in economics from Duke University and his juris doctor from the University of Virginia School of Law, where he served as Submissions Editor of the Virginia Journal of Law and Technology. He is admitted to practice law in the District of Columbia and California (inactive).

[2] Exec. Order No. 13925, 85 FR 34079 (May 28, 2020), https://www.whitehouse.gov/presidential-actions/executive-order-preventing-online-censorship/.

[3] Press Release, Dep't of Justice, Justice Department Issues Recommendations for Section 230 Reform (June 17, 2020), https://www.justice.gov/opa/pr/justice-department-issues-recommendations-section-230-reform.

[4] *Senator Hawley Announces Bill Empowering Americans to Sue Big Tech Companies Acting in Bad Faith*, Josh Hawley (June 17, 2020), https://www.hawley.senate.gov/senator-hawley-announces-bill-empowering-americans-sue-big-tech-companies-acting-bad-faith.

[5] Limiting Section 230 Immunity to Good Samaritans Act, S.3983, 116th Cong. (2020).

carefully crafted a bill that turns that dependence against them — to do something the First Amendment clearly forbids: to force them to cede editorial control over their services.

Second, Hawley claims that his bill "protects consumers" by holding companies to their promises. In reality, it defines "good faith" so broadly that "edge providers" would face a constant threat of being sued under consumer protection and contract laws for how they exercise their editorial discretion over user content. Given the fines involved ($5,000/user plus attorneys' fees), a single court decision could bankrupt even the largest tech company.

No one should have any illusion about what Hawley's bill really does: use state power to advance a political agenda. The bill's complicated structure merely masks the elaborate ways it violates the First Amendment. Conditioning 230 immunity on opening yourself up to legal liability under consumer protection law is a Rube-Goldberg-esque legal contraption intended to do what the First Amendment clearly forbids: forcing websites to host user-generated content they find objectionable.

I. How the Hawley Bill Works

Section 230(c)(1) says: "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider." These have been called the The Twenty-Six Words That Created the Internet.[6] When websites and services are sued for third party content they host, Section 230 allows them to cheaply get lawsuits against them thrown out with a motion to dismiss. Consequently, lawsuits are far rarer than they would be in a world without 230. Section 230(c)(1) ensures that those who create content are the ones to be sued. Courts resolve nearly all 230 cases under this provision.

Republicans have insisted angrily that *all* of Section 230 was intended to depend on a showing of good faith, including political neutrality; however, the plain text of the statute is clear. Only Subsection 230(c)(2)(A) requires such a showing — and the statute's operative language *doesn't* mention neutrality. As Justice Neil Gorsuch recently declared, "When the express terms of a statute give us one answer and extratextual considerations suggest another, it's no contest. Only the written word is the law, and all persons are entitled to its benefit."[7] By proposing to amend Section 230(c)(1) to require both good faith *and* neutrality, Trump's DOJ and Hawley both concede that the President's Executive Order and other Republican clamoring for immediate legal action are simply wrong about the current state of the law.

The real aim of Hawley's bill is to force the largest social media services to change how they treat content that serves the "MAGA" political agenda — *e.g.*, not labeling Trump's tweets, allowing far-right provocateurs to engage in bannable conduct, treating Diamond and Silk or Gateway Pundit as the journalistic equivalents of *The New York Times*. The bill is almost perfectly tailored to do just that while avoiding damage to smaller, alternative social networks favored by conservative activists for their "anything goes" approach to content moderation.

[6] JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET (1st ed. 2019).

[7] *Bostock v. Clayton County*, 590 U.S. ___ (2020).

Hawley's bill applies only to "edge providers": websites or services with 30+ million annual unique users, or more than 300 million unique global users, in the past year, and more than $1.5 billion in global revenue. To maintain 230(c)(1) protections, they would have to attest to "good faith" — essentially, political neutrality — in their content moderation practices. Thus, an edge provider has to choose between two litigation risks: If it "voluntarily" exposes itself to suit for the "fairness" of its content moderation, it cedes editorial control to judges and regulators. If it surrenders Section 230 protections, it risks being sued for anything its users say — which may simply make it impossible for them to operate.

Trump's Executive Order asks the Federal Communications Commission to collapse Section 230's three distinct immunities into a single immunity dependent on "good faith" — and then define that term broadly to include neutrality and potentially much more. The Hawley bill does roughly the same thing by requiring large "edge providers" to promise "good faith." Both would change the dynamics of litigation completely: A plaintiff with a facially plausible complaint would (1) prevail on a motion to dismiss, (2) get court-ordered discovery of internal documents and depositions of employees to assess "good faith" (however that term is expanded), and (3) force the company to litigate all the way through a motion for summary judgment. Whether or not the plaintiff ultimately wins, this pre-trial phase of litigation is where the defendant will incur the vast majority of their legal costs — and where plaintiffs force settlements. Multiply those costs of litigation, and settlement, times the millions or billions of pieces of content posted to social media sites every day and you get "death by ten thousand duck-bites."[8] That's why Judge Alex Kozinski (a longtime conservative champion once short-listed for the Supreme Court) declared: "section 230 must be interpreted to protect websites not merely from ultimate liability, but from having to fight costly and protracted legal battles."[9]

Having to prove good faith to resolve litigation would kill most social media websites, which exist to host content by others. Ironically, it's possible that the best established social media sites with the biggest legal departments might cope; they might even be grateful that Hawley's bill had made it impossible for new competitors to get off the ground. At the same time, if (c)(1) is no longer an immunity from suit but merely a defense raised only after great expense, websites across the Internet would simply turn off their comments sections.

Today, Section 230 doesn't define "good faith." Courts assessing eligibility for the 230(c)(2)(A) immunity have defined the term narrowly.[10] Hawley's bill would add a five-factor definition of "good faith" in a new Subsection 230(c)(3). These factors would give plaintiffs ample room to declare that an edge provider had been politically biased against them. Inevitably, courts would have to analyze the nature of third-party content, comparing content that had been removed with content that had not in order to judge overall patterns.

[8] *Fair v. Roommates*, 521 F.3d 1157, 1174 (9th Cir. 2008).

[9] *Id*.

[10] *See e.g., BFS Fin. v. My Triggers Co.*, No. 09CV-14836 (Franklin Cnty. Ct. Com. Pl. Aug. 31, 2011) (allowing antitrust claims); *Smith v. Trusted Universal Standards in Elec. Transactions*, 2011 WL 900096, at *25–26 (D.N.J. Mar. 15, 2011).

To maintain 230 protections, an edge provider must also agree to pay up to $5,000 damages to users if it is found to have breached its (compelled) promises of "neutrality." Three hundred million users times $5,000 is $1.5 *trillion* dollars, exceeding the entire market cap of Google. The bill also adds attorneys fees, threatening to create a cottage industry of litigation against edge providers. The mere threat of such massive fines will fundamentally change how websites operate — precisely Hawley's goal.

Perhaps most important is what the bill *doesn't* say: unlike Trump's Order, Hawley's bill doesn't directly call on the FTC or state AGs to sue websites for bias. But make no mistake; his bill would weaponize federal and state consumer protection laws to allow politicians to coerce social media into favoring their side of the culture wars. The FTC might hesitate to bring such suits, because of all the constitutional problems discussed below, but multiple Republican attorneys general have already made political hay out of grandstanding against[11] "liberal San Francisco tech giants." They would surely use Hawley's bill to harass edge providers, raise money for their campaigns, and run for governor — or Senate.

II. A New Fairness Doctrine — with Even Greater First Amendment Problems

The Original Fairness Doctrine required broadcasters (1) to "adequately cover issues of public importance" and (2) to ensure that "the various positions taken by responsible groups" were aired, thus mandating the availability of airtime to those seeking to voice an alternative opinion. President Reagan's FCC abolished these requirements in 1987. When Reagan vetoed Democratic legislation to restore them, he noted that "the FCC found that the doctrine in fact inhibits broadcasters from presenting controversial issues of public importance, and thus defeats its own purpose."

The Republican Party has steadfastly opposed the Fairness Doctrine for decades. The 2016 Republican platform (re-adopted verbatim for 2020) states: "We likewise call for an end to the so-called Fairness Doctrine, and support free-market approaches to free speech unregulated by government." Yet now, Hawley and Trump propose a version of the Fairness Doctrine for the Internet that would be more vague, intrusive, and arbitrary than the original.

In *Miami Herald Publishing Co. v. Tornillo*,[12] the Supreme Court struck down a 1913 state law imposing a version of the Fairness Doctrine on newspapers that required them to grant a "right of reply" to candidates for public office criticized in their pages. The Court acknowledged that there had been a technological "revolution" since the enactment of the First Amendment. The arguments made then about newspapers, as summarized by the Court, are essentially the same arguments conservatives make about digital media:

> The result of these vast changes has been to place in a few hands the power to inform the American people and shape public opinion…. The abuses of bias and manipulative reportage are, likewise, said to be the result of the vast accumulations of unreviewable power in the modern media empires. The First Amendment interest

[11] E-Mail from Ken Paxton, Tex. Att'y Gen., available at https://mailchi.mp/kenpaxton/texas-is-leading-were-taking-on-silicon-valley?e=9718826c0b.

[12] 418 U.S. 241 (1974).

> of the public in being informed is said to be in peril because the 'marketplace of ideas' is today a monopoly controlled by the owners of the market.[13]

And yet, the court struck down the law as unconstitutional because:

> a compulsion to publish that which "'reason' tells them should not be published" is unconstitutional. A responsible press is an undoubtedly desirable goal, but press responsibility is not mandated by the Constitution and like many other virtues it cannot be legislated.[14]

"Government-enforced right of access inescapably 'dampens the vigor and limits the variety of public debate.'" *Id*. at 257. Critically, the Court rejected the intrusion into the editorial discretion "[e]ven if a newspaper would face no additional costs to comply," because:

> A newspaper is more than a passive receptacle or conduit for news, comment, and advertising. The choice of material to go into a newspaper, and the decisions made as to limitations on the size and content of the paper, and treatment of public issues and public officials — whether fair or unfair — constitute the exercise of editorial control and judgment.[15]

The Trump/Hawley Fairness Doctrine would impose the very same intrusion upon editorial judgments of edge providers. In addition, determining whether a website has operated "fairly" would be "void for vagueness since no editor could know exactly what words would call the statute into operation."[16]

The Supreme Court upheld the Fairness Doctrine for broadcasters in *Red Lion Broadcasting Co. v. FCC*,[17] but only because the Court denied broadcasters full First Amendment protection: "Although broadcasting is clearly a medium affected by a First Amendment interest, differences in the characteristics of new media justify differences in the First Amendment standards." The same arguments have been made about the Internet, and the Supreme Court explicitly rejected them.

When the Court struck down Congress' first attempt to regulate the Internet, the Communications Decency Act (everything *except* Section 230), it held: "our cases provide no basis for qualifying the level of First Amendment scrutiny that should be applied to this medium."[18] The Court has since repeatedly reaffirmed this holding. While striking down a state law restricting the purchase of violent video games, Justice Scalia declared: "the basic principles of freedom of speech and the press, like the First Amendment's command, do not vary when a new and different medium for

---

[13] *Id*. at 250.

[14] *Id* at 256.

[15] *Id*. at 258.

[16] *Id*. at 247.

[17] 395 U.S. 367 (1969).

[18] *Reno v. American Civil Liberties Union*, 521 U.S. 844, 870 (1997).

communication appears."[19] In short, *Red Lion* represented an exception, and even that exception may not survive much longer.

III. Social Media Aren't Public Fora, So the First Amendment Protects Them

The President's Executive Order attempts to sidestep the Supreme Court's consistent protection of digital speech by claiming that social media are effectively "public fora" and thus that the First Amendment limits, rather than protects, their editorial discretion — as if they were extensions of the government: "It is the policy of the United States that large online platforms, such as Twitter and Facebook, as the critical means of promoting the free flow of speech and ideas today, should not restrict protected speech." The Order also cites the Supreme Court's decision that shopping malls *were* public fora under California's constitution in *Pruneyard Shopping Center v. Robins,* 447 U.S. 74, 85-89 (1980).

But Justice Kavanaugh, leading the five conservatives, explicitly rejected such arguments last year: "merely hosting speech by others is not a traditional, exclusive public function and does not alone transform private entities into state actors subject to First Amendment constraints."[20] *Pruneyard* simply doesn't apply to social media.

Trump's Order cites the Supreme Court's recent decision in *Packingham v. North Carolina*,[21] but omits the critical legal detail: it involved a *state law* restricting the Internet use of convicted sex offenders. Thus, *Packingham* changed nothing: the First Amendment still fully protects, rather than limits, the editorial discretion of website operators under *Miami Herald* and *Reno*.

IV. Hawley's Bill Imposes an Unconstitutional Condition

Hawley's bill turns on one underlying legal claim more than any other: that Section 230 is a special privilege granted only to large websites, and withholding it does not violate the First Amendment. The factual claim is false: the law applies equally to *all* websites, protecting newspapers, NationalReview.com, FoxNews.com and every local broadcaster from liability for user comments posted on their website in exactly the same way it protects social media websites for user content. The legal claim is also wrong.

The Supreme Court has clearly barred the government from forcing the surrender of First Amendment rights in order to qualify for a benefit or legal status. In *Agency for Int'l Dev. v. All. for Open Soc'y Int'l, Inc.*,[22] the Court said that the government couldn't condition the receipt of AIDS-related funding on the recipients' adoption of a policy opposing prostitution (a form of compelled speech). Much earlier, in *Speiser v. Randall*,[23] the Court made it clear that denying a

[19] *Brown v. Entertainment Merchants Assn.*, 564 U.S. 786, 790 (2011).

[20] *Manhattan Community Access Corp. v. Halleck*, 139 S. Ct. 1921, 1930 (2019).

[21] 137 S. Ct. 1730, 1737 (2017) (social media "can provide perhaps the most powerful mechanisms available to a private citizen to make his or her voice heard").

[22] 570 U.S. 205 (2013).

[23] 357 U.S. 513, 518 (1958).

tax exemption to claimants who engage in certain forms of speech effectively penalizes them for that speech — essentially fining them for exercising their First Amendment rights.

Using Section 230 to coerce social media companies into surrendering their First Amendment rights is no different. Consider how clearly the same kind of coercion would violate the First Amendment in other contexts. Pending legislation would immunize businesses that re-open during the pandemic from liability for those who might be infected by COVID-19 on their premises. Suppose the bill included a provision requiring such businesses to be politically neutral in any signage displayed on their stores — such that, if a business put up, or allowed a Black Lives Matter sign, they would have to allow a "right of reply" in the form of a sign from "the other side." The constitutional problem would be obvious and in no way ameliorated by the "voluntary" nature of the immunity program.

V. Social Media Companies Can't Be Forced to Risk Being Associated with Content They Find Objectionable

The case against unconstitutional conditions and public forum status is even clearer for websites than it would be for retailers or shopping malls, for two reasons. First, social *media* companies are in the speech business, unlike businesses whose storefronts might incidentally post their own speech or host the speech of others. *Reno* makes clear that websites enjoy the same First Amendment right as newspapers, and "[t]he choice of material to go into a newspaper, and the decisions made as to limitations on the size and content of the paper, and treatment of public issues and public officials — whether fair or unfair — constitute the exercise of editorial control and judgment."[24]

Second, *Pruneyard* emphasized that shopping malls could "expressly disavow any connection with the message by simply posting signs in the area where the speakers or handbillers stand." But users will naturally assume speech carried by a social network reflects their decision to carry it — just as Twitter and Facebook have been attacked for *not* removing President Trump's tweets or banning him from their services.

Disclaimers may actually be less effective online. Consider the three labels Twitter has applied to President Trump's tweets (the first two of which provoked the issuance of his Executive Order).

[24] *Miami Herald,* 418 U.S. at 258.

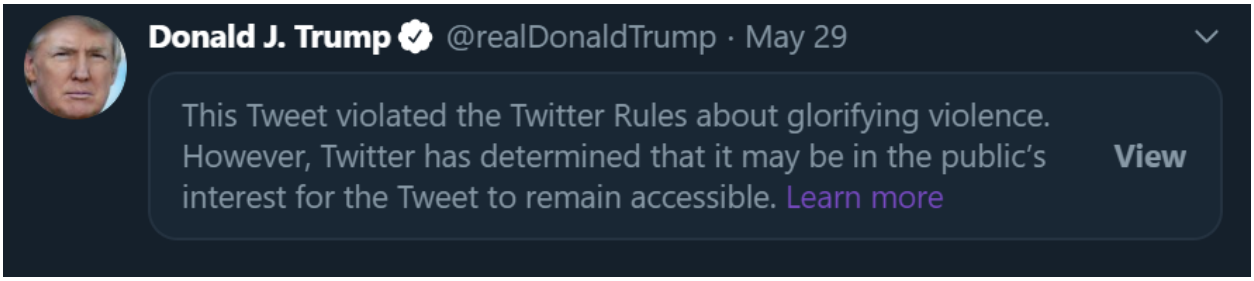
The first example[25] not only fails to clearly "disavow any connection with the message," it is also ambiguous: it could be interpreted to mean there really is some problem with mail-in ballots.

Similarly, Twitter applied a "(!) Manipulated Media" label to Trump's tweet of a video purporting to show CNN's anti-Trump bias. Twitter's label is once again ambiguous: since Trump's video

[25] @realDonaldTrump, TWITTER (May 26, 2020, 8:17 AM), https://twitter.com/realDonaldTrump/status/1265255835124539392.

claims that CNN had manipulated the original footage, the "manipulated media" claim could be interpreted to refer to either Trump's video or CNN's. Although the label links to an "event" page[26] explaining the controversy, the warning only works if users actually click through. It's far from clear to many users that the label is actually a link that will take them to a page with more information.

Finally, when Trump tweeted,[27] in reference to Black Lives Matter protests, "when the looting starts, the shooting starts," Twitter did not merely add a label below the tweet. Instead, it hid the tweet behind a disclaimer. Clicking on "view" allows the user to view the original tweet:



And yet Twitter has still been lambasted for not taking the tweet down completely, a decision interpreted by some as an acceptance of the validity of such an extreme position.

Further, disclaimers risk creating increased liability; indeed, they may trigger lawsuits from scorned politicians. For example, labeling (and hiding) Trump's tweets provoked issuance of the Executive Order. In the end, the only truly effective way for Twitter to disavow Trump's comments would be to ban him from their platform — precisely what the Hawley bill aims to deter.

In this sense, the Trump/Hawley version of the Fairness Doctrine is hugely *more* intrusive than the right of reply in the original Fairness Doctrine; it puts edge providers in the doubly unconstitutional position of (a) hosting content they do not want to host and (b) being afraid even to label it as content they find objectionable.

VI. Why the Hawley Bill's Good Faith Requirement Violates the First Amendment

To maintain 230 immunity, edge providers would be required to promise to moderate content in "good faith" — which the Hawley bill defines very loosely as "honest belief and purpose...fair dealing standards, and…[no] fraudulent intent" — in other words, political neutrality (and more). The bill adds this to Section 230's list of exceptions: "Nothing in this section shall be construed to impair or limit any claim for breach of contract, promissory estoppel, or breach of a duty of good faith.'' Thus, an edge provider's compelled "promises" could be enforced by the Federal Trade Commission, state AGs, or private plaintiffs under various federal and state consumer protection laws and common law contract theories. These enforcement mechanisms raise slightly different

[26] *Video being shared of CNN report on toddlers is doctored, journalists confirm*, Twitter (June 18, 2020), https://twitter.com/i/events/1273790055513903104.

[27] @realDonaldTrump, TWITTER (May 29, 2020, 12:53 AM), https://twitter.com/realDonaldTrump/status/1266231100780744704.

legal issues, but they all violate the First Amendment in essentially the same way: state action interfering with edge providers' exercise of editorial discretion.

VII. Consumer Protection Law Can't Police "Fairness" Claims

Republicans used to *oppose* weaponizing consumer protection laws against media companies. In 2004, MoveOn.org and Common Cause asked the FTC to proscribe Fox News' use of the slogan "Fair and Balanced" as a deceptive trade practice. Republican Chairman Tim Muris responded[28] pithily: "I am not aware of any instance in which the [FTC] has investigated the slogan of a news organization. There is no way to evaluate this petition without evaluating the content of the news at issue. That is a task the First Amendment leaves to the American people, not a government agency."

Similarly, the Hawley bill would necessarily embroil the FTC, state AGs, and judges in "evaluating the content … at issue." Media companies aren't exempt from consumer protection or antitrust laws, but the First Amendment makes suing them for how they exercise their editorial discretion extremely difficult, if not impossible — which is why the FTC has never attempted to police marketing claims about editorial practices the way it polices marketing claims generally.

As Chairman Muris noted, general statements about "fairness" or "neutrality" simply are not verifiable. This is why the Ninth Circuit recently dismissed Prager University's deceptive marketing claims against YouTube. Despite having over 2.52 million subscribers and more than a billion views, this right-wing producer[29] of "5-minute videos on things ranging from history and economics to science and happiness," sued[30] YouTube for "unlawfully censoring its educational videos and discriminating against its right to freedom of speech." Specifically, Dennis Prager alleged[31] that roughly a sixth of the site's videos had been flagged for YouTube's Restricted Mode,[32] an opt-in feature that allows parents, schools and libraries to restrict access to potentially sensitive (and is turned on by fewer than 1.5% of YouTube users). The Ninth Circuit ruled:

> YouTube's braggadocio about its commitment to free speech constitutes ***opinions*** that are not subject to the Lanham Act. Lofty but vague statements like "everyone deserves to have a voice, and that the world is a better place when we listen, share and build community through our stories" or that YouTube believes that "people should be able to speak freely, share opinions, foster open dialogue, and that creative freedom leads to new voices, formats and possibilities" are classic, non-

[28] Statement of Federal Trade Commission Chairman Timothy J. Muris on the Complaint Filed Today by MoveOn.org (July 19, 2004), https://www.ftc.gov/news-events/press-releases/2004/07/statement-federal-trade-commission-chairman-timothy-j-muris.

[29] PragerU, YOUTUBE, https://www.youtube.com/user/PragerUniversity/about (last visited July 26, 2020).

[30] *PragerU Takes Legal Action Against Google and YouTube for Discrimination*, PragerU (2020), https://www.prageru.com/press-release/prageru-takes-legal-action-against-google-and-youtube-for-discrimination/.

[31] Dennis Prager, *Don't Let Google Get Away With Censorship*, The Wall Street Journal (Aug. 6, 2019), https://www.wsj.com/articles/dont-let-google-get-away-with-censorship-11565132175.

[32] Your content & Restricted Mode. YouTube Help (2020), https://support.google.com/youtube/answer/7354993?hl=en.

> actionable opinions or puffery.[33] Similarly, YouTube's statements that the platform will "help [one] grow," "discover what works best," and "giv[e] [one] tools, insights and best practices" for using YouTube's products are ***impervious to being "quantifiable," and thus are non-actionable "puffery."***[34] The district court correctly dismissed the Lanham Act claim.[35]

Websites *can't* be sued today for making statements that may sound like offering neutrality — contrary to Republican claims that they should be, and Trump's call for such lawsuits in this Executive Order. The Hawley bill implicitly concedes this point.

But simply forcing edge providers to be more *specific* in their claims about neutrality will not overcome the ultimate constitutional problem. Puffery includes "claims [which] are either vague or *highly subjective*."[36] It would be difficult to imagine a more subjective marketing claim than one about "good faith," "neutrality" or "fairness." Ultimately, the reason consumer protection law does not attempt to police marketing claims about neutrality is not their lack of specificity but their *subjectivity.*

In theory, the FTC might be able to base a deception case on certain very clear, *objective* claims about editorial practices; that category of deception, however, would be narrow — the use of human moderators to evaluate particular pieces of content or to decide which topics are "trending," or the application of community standards to elected officials, for example. These deception cases would do little to address the complaints of conservatives, and even such narrow complaints might be unconstitutional.

VIII. Consumer Protection Law Can't Police Non-Commercial Speech

The FTC can police marketing claims for being misleading to the extent they "propose a commercial transaction."[37] Community standards documents do much more than that: they are essentially statements of values, comparable to Christian retailer Hobby Lobby's statement that the company is committed to "[h]onoring the Lord in all we do by operating the company in a manner consistent with Biblical principles."

Such statements are non-commercial speech, which is fully protected by the First Amendment under strict scrutiny *even when it is misleading*.[38] To overcome strict scrutiny, the government must show that the bill is (1) necessary to address a compelling government interest (2) to which the law is narrowly tailored, and (3) that the government uses the least restrictive means possible

---

[33] *See Newcal Indus., Inc. v. Ikon Office Sol.*, 513 F.3d 1038, 1053 (9th Cir. 2008).

[34] *Id.*

[35] *Prager Univ. v. Google LLC*,[35] 951 F.3d 991, 1000 (9th Cir. 2020).

[36] *Sterling Drug, Inc. v. FTC*, 741 F.2d 1146, 1150 (9th Cir. 1984) (emphasis added).

[37] *Central Hudson Gas & Elec. Corp. v. Public Service Comm'n of New York*, 447 U.S. 557,561 (1980); *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 762 (1976).

[38] *United States v. Alvarez*, 567 U.S. 709 (2012).

to address that interest.[39] In *Miami Herald*, the court noted that Florida's interest in "ensuring free and fair elections" was a "concededly important interest," but had to yield to the "unexceptionable, but nonetheless timeless, sentiment that liberty of the press is in peril as soon as the government tries to compel what is to go into a newspaper."[40] The bill also fails on the second two prongs of strict scrutiny.

If the Hawley bill passes, the Trump Administration will undoubtedly argue that edge providers' community standards are ads for their services. But when speech has commercial aspects that are "inextricably intertwined" with other fully protected speech, that speech is generally fully protected.[41] For example, corporate statements endorsing Black Lives Matter receive First Amendment protection even when embedded in marketing claims.

Courts are generally reluctant to label content as commercial speech because that denies the speech full First Amendment protection. Although community standards and terms of service may "refer[] to a specific product," they in no way resemble traditional advertising — two of the factors courts assess in drawing the line between commercial and noncommercial speech.[42] The third factor, the profit motive — which Hawley harps on in his public statements — is not dispositive: "If a newspaper's profit motive were determinative, all aspects of its operations—***from the selection of news stories to the choice of editorial position***—would be subject to regulation if it could be established that they were conducted with a view toward increased sales."[43]

Pittsburgh Press makes clear that statements about the way publishers exercise their editorial discretion are fundamentally different from statements about the health benefits of drug products, for example.

Even if a court decided to treat community standards as commercial speech, the government would still face an uphill battle. "The party seeking to uphold a restriction on commercial speech carries the burden of justifying it,"[44] and "must demonstrate that the harms it recites are real, and that its restriction will in fact alleviate them to a material degree."[45] Because the government's interest in regulating commercial speech lies in its misleading or false nature, it would have to show that statements about a website's editorial practices are misleading. General claims about "fairness," however, are simply not verifiable.

IX. Why the Government Can't Compel Disclosures about Editorial Policies

Compelling edge providers to change what they say about their community standards violates the First Amendment even apart from enforcement of such claims. As a condition for maintaining 230

[39] *Reed v. Town of Gilbert*, 576 U.S. 155, 163, 171 (2015).

[40] 418 U.S. at 260.

[41] *Riley v. Nat'l Fed'n of the Blind of N.C., Inc*., 487 U.S. 781, 783 (1988).

[42] *Bolger v. Youngs Drug Prods. Corp*., 463 U.S. 60, 66-67 (1983).

[43] *Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations*, 413 U.S. 376, 385 (1973) (emphasis added).

[44] *Bolger*, 463 U.S. at 71, n. 20.

[45] *Edenfield v. Fane*, 507 U.S. 763, 771 (1993).

protection, the Hawley bill requires edge providers to (1) "describe any policies … relating to restricting access to or availability of [user-generated] material" and (2) "promise that the edge provider shall … design and operate the provided service in good faith." The first requirement seems hands-off: it does not directly dictate what an edge provider's terms of service must say. But this is simply a trick of clever drafting: this requirement does not *need* to be specific, because the second requirement ("good faith") will, in practice, govern both. The two inquiries will collapse into one, allowing complaints about both the fairness of content moderation practices as compared to community standards, and the adequacy of those standards.

As a result, companies would (1) make their community standards as opaque or unspecific as possible and (2) minimize transparency about content moderation generally (*e.g.*, avoiding public statements or reporting on content removals). But relying on "good faith" does not solve the compelled speech First Amendment problem.

Suppose that, instead of suing to enforce Fox News' "Fair and Balanced" slogan in 2004, Congressional Democrats had proposed a bill like Hawley's: just replace "community standards" with "editorial standards" and apply the bill to cable programming networks over a certain size. It would be obvious that the government cannot compel traditional media companies to "describe any policies … relating to [selection] of [programming] material."

By contrast, the government *may* (and does) compel food manufacturers to disclose ingredient lists and nutritional information. The First Amendment permits such mandates because they apply to statements of *objective fact*, not the disclosure of opinions. This is why the seemingly simple age-based ratings systems for video games and movies have evolved as purely private undertakings. Behind each label is an editorial judgment, an *opinion*, about how to apply rating criteria. The government can compel neither the rating system overall, nor specific disclosures about the contents of specific films, nor disclosure of the rating methodology. By the same token, it cannot compel websites to disclose their editorial methodologies, whether implemented by humans or algorithms.[46]

X. The Hawley Bill Is Designed to Chill the Exercise of Editorial Discretion

The Hawley bill proposes four criteria for assessing a website's "good faith." The first two concern "selective enforcement," whether by humans or algorithms. But what purports to be a regulation only of marketing claims would actually, inevitably embroil regulators and/or judges in evaluating the editorial discretion of edge providers — conduct that would clearly qualify for the full protection of the First Amendment as non-commercial speech under *Miami Herald.* Twitter's alleged political bias in applying its community standards is no more actionable under consumer protection law than would be Fox News' political bias in its editorial policies.

The third criterion — "the intentional failure to honor a public or private promise made by, or on behalf of, the provider" — appears to preserve consumer protection claims, but its aim is significantly broader. In *Barnes v. Yahoo!, Inc.*,[47] the court allowed the plaintiff's suit against Yahoo! to proceed. Barnes sued the company for failing to stop her ex-boyfriend from posting

[46] *Brown*, 131 S. Ct. at 2740.

[47] 565 F.3d 560 (9th Cir. 2009).

revenge porn. The court ruled that the company had essentially waived its Section 230 immunity when its Director of Communications promised the plaintiff she would "personally walk the statements over to the division responsible for stopping unauthorized profiles and they would take care of it."

This promissory estoppel theory was limited to the particular facts of that case: a clear promise made directly to a *specific* user. The Hawley bill's "public or private promise" language could be read to allow plaintiffs to set aside Section 230 immunity and sue edge providers for far more general statements about content moderation practices that would never qualify for promissory estoppel. By holding companies to every past statement, the Hawley bill aims to stop companies from changing their content moderation policies over time as new challenges emerge — a critical dimension of any company's editorial discretion.

The fourth criterion — "any other intentional action taken by the provider without an honest belief and purpose, without observing fair dealing standards, or with fraudulent intent" — seems tailor-made for a law school exam on the "void for vagueness" standard. In particular, it is considerably more expansive than the narrow standard the Supreme Court set forth in *Central Hudson Gas Elec. v. Public Serv. Comm'n*, 447 U.S. 557 (1980), for regulating commercial speech: "there can be no constitutional objection to the suppression of commercial messages that do not accurately inform the public about lawful activity." In other words, the Court allows the regulation of commercial speech only because of its *effects*, not its intent. Applying a subjective, rather than an objective standard, would make litigation significantly easier. Thus, this criterion would not be constitutional even if it were applied solely to commercial speech. But as we have already seen with the Fox News example, there would be no way to apply this standard "without evaluating the content … at issue," as FTC Chairman Muris put it.

XI. The Bill Unconstitutionally Targets Specific Websites

The bill applies to "edge providers," defined as providers of a website, mobile application or web application with more than $1.5 billion in global revenue and more than 30 million U.S. users or more than 300 million global users, that have accessed the site by any means in the past year. This tailors the bill to apply to just a handful of services: Google (Alphabet), Apple, Facebook (including Instagram and Whatsapp) and Amazon (the so-called "GAFA") as well as Twitter, eBay, Microsoft, Apple, and TikTok (because the revenue threshold is global). Reddit, Flickr, and Etsy would meet the user thresholds but not the revenue thresholds. Wikipedia wouldn't be covered because it's a non-profit.

What may at first seem like a sensible way to focus the effect of the bill actually creates a host of problems. First, it's possible that, despite posing an existential threat to "Big Tech" companies, Hawley's bill could actually protect them from competition. By penalizing smaller market entrants for getting too big, Hawley's bill creates an incentive for small players to get bought-out by their "big tech" counterparts before crossing Hawley's size threshold — big companies better equipped to handle the legal risks Hawley's bill would create.

The bill's scope raises three distinct constitutional problems. First, singling out a small group of websites provides further reason for applying stricter scrutiny. "Minnesota's ink and paper tax violates the First Amendment not only because it singles out the press, but also because it targets

a small group of newspapers…. And when the exemption selects such a narrowly defined group to bear the full burden of the tax, the tax begins to resemble more a penalty for a few of the largest newspapers than an attempt to favor struggling smaller enterprises."[48] Applying taxes only to large newspapers "poses a particular danger of abuse by the State."[49]

Hawley's bill poses a "danger of abuse" by focusing on only the largest social networks — *all* of the ones conservatives complain about being biased against them — while excluding sites with a *laissez-faire* approach to content moderation, where extremist right-wing content has been allowed to flourish, such as Reddit. The relatively high revenue threshold excludes Reddit as well as other popular social media sites like Yelp (business reviews), IMDB (movie reviews), Fandom (a hosting platform), and Pinterest. The user threshold also excludes smaller social networks that have become gathering places for the Alt Right, like Gab (1.8 million monthly users users) and Minds (1.25 million users total).

The bill *might* apply to websites for traditional media, but even this is difficult to predict. Websites the largest newspapers and cable channels all meet the monthly user threshold, but won't qualify for the revenue threshold if separate corporate digital divisions are treated as the "edge providers" covered by the bill. In theory, it might be possible to "pierce the corporate veil" to argue that the parent companies' revenue should be counted, but this is *not* what the bill says — which further suggests the bill is tailored to social media sites. In any event, including some large traditional media websites in its scope wouldn't come anywhere near making the bill broad enough to avoid the concerns of *Minneapolis Star* or *Arkansas Writers' Project*.

Second, the bill applies only to a particular subset of Internet media — websites, apps and services that host user content, not services like Netflix or non Internet media. On its own, this all but ensures that the bill would be subject to strict scrutiny — which it would surely fail.[50]

Arguably, a bill that applied equally to all "interactive computer service providers" would be *less* problematic because it would not single out a "small group" of sites for what amounts to punishment. Abandoning user count or revenue thresholds would avoid the problem of retaliatory targeting, but additional First Amendment problems would remain.

XII. Hawley's Bill Would Backfire Against Conservatives

It's impossible to anticipate, *ex ante*, the net effect of the law upon the decision-making of each social media service — *i.e.*, whether they will do more or less moderation, and whether conservatives would actually benefit overall. The chief purpose of Section 230 was to avoid the "Moderator's Dilemma," created by *Stratton Oakmont, Inc. v. Prodigy Services Co.*, 1995 WL 323710 (N.Y. Sup. Ct. 1995). The court held Prodigy *more* liable because it actively engaged in content moderation to create a "family-friendly" service. If edge providers fear that removing

[48] *Minneapolis Star*, 460 U.S. at 591-92.

[49] *Arkansas Writers' Project, Inc. v. Ragland*, 481 U.S. 221 (1987).

[50] *See Turner Broadcasting System, Inc. v. FCC*, 512 U.S. 622 (1994) ("Regulations that discriminate among media ... often present serious First Amendment concerns."); *Minneapolis Star Tribune, Co. v. Minnesota Commr of Revenue*, 460 U.S. 575, 583 (1983) (a tax applied only to newspapers).

certain content may increase their legal risks, they will moderate less. On the other hand, they may calculate that *more* moderation will allow them to claim a more consistent approach.

That the same law could produce diametrically opposite results is not at all unusual in First Amendment jurisprudence. This is precisely the constitutional problem with vague laws: they are both unpredictable and highly subject to manipulation by those charged with enforcement.

Empowering the government to determine political neutrality cuts both ways. Discouraging edge providers from moderating incendiary or abusive speech from the right will have the same kinds of effects on the left. Democrats will just as easily claim "bias" when speech they like is removed. Consequently, social media sites will hesitate to take down content from extremist groups like Antifa or radical anti-police activists for fear that a Democratic FTC or state attorney general will sue them.

More generally, if Republicans start suing edge providers for failing to deliver on the claim of neutrality required by the new Hawley bill, you could count on Democrats — when they have the chance — to start suing social media operators for not living up to other provisions in their community standards. Consider Twitter's Community Standards:

Safety

Violence: You may not threaten violence against an individual or a group of people. We also prohibit the glorification of violence. Learn more about our violent threat and glorification of violence policies.

Terrorism/violent extremism: You may not threaten or promote terrorism or violent extremism. Learn more.

Child sexual exploitation: We have zero tolerance for child sexual exploitation on Twitter. Learn more.

Abuse/harassment: You may not engage in the targeted harassment of someone, or incite other people to do so. This includes wishing or hoping that someone experiences physical harm. Learn more.

Hateful conduct: You may not promote violence against, threaten, or harass other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. Learn more.

Twitter has made an editorial decision not to remove tweets posted by President Trump that seem to violate all of these prongs (minus the one about child sexual exploitation). The First Amendment clearly protects their right to make that decision, but if the government could hold a company to such statements about its editorial practices, as Hawley claims, without violating the First Amendment, why couldn't a Democratic FTC make the same argument about Twitter not living up to its promise to enforce its community standards? Indeed, Facebook has been heavily criticized[51] by groups on the left for failing to do more to take down racist content that may even incite users to violence.

For better or worse, the First Amendment prevents the government from forcing Facebook, Twitter or any other social media sites to change how they favor, disfavor, or remove user content. But if

[51] Kevin Rose, *Social Media Giants Support Racial Justice. Their Products Undermine It.*, NEW YORK TIMES (June 22, 2020), https://www.nytimes.com/2020/06/19/technology/facebook-youtube-twitter-black-lives-matter.html.

Hawley's bill were somehow to pass now, it could just as easily be used by a Biden administration to pressure social media sites to take down right-leaning content in the years it would take for the complex legal questions outlined here to work their way through the courts.

XIII. The "Problem" for Republicans Isn't 230, but the First Amendment

In the end, Republicans' complaints aren't really about Section 230, but about the First Amendment. Yes, Section 230 protects websites from liability for user content — "death by ten thousand duck-bites."[52] While the Hawley bill and Trump's Executive Order both make edge providers liable for what users say, this is only a means to an end; their real focus is not on the decision made by edge providers to host potentially unlawful content, but on their decision *not* to host content they deem objectionable. That decision is one the First Amendment protects as fully for websites as it does for newspapers or Fox News.

Trump, Hawley and other Republicans would do well to remember what President Reagan said when he vetoed legislation to restore the Fairness Doctrine back in 1987:

> We must not ignore the obvious intent of the First Amendment, which is to promote vigorous public debate and a diversity of viewpoints in the public forum as a whole, not in any particular medium, let alone in any particular journalistic outlet. History has shown that the dangers of an overly timid or biased press cannot be averted through bureaucratic regulation, but only through the freedom and competition that the First Amendment sought to guarantee.

Republicans should ask themselves: "WWRD—What Would Reagan Do?" The answer should, by now, be clear: "Congress shall make no law…"

[52] *Roommates*, 521 F.3d at 1174.

The Internet Archive National Emergency Library – A National Treasure In Uncertain Times or Mass Piracy?

By Joseph Petersen, James Trigg, and Olivia Poppens[1]

As libraries across the nation shuttered their doors to combat the spread of COVID-19, students and academics of all ages and disciplines suddenly faced an unprecedented obstacle. From the high school student writing a paper on Salinger, to the math professor studying ancient equations—virtually all lacked access to a physical library. The vast majority of scholarly work requires citing one or two sentences from ten, twenty or perhaps hundreds of diverse sources, which is why libraries have long been at the center of scholarly research. Without the traditional library, students and academics must find alternative sources for books that they would typically skim in a library, an often impossible task given the percentage of out-of-print works housed in libraries.

Enter the Internet Archive's National Emergency Library ("NEL"). Offering an online library is nothing new for the Internet Archive. It has in fact done so for years without legal challenge. Pre-pandemic, that library operated similarly to a traditional library by requiring borrowers to wait in line to rent copies that are already checked out. However, when libraries across the country began shutting down in the face of the pandemic, the Internet Archive abruptly eliminated its waitlist policy, opening its collection of nearly two million books to anyone with an internet connection, and allowing two-week rentals of all works without constraints.[2] This unprecedented action prompted acclaim from certain quarters and outrage in others. In fact, while this article was undergoing final edits, four major publishers—Hachette Books, HarperCollins, John Wiley & Sons, and Penguin Random House—collectively sued the Internet Archive for copyright infringement.[3] The publishers allege not only that the NEL constitutes broad copyright infringement, but that the entirety of the Internet Archive's library services do as well. Only days after the complaint was filed, the Internet Archive announced its decision to retire the NEL. Despite this fact, the legality of the NEL is likely to be litigated for purposes of determining whether the Internet Archive owed damages to the publishers, and the case will also address the Internet Archive's practice of Controlled Digital Lending ("CDL").

---

[1] Joe Petersen is partner in Kilpatrick Townsend's Silicon Valley office with more than two decades of experience representing a broad array of clients in litigation, arbitration and administrative proceedings involving copyright and trademark law. He frequently speaks on copyright and trademark issues before national organizations and routinely advises legal publications on cutting edge copyright issues. James Trigg is a partner in Kilpatrick's New York and Atlanta offices, and practices in the areas of copyright, trademark, and entertainment law, with experience in litigation, licensing, and general trademark and copyright counseling. James advises and advocates for clients across an array of industries, including media, fashion, textiles, and hospitality. Olivia Poppens is an associate in Kilpatrick's Los Angeles office and practices copyright and trademark law, with a specific focus on copyright litigation, as well as trademark and copyright enforcement both domestically and abroad.

[2] Chris Freeland, *Internet Archive Responds: Why we released the National Emergency Library*, Internet Archive Blog (Mar. 30, 2020), https://blog.archive.org/2020/03/30/internet-archive-responds-why-we-released-the-national-emergency-library

[3] Complaint & Demand for Jury Trial, *Hachette Book Group, Inc. v. Internet Archive*, No. 1:20-cv-04160 (S.D.N.Y. Jun. 1, 2020).

The Internet Archive Pre-COVID & The Recent Controversy

The Internet Archive has existed since 1996, functioning as a 501(c)(3) nonprofit that literally archives the internet. In addition, Internet Archive operates an internet library.[4] With over 1.8 million books available, the majority of which were published during the 20th century, California has recognized the Internet Archive as a library since 2007.[5] The library obtains hard copy books either through purchase or donation and scans the contents of those works and loads the digital copies to a server that is accessible via the internet.[6]

Prior to the pandemic, the Internet Archive would only lend the number of scanned copies that it physically owned. In other words, if the Internet Archive owned five copies of *Things Fall Apart*, only five scanned copies of that work would be offered at a time, and anyone beyond those five borrowers would be relegated to a waiting list until one of the copies was returned. Additionally, the Internet Archive focuses on books without corresponding e-books, specifically those from the 20th century, and repeatedly contended that its library is not a source for the latest potboiler.[7]

The Internet Archive's lending strategy, called CDL, has been endorsed by a variety of legal scholars.[8] Proponents of CDL have explicitly laid out how libraries can stay within the bounds of copyright: "a library may only loan simultaneously the number of copies that it has legitimately acquired, usually through purchase or donation."[9] In other words, CDL has relied on an "owned-to-loaned" ratio to support its legality under the copyright law.[10] This strategy allows digitized lending to function as closely to a physical, traditional library as possible. Supporters of CDL rely on two legal principles: the first sale doctrine (also called exhaustion) and fair use.[11] On the other hand, the complaint filed against the Internet Archive takes aim at the CDL scheme, calling controlled digital lending "a manufactured legal paradigm, conceived by IA, to cast aside well-established copyright jurisprudence," and accusing the non-profit of sponsoring the "White Paper

[4] *About*, Internet Archive Blog, http://blog.archive.org/about/ (last visited May 11, 2020).

[5] Adrian McCoy, *The Internet Gives Birth to an 'Official' Online Library*, post-gazette.com (Jun. 24, 2007), https://old.post-gazette.com/pg/07175/796164-96.stm.

[6] Chris Freeland, *Internet Archive Responds: Why we released the National Emergency Library*, Internet Archive Blog (Mar. 30, 2020), https://blog.archive.org/2020/03/30/internet-archive-responds-why-we-released-the-national-emergency-library; Constance Grady, *Why authors are so angry about the Internet Archive's Emergency Library*, Vox (Apr. 2, 2020), https://www.vox.com/culture/2020/4/2/21201193/emergency-library-internet-archive-controversy-coronavirus-pandemic

[7] Chris Freeland, *Internet Archive Responds: Why we released the National Emergency Library*, Internet Archive Blog (Mar. 30, 2020), https://blog.archive.org/2020/03/30/internet-archive-responds-why-we-released-the-national-emergency-library.

[8] David Hansen and Kyle Courtney, *A White Paper on Controlled Digital Lending of Library Books*, Controlled Digital Lending by Libraries, https://controlleddigitallending.org/whitepaper#_ftn1 (last accessed May 10, 2020).

[9] *Id.*

[10] *Id*.

[11] *See id.*

on Controlled Digital Lending of Library Books."[12] Opponents base their attacks on the claims that neither the first sale doctrine nor the fair use doctrine apply to CDL.

Despite the reasoning in the publishers' complaint, both publishers and authors alike have tolerated the Internet Archive's CDL service for years. The straw that broke the camel's proverbial back, however, was the library's decision to offer its digital collection to borrowers without regard to the number of physical copies of the works residing in the library's collection—an issue the publishers' complaint addresses on its head.

Anticipating backlash from authors and publishers alike, the Internet Archive created a form of take-down procedure for those who would prefer to have their works excluded from the NEL.[13] The FAQ page for the NEL directs authors to an email address, and requests that they list the URL for each book they would like removed, noting a 72 hour turn-around on such requests.[14] This system requires the copyright owner to find each of his works available on the website, and submit those titles to the Internet Archive for removal. Not surprisingly, publishers and authors have spoken out against the NEL, with some labeling it "piracy" and "illegal."[15] The publishers' complaint filed in June goes as far as to allege the "opt out" system is not always honored.[16] Authors like Chuck Wendig critiqued the service for "disrupting the chain of royalties that lead from books to author," as it "endangers" the authors' ability to "continue to produce art."[17] Notably, Wendig's works are not available in the National Emergency Library, either because his works are too recent, or because he decided to use the "opt out" feature.[18]

After the Internet Archive announced the NEL, the Author's Guild released a statement expressing its disdain for the NEL, and citing the plight of authors who are struggling more than ever in a rapidly contracting economy.[19] Other publishing groups made similar statements, with the President of the Association of American Publishers calling the Internet Archive's actions an

[12] Complaint & Demand for Jury Trial at 40, *Hachette Book Group, Inc. v. Internet Archive*, No. 1:20-cv-04160 (S.D.N.Y. Jun. 1, 2020).

[13] *National Emergency Library FAQs*, Internet Archive, https://help.archive.org/hc/en-us/articles/360042654251-National-Emergency-Library-FAQs (last accessed May 15, 2020).

[14] *See id.*

[15] Rachelle Hampton, *The Internet Archive Started an "Emergency" Online Library. Authors are Furious.*, Slate (Apr. 1, 2020), https://slate.com/culture/2020/04/internet-archive-national-emergency-library-controversy.html

[16] Complaint & Demand for Jury Trial at 44, *Hachette Book Group, Inc. v. Internet Archive*, No. 1:20-cv-04160 (S.D.N.Y. Jun. 1, 2020).

[17] Chuck Wendig, *My Statement to NPR on the Internet Archive's Emergency Library*, Chuck Wendig: Terribleminds (Mar. 31, 2020), http://terribleminds.com/ramble/2020/03/31/my-statement-to-npr-on-the-internet-archives-emergency-library/

[18] National Emergency Library, https://archive.org/details/nationalemergencylibrary (last accessed May 15, 2020).

[19] *Internet Archive's National Emergency Library Harms Authors*, The Authors Guild (Mar. 27, 2020), https://www.authorsguild.org/industry-advocacy/internet-archives-uncontrolled-digital-lending/

"aggressive, unlawful and opportunistic attack on the rights of authors and publishers in the midst of the novel coronavirus pandemic."[20]

Politicians have jumped into the fracas. For example, the head of the Senate Intellectual Property Committee, Thom Tillis, expressed his concern that the "'Library' is operating outside the boundaries of copyright law," and further noting that he is unaware of "any measure under copyright law that permits a user of copyright works to unilaterally create an emergency copyright act."[21] The senior senator of New Mexico, Tom Udall, wrote a letter to the Copyright Office regarding the National Emergency Library, asking the Office to provide guidance to "libraries, authors, and online outlets to identify potential solutions" that accommodate both copyright owners and those who need to access online collections.[22]

The Copyright Office responded to Senator Udall's letter, explaining that while it does not "provide legal advice about specific factual scenarios" it could provide impartial advice on matters of "copyright law and policy."[23] The letter lays out the ways libraries and school activities may be permitted under the Copyright Act, and ends by directly addressing the NEL in a manner strongly suggesting that the Copyright Office does not agree that the exigencies of the moment justify the opening of the Internet Archive's library in this manner.[24]

Readers and academics, however, have praised the NEL, with the *New Yorker* calling it a "gift to readers everywhere."[25] All readers who know about the library stand to benefit from this online repository, but the Internet Archive maintains that the main beneficiaries are students, academics, and researchers across the globe. The Internet Archive has also stated that the US taxpayer should enjoy this online privilege, as an estimated 650 million books are locked up in taxpayer-funded libraries now inaccessible due to COVID-19.[26] Some legal scholars, including Professor Pamela Samuelson of Berkeley Law, believe "that exigent circumstances—the pandemic, the indefinite closures of schools and libraries, and the many new burdens on families, including taking on educating their children and building their reading skills at a time when the parents are themselves burdened with their own or family illnesses and work or loss of work—

[20] *Comment from AAP President and CEO Maria Pallante on the Internet Archive's "National Emergency Library*, AAP Association of American Publishers (Mar. 27, 2020), https://publishers.org/news/comment-from-aap-president-and-ceo-maria-pallante-on-the-internet-archives-national-emergency-library/

[21] Porter Anderson, *US Senate IP Chief Questions Internet Archive's 'National Emergency Library,'* Publishing Perspectives (Apr. 9, 2020), https://publishingperspectives.com/2020/04/us-senate-subcommittee-chair-questions-internet-archives-national-emergency-library-covid19/

[22] Letter from Tom Udall, Senator of New Mexico, to Maria Strong, Acting Register of Copyrights (Apr. 16, 2020), https://www.copyright.gov/laws/hearings/Sen-Udall-Response-National-Emergency-Library.pdf?twcop=loclr.

[23] Letter from Maria Strong, Acting Register of Copyrights, to Tom Udall, Senator of New Mexico (May 15, 2020) https://www.copyright.gov/laws/hearings/Sen-Udall-Response-National-Emergency-Library.pdf?twcop=loclr.

[24] *See id.*

[25] Nicole Lepore, *The National Emergency Library is a Gift to Readers Everywhere*, The New Yorker (Mar. 26, 2020), https://www.newyorker.com/books/page-turner/the-national-emergency-library-is-a-gift-to-readers-everywhere.

[26] Chris Freeland, *Internet Archive Responds: Why we released the National Emergency Library*, Internet Archive Blog (Mar. 30, 2020), https://blog.archive.org/2020/03/30/internet-archive-responds-why-we-released-the-national-emergency-library.

alter the usual copyright calculus. Because the NEL collection mainly consists of older books which are either commercially inactive or unavailable in e-book form, the argument that it is depriving authors and publishers of substantial revenues is strained."[27] Also, nonprofit organizations like the Electronic Frontier Foundation have expressed their advocacy for the program on the grounds that the library is covered by fair use.[28]

The Legality of the Internet Archive's NEL and the Practice of CDL

Notwithstanding the polarized views on the benefits and burdens of the National Emergency Library and the general lending strategy of CDL, all appear in agreement that their legality under copyright law will turn on the doctrines of first sale and fair use.[29]

First Sale doctrine

The first sale doctrine, often called the "exhaustion" doctrine, is codified in section 109 of the Copyright Act. It provides that a copyright owner does not control the resale or other distribution of copies or phonorecords of a work that have been lawfully sold. It is this doctrine that authorizes libraries to lend physical copies of works without the permission of the rights holder, the idea being that once a copyright holder enjoys a benefit from the first sale of a particular copy of her work, her distribution right is "exhausted" as to that particular copy. The first sale doctrine is generally understood to implicate the right of distribution and not reproduction.[30]

The lead case on the first sale doctrine and digital reproduction and resale is the Second Circuit's decision in *Capitol Records, LLC v. ReDigi, Inc*.[31] In that case, Capitol Records sued ReDigi for facilitating the resale of lawfully purchased digital music files on its platform.[32] The Second Circuit held that ReDigi effectuated resales of digital music files in a way that resulted in the creation of at least one unauthorized reproduction, which in turn violated the copyright holder's

---

[27] Tim Zubizarreta, *All Your Books Are Belong to Us**, Jurist (Apr. 11, 2020), https://www.jurist.org/commentary/2020/04/brian-frye-emergency-library/; Telephone Interview with Pamela Samuelson, Professor of Law at Berkeley Law School, President and Chair of Board of Directors for the Authors Alliance (May 11, 2020).

[28] Corynne McSherry and Katharine Trendacosta, *Sharing our Common Culture in Uncommon Times*, Electronic Frontier Foundation (Apr. 10, 2020), https://www.eff.org/deeplinks/2020/04/sharing-our-common-culture-uncommon-times.

[29] *See* 17 U.S.C. § 109; 17 U.S.C. § 107; h David Hansen and Kyle Courtney, *A White Paper on Controlled Digital Lending of Library Books*, Controlled Digital Lending by Libraries, https://controlleddigitallending.org/whitepaper#_ftn1 (last accessed May 10, 2020).

[30] Brief for the Copyright Alliance as Amicus Curiae, p. 5-9, *Capitol Records, LLC v. ReDigi Inc.*, 910 F.3d 649 (2d. Cir. 2018), cert. denied, 139 S. Ct. 2760, 204 L. Ed. 2d 1148 (2019); *see also* 17 U.S.C. § 109(a); U.S. COPYRIGHT OFFICE, DMCA SECTION 104 REPORT 79-80 (2001), https://www.copyright.gov/reports/studies/dmca/sec-104-report-vol-1.pdf ("Section 109 limits a copyright owner's exclusive right of distribution. It does not, by its terms, serve as a defense to a claim of infringement of any of the other exclusive rights."); *see also ReDigi*, 910 F.3d at 655 ("Under the first sale doctrine, codified in § 109A), the rights holder's control over the distribution of any particular copy or phonorecord that was lawfully made effectively terminates when that copy or phonorecord is distributed to its first recipient.").

[31] *See ReDigi Inc.*, 910 F.3d 649.

[32] *See id*. at 652-53.

exclusive right of reproduction under 17 U.S.C. § 106(1).[33] The court cited the Copyright Register's conclusion that § 109 "does not apply to otherwise authorized digital transmissions of a copyrighted work, reasoning that such transmissions cause the recipient to obtain a new copy of the work."[34]

As explained, the first sale doctrine allows the owner of a text to copy, "sell or otherwise dispose" of that *particular* copy of a book.[35] Proponents of CDL note that courts struggle to identify what a "particular" copy is for purposes of the first sale doctrine, but tend to agree that the CDL scheme would fall into this doctrine. On the other hand, the publishers' complaint specifically addresses the first sale doctrine as it relates to CDL and argues that the first sale doctrine only applies to the distribution right, *not* the reproduction right. They contend that scanning and uploading a work is a reproduction, rather than a distribution as occurs with the lending of a physical book.[36]

As for the NEL, the Copyright Office's response to Senator Udall's April 16, 2020 letter leans heavily on the *ReDigi* decision in its summary of the first sale doctrine. The Office interprets *ReDigi* broadly as foreclosing reliance on the first sale doctrine for any reproductions of copyrighted works.[37] Under this interpretation, since the NEL is distributing digital *reproductions* of the books in its collection, and not the physical books themselves, the limitation set out in Section 109 simply does not apply. Similarly, this is the same argument the publishers' complaint sets forth for why CDL is not legal under the copyright law.

Kyle Courtney, the Copyright Advisor for Harvard University, differs with the Office's expansive interpretation of *ReDigi*.[38] He argues that the Second Circuit did not intend to shut the door on "digital first sale for non-profit, educational uses," but rather shut the door on "bit-for-bit replicas" of the original mp3 files that caused market harm.[39] The difference between books and licensed mp3 copies is a difference Courtney believes cannot be overlooked, specifically as to the fact that the substitutionary effect of these digitized books cannot be easily ascertained.[40]

Mr. Courtney's argument is certainly the more nuanced (which often in litigation can be reason enough to handicap its chance of success). Its success may require detailed proofs regarding the substitutionary effects of the NEL's open access policy. Although this argument is constrained as

---

[33] *See id.* at 659.

[34] *See id.* (quoting Digital Millennium Copyright Act, Pub. L. No. 105-304, 122 Stat. 2860, 2876 (1998))

[35] David Hansen and Kyle Courtney, *A White Paper on Controlled Digital Lending of Library Books*, Controlled Digital Lending by Libraries, https://controlleddigitallending.org/whitepaper#_ftn1 (last accessed May 10, 2020).

[36] Complaint & Demand for Jury Trial at 41-42, *Hachette Book Group, Inc. v. Internet Archive*, No. 1:20-cv-04160 (S.D.N.Y. Jun. 1, 2020).

[37] Letter from Maria Strong, Acting Register of Copyrights to Tom Udall, Senator of New Mexico (May 15, 2020), https://www.copyright.gov/laws/hearings/Sen-Udall-Response-National-Emergency-Library.pdf?twcop=loclr.

[38] Kyle Courtney, *Libraries Do Not Need Permission to Lend Books: Fair Use, First Sale, and the Fallacy of Licensing Culture*, Kyle K. Courtney (May 18, 2020), https://kylecourtney.com/2020/05/18/libraries-do-not-need-permission-fair-use-first-sale-and-the-fallacy-of-permission-culture/

[39] *Id.*

[40] *See id.*

applied to the NEL, it will likely be fleshed out in litigation as it applies to CDL.[41] In reality, the first sale argument is arguably more persuasive under CDL, given the "one-to-one" ratio is generally maintained under that lending strategy. Nonetheless, such evidence may be equally if not more relevant to the question of fair use.

Fair Use

The main event of the NEL's legality, as well as that of CDL, likely turns on whether the use is defensible under the fair use doctrine. Section 107 of the Copyright Act provides that "the fair use of a copyrighted work, including such use by reproduction in copies . . . is not an infringement of copyright." The statutory preamble lists several illustrative, but not limitative, potentially fair uses, including use "for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research."

In determining whether a use is fair use, courts are directed to consider four factors: "(1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work."[42] Although the first and fourth factors of this test have repeatedly driven the fair use analysis, the balancing test allows for flexibility and robust debate as to whether a particular use is "fair" under copyright law.[43]

Two cases from the last decade guide the analysis here: *Authors Guild v. HathiTrust* and *Authors Guild v. Google* (often known as the "Google Books" case).[44] In both cases, the Authors Guild sued respective defendants for copyright infringement arising from the mass digitization of millions of books previously found exclusively in physical libraries across the country. Most analogous to the current controversy, the Second Circuit in *HathiTrust* reviewed the Authors Guild's challenge to the HathiTrust Digital Library, an electronic library containing over 10 million works.[45] The HathiTrust uploaded scanned books to allow for, among other things, a "full-text searchable database" available to all members, as well as digitized versions of copyrighted works for print disabled patrons.[46] Significantly, access to the full text of in-copyright works was denied to all except those with certified print-disabilities.

---

[41] Complaint & Demand for Jury Trial at 41-42, *Hachette Book Group, Inc. v. Internet Archive*, No. 1:20-cv-04160 (S.D.N.Y. Jun. 1, 2020).

[42] 17 U.S.C. § 107

[43] *See Authors Guild v. Google Inc.*, 804 F.3d 202, 220 (2d Cir. 2015).

[44] *See Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d. Cir. 2014); *Google Books*, 804 F.3d 202. Joe Petersen and Kilpatrick Townsend & Stockton represented Hathitrust in the *HathiTrust* case, and the firm was similarly involved in the *Google Books* case in early proceedings.

[45] *See HathiTrust*, 755 F.3d 87.

[46] *See id*. at 105.

The *Google Books* case was decided soon after *HathiTrust,* and involved a copyright challenge by the Authors Guild against Google for its digital library.[47] Google scanned millions of books to its digital library, ultimately allowing internet users to search for a particular word or phrase in over 20 million books. The digital library provided the user with a "snippet view" of the page in which the search term appeared.[48] This function only provided the user with about one-eighth of the page, and prohibited users from reading the entire book by "blacklisting" one page of every ten.[49] The court held that Google's "unauthorized digitizing of copyright-protected works, creation of search functionality, and display of snippets from those works are non-infringing fair uses."[50] Both cases prove highly relevant to the current debate.

Fair Use: Factor One

Courts have held that under the first factor analysis, more transformative uses are likely to tip the scales in favor of a finding of fair use. Sufficient transformation has been found when the secondary use adds a new purpose or character to the work, and when the new use expands the utility of the original.[51] Further, an educational purpose or a purpose that promotes the "copyright law's goal of 'promoting the Progress of Science and useful Arts,' U.S. Const., art. I, § 8, cl. 8" also weighs in favor of this factor.[52] Some courts have recognized that uses mentioned specifically in the text of § 107—"criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research""—can tip the scales in favor of fair use.[53] Finally, courts are more likely to find a secondary use "fair when it produces a value that benefits the broader public interest."[54]

Opponents of the NEL and CDL may claim that not only does the Internet Archive fail to transform the original copyrighted work in any way but it also does not expand the copyrighted work's utility. Proponents for both the NEL and CDL argue that the use is transformative under a "format shifting" approach, essentially that the services shift the format from physical to digital. This argument is dismissed as illegitimate by the publishers in their complaint, with respect to both CDL and the NEL.[55] Further, critics argue that unlike the secondary uses in *HathiTrust* and

---

[47] *Google Books*, 755 F.3d at 209-10.

[48] *See id.*

[49] *Id.*

[50] *See id*. at 229.

[51] *See id*.; *see Google Books,* 804 F.3d 202, 214 (2d Cir. 2015).

[52] *Castle Rock Entm't, Inc. v. Carol Pub. Grp., Inc*., 150 F.3d 132, 141 (quoting *Arica Inst., Inc. v. Palmer,* 970 F.2d 1067, 1077 (2d Cir. 1992) (alteration incorporated)); see also *Bill Graham Archives v. Dorling Kindersley Ltd*., 448 F.3d 605, 608 (2d Cir. 2006).

[53] *See NXIVM Corp. v. The Ross Inst.*, 364 F.3d 471, 477 (2d Cir. 2004); *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578-79, 114 S.Ct. 1164, 127 L.E2d 500 (1994) (holding that this analysis may be guided by the preamble of § 107).

[54] *Blanch v. Koons*, 467 F.3d 244, 253 (2d Cir. 2006) (citing *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 922 (2d Cir. 1994)).

[55] Complaint & Demand for Jury Trial at 42, *Hachette Book Group, Inc. v. Internet Archive*, No. 1:20-cv-04160 (S.D.N.Y. Jun. 1, 2020) ("False premise that a print book and a digital book share the same qualities. But, as outlined above, they are fundamentally different mediums, and they exist as distinct economic markets.").

Google Books, the Internet Archive is not adding utility in a way that makes the use transformative.[56] In *Google Books*, the defendants only allowed users small snippets of their collection for the purpose of search, while in *HathiTrust*, defendants limited the uses to "full-text" searches and access for the print-disabled.[57]

Proponents of the Internet Archive's position may claim that the library expands the utility of the hard copy books to internet users in a time when those hard copy books would be wholly inaccessible otherwise. Such arguments certainly gain force with respect to books that are currently out of print and which are unavailable in digital form. It is unclear, however, the extent to which the NEL is also comprised of works that are only a click away on popular online bookseller sites.

Those in favor of CDL also claim that that first sale doctrine should "positively influence the 'purpose and character' assessment in fair use," because (in proponents' view) CDL is an updated version of lending that is protected under the Copyright Act.[58] Recognizing that this may not be a slam dunk argument, proponents also rely on the concept of "format shifting" to justify their fair use defense.[59] Although there have been cases that reject "format shifting," in and of itself, as a basis for finding transformative use, those who support CDL distinguish this process from those cases as they believe CDL goes to the core of the copyright law, advancing "public knowledge" and "enriching" the public through access to creative works.[60]

Creative arguments aside, the NEL's use is not likely to be deemed transformative. The legality of CDL under this factor is a closer case. The central issue for both the NEL and CDL hinges on whether the public benefit of the service and lending strategy can compensate for its relative lack of transformative purpose. In the *HathiTrust* case, full text access for the print-disabled was not "transformative" in the way that term is typically construed in the fair use analysis. Nonetheless, the public benefit was so overwhelming that both the district court and the Second Circuit had little hesitation in finding that the first fair use factor favored the universities. A court could similarly conclude that the overwhelming public benefit of the NEL is compelling enough reason to overcome the lack of transformation given the unprecedented circumstances of the time.[61] Similarly, a court could find that CDL's goal of advancing "public knowledge" could also satisfy the transformative requirement under the first factor. On the other hand, the lack of transformative use in conjunction with the differences between the use in *HathiTrust* and the use

[56] *Id.*

[57] *See Google Books*, 804 F.3d at 217 (citing *HathiTrust*, 755 F.3d at 98).

[58] David Hansen and Kyle Courtney, *A White Paper on Controlled Digital Lending of Library Books*, Controlled Digital Lending by Libraries, https://controlleddigitallending.org/whitepaper#_ftn1 (last accessed May 10, 2020).

[59] *Id.*

[60] *Id.*; c.f. *Disney Enterprises, Inc. v. VidAngel, Inc.*, 869 F.2d 848, 861-62 (9th Cir. 2017); *Wall Data Inc. v. Los Angeles County Sheriff's Dept.*, 447 F.3d 769, 778-79 (9th Cir. 2006).

[61] *See NXIVM*, 364 F.3d at 477; *Campbell*, 510 U.S. at 578-79 (holding that this analysis may be guided by the preamble of § 107).

at issue both with the NEL and CDL, may lead a court to find that the first factor weighs against a finding of fair use.[62]

Fair Use: Factors Two and Three

Factors two and three, the nature of the work, and the amount and substantiality of the portion used in relation to the copyrighted work as a whole, are rarely if ever outcome determinative. Nonetheless, they probably do weigh against the Internet Archive with regard to both the CDL and NEL.[63] While academic, informational, or nonfiction works are more likely to fall in the purview of fair use, the Internet Archive lists hundreds of thousands of books, not all of which are purely academic. Moreover, the amount and substantiality factor requires an evaluation of whether the "quantity and value of the materials used" are reasonable.[64] Here, the Internet Archive will argue, with some force, that in order to provide a substitute for now shuttered physical libraries it had no choice but to open the entirety of its library through the NEL in this manner (and thus the use is reasonable in light of the project's goals). Similarly, the CDL strategy may be defended on these grounds, given a variety of books are necessary to maintain a digital library that is analogous to physical libraries. That said, here as well this argument could be undermined if it can be demonstrated that the library includes an inordinate amount of works that may be purchased at the click of the button on the internet.

Fair Use: Factor Four

The fourth factor has been called "undoubtedly the single most important element of fair use."[65] Undergirding the significance of the fourth factor in the fair use analysis is the simple fact that the fair use doctrine seeks an ever elusive balance of promoting the progress of science through access to works while attempting not to stifle the creation of those works in the first place. The application of this factor to the NEL and the CDL is likely to diverge more than any of the other legal doctrines discussed.

Reasonably, courts have found that when the secondary use is more transformative, an adverse commercial effect is less certain and harder to infer. [66] The entire analysis focuses "on whether the copy brings to the marketplace a competing substitute for the original, or its derivative, so as to deprive the rights holder of significant revenues because of the likelihood that potential purchasers may opt to acquire the copy in preference to the original."[67]

The Internet Archive contends that both CDL and the NEL are being used in the same way traditional libraries always have, with the majority of borrowers only using books for thirty

---

[62] *See HathiTrust*, 755 F.3d 87.

[63] 17 U.S.C. § 107

[64] *Campbell*, 510 U.S. at 586, 144 S.Ct. 1164 (quoting *Folsom v. Marsh*, 9 F. Cas. 342, 348 (C.C.D. Mass. 1841)).

[65] *Harper & Row Publishers, Inc. v. Nation Enter.*, 471 U.S. 539, 566, 105 S.Ct. 2218, 85 L.Ed.2d 588 (1985).

[66] *See Campbell*, 510 U.S. at 591; *Blanch*, 467 F.3d at 258.

[67] *Fox News Network, LLC v. Tveyes, Inc.*, 883 F.3d 169 (2d Cir. 2018), cert. denied, 139 S. Ct. 595, 202 L. Ed. 2d 428 (2018) (quoting *Google Books.*, 804 F.3d at 223 (2d Cir. 2015)).

minutes or less, and with fewer than 10% of users opening the rental books after the first day.[68] Those who argued for the CDL prior to the release of the NEL maintained that there is no recognized market that the Internet Archive affects as the library functions analogously to a traditional library.[69] Those proponents ultimately concluded that there is not a cognizable copyright injury caused by CDL.[70] Those in favor of the NEL argue that users are skimming books for research and scholastic purposes, essentially arguing the emergency library has no more an effect on the potential market for a copyrighted work than traditional libraries. Perhaps most illustrative of this argument is the fact the NEL does not include *any* books published in the last five years. In fact, 90% of the books in the library were published over 10 years ago, and two-thirds of the books available were published during the 20th century.[71] Studies have shown that the commercial value of a book declines rapidly after its initial release date, making it all the more likely, proponents assert, that the commercial impact of this endeavor is negligible to authors and publishers.[72]

Opponents of the Internet Archive's NEL may argue that regardless of the age of the books, the NEL is indeed serving as a substitute for the physical book, as the user is choosing to borrow that copy rather than buy the book in store (an argument that necessarily - and in many instances dubiously - assumes that a book would be available for purchase). Despite the fact that the books are predominately used for scholarship, the fact that anyone may check out the books gives weight to the argument that the NEL *is* substituting the market for these in-copyright works by lending unlimited copies of those in the repository. The publishers' complaint even goes as far argue that, among other harms, the NEL devalues the *entire* book market by influencing consumers into believing books are "cheap" and actually buying a physical book or ebook is unnecessary.[73]

This same substitution argument is extended to the entire concept of CDL, as opponents argue that digital books do not deteriorate over time like physical books, nor do they require physical transportation. These facts drive opponents' argument that the economic rights of authors will be "severely diminished" if libraries are able to "circumvent the rightsholder entirely and copy

---

[68] Brewster Kahle, *The National Emergency Library – Who Needs it? Who Reads It? Lessons from the First Two Weeks*, Internet Archive Blogs (Apr. 7, 2020), http://blog.archive.org/2020/04/07/the-national-emergency-library-who-needs-it-who-reads-it-lessons-from-the-first-two-weeks/; *see also* Bailey, Lila Bailey, Kyle Courtney, et al., David Hansen and Kyle Courtney, *Position Statement on Controlled Digital Lending*, Controlled Digital Lending by Libraries, https://controlleddigitallending.org/statement (last accessed June 17, 2020).

[69] Bailey, Lila Bailey, Kyle Courtney, et al., David Hansen and Kyle Courtney, *Position Statement on Controlled Digital Lending*, Controlled Digital Lending by Libraries, https://controlleddigitallending.org/statement (last accessed June 17, 2020).

[70] *See id.*

[71] Chris Freeland, *Internet Archive Responds: Why we released the National Emergency Library*, Internet Archive Blog (Mar. 30, 2020), https://blog.archive.org/2020/03/30/internet-archive-responds-why-we-released-the-national-emergency-library.

[72] Chris Freeland, *Internet Archive Responds: Why we released the National Emergency Library*, Internet Archive Blog (Mar. 30, 2020), https://blog.archive.org/2020/03/30/internet-archive-responds-why-we-released-the-national-emergency-library.

[73] Complaint & Demand for Jury Trial at 45, *Hachette Book Group, Inc. v. Internet Archive*, No. 1:20-cv-04160 (S.D.N.Y. Jun. 1, 2020).

millions of print books into digital copies to be widely distributed, even if it maintains an 'owned to loaned ratio.'"[74]

Given the potential for some substitutionary effects of both the NEL and CDL, the critical question will involve determining the degree of that effect. In *HathiTrust*, the Second Circuit found full-text access for the print disabled to be fair use based, in part, on a finding that the "present-day market for books accessible to the handicapped is so insignificant" that authors often forgo royalties "generated through the sale of books manufactured in specialized formats for the blind."[75] This brings up an interesting argument for both sides.

On the one hand, those against the NEL may argue that unlike the copies created for the blind created by HathiTrust, the copies created by the NEL are of conventional books for which authors still recoup royalties. Proponents of the NEL, however, are likely to point to studies evidencing that authors' royalties greatly decline after the initial release of a work, and further argue that a large amount of these books *do not* have corresponding e-books.[76]

The Internet Archive and its proponents might also argue that the PDF scans are inferior to e-books or physical copies which a reader could buy or check out from a traditional library. Based on this assertion they could in turn argue that the market for authorized reproductions is not significantly usurped.[77] While the quality of the Internet Archive's books may be debatable, the fact that the books are currently free for all suggests at least some degree of substitutionary effects. In contrast, the Internet Archive's argument as it relates to CDL generally will hinge on the fact that the "one-to-one" ratio does not hinder the market for authors and publishers any more so than a physical library. Since users must wait in line to check out a book, those who don't have time to wait may still buy the book, making a limited impact on the market for books available on the Internet Archive before the release of the NEL.

The United States has not experienced a crisis at the scale of COVID-19 in living memory. As such, case law discussing the applicability of copyright law under such exigent circumstances is non-existent.[78] From a purely precedential perspective, the Internet Archive's NEL may not sustain a fair use argument in normal times. That said, these are extraordinary times and whether they call for extraordinary copyright measures is subject to question and robust debate. Even if the court finds the NEL to have been a bridge too far, it may nonetheless conclude that the CDL strikes the appropriate balance between protection of authors' rights, on the one hand, and the furtherance of science and the arts, on the other.

---

[74] *Id.* at 43.

[75] *See id.* at 103.

[76] Chris Freeland, *Internet Archive Responds: Why we released the National Emergency Library*, Internet Archive Blog (Mar. 30, 2020), https://blog.archive.org/2020/03/30/internet-archive-responds-why-we-released-the-national-emergency-library.

[77] *See id.*

[78] *Golan v. Gonzales*, 501 F.3d 1179, 1191-92 (10th Cir. 2007) discussed exigent circumstances under the Emergency Copyright Act of 1941 in the context of compliance with procedural rules, as authors during the war may not have been able to comply with copyright formalities. Unlike the case at issue here, this altered copyright holders' requirements, as opposed to what others could do with copyrighted works.

What to Expect Next?

While the NEL is coming to a close, the fight surrounding its legality will continue in the event the publishers seek damages arising from its operation between March and June 2020. And, notwithstanding the NEL's cessation, the parties presumably will continue to battle over the legality of CDL. Standing requirements, however, may limit the extent to which the publishers can prevail.[79] The Copyright Act requires that only "the legal or beneficial owner of an exclusive right . . . [can] institute an action for any infringement of that particular right committed while he or she is the owner of it."[80] Courts have found that groups like the Authors Guild do not have "standing to bring suit on behalf of their members."[81] Essentially, only copyright owners (or licensees) can bring suit against the Internet Archive, and those copyright owners can only bring suit for works they specifically own. While the four publishers that brought suit against the Internet Archive maintain a vast amount of copyrights, including copyrights of works housed in the Internet Archive, they will only succeed in enjoining the Internet Archive from distributing the copyrighted works they own.

Whether the publishers' case proceeds to trial, more litigation is brought against the Internet Archive or, perhaps Congress steps in to tackle these issues head on, the debate around the legality of both the NEL and the practice of CDL will likely continue for years to come.

[79] Complaint & Demand for Jury Trial at 51, *Hachette Book Group, Inc. v. Internet Archive*, No. 1:20-cv-04160 (S.D.N.Y. Jun. 1, 2020).

[80] *See* 17 U.S.C. § 501(b).

[81] *HathiTrust*, 755 F.3d at 94 (citing *ABKO Music, Inc. V. Harrisongs Music, Ltd.*, 944 F.2d 971, 980 (2d Cir. 1991)); *see also Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 104 S. Ct. 774, 78 L. Ed. 2d 574 (1984) (holding that the infringement of plaintiffs' own copyrighted works provided the requisite standing, and that they had "no right to invoke whatever rights other copyright holders may have to bring infringement actions" against Sony).

CCPA Private Litigation, Recent Developments and Potential New Privacy Legislation on the Horizon

By Jim Snell, Marina Gatto and Gabriella Gallego[1]

The California Consumer Privacy Act (CCPA) went into effect over five months ago, on January 1, 2020. Although enforcement by the California Attorney General cannot begin until July 1, private plaintiffs have been bringing claims under the law's limited private right of action since before the beginning of the year.

In addition, the California Attorney General has been working on regulations that are now in final form but still not effective and may not be effective until after the July 1 date by which Attorney General enforcement can begin. Companies are struggling to work on compliance with just finalized but still not legally enforceable regulations.

Further, there have been efforts to put a ballot initiative on the November ballot that would substantially amend the CCPA.

This article will (1) discuss CCPA private litigation trends to date, (2) summarize the California Attorney General regulations and (3) summarize a ballot initiative that may be on the November 2020 ballot.

CCPA Litigation and the Private Right of Action

The CCPA provides a limited private right of action for data breaches affecting certain categories of personal information where the breach is the result of a lack of reasonable security. More specifically, section 1798.150(a)(1)(A) of the CCPA states that a private litigant may bring a cause of action only if their "nonencrypted and nonredacted personal information . . . is subject to an unauthorized access and exfiltration, theft, or disclosure as a result of the business's violation of the duty to implement and maintain reasonable security procedures and practices appropriate to the nature of the information to protect the personal information." Several of the private complaints filed thus far allege claims under this section of the CCPA; however, private claims also allege violations of the CCPA not subject to the private right of action as well as claims for violations of other laws by referencing alleged CCPA violations. For example, in response to an alleged data breach by online video communications company Zoom, a plaintiff alleges a CCPA claim under the private right of action--that Zoom failed to implement reasonable security standards which resulted in the unauthorized disclosure of unredacted personal information. Complaint at 11, *Robert Cullen v. Zoom Video Commc'ns, Inc.*, No. 5:20-cv-02155 SVK (N.D. Cal. filed March 3, 2020). However, the plaintiff also alleges violations outside of the private right of action, including that Zoom allegedly violated the CCPA "by, among other things, collecting and using personal information without providing consumers with

[1] Partner James Snell represents clients in a wide range of complex commercial matters, including privacy and security, internet, marketing, and intellectual property litigation. Associates Marina Gatto and Gabriella Gallego assist companies with data security and privacy compliance matters, including compliance with the California Consumer Privacy Act (CCPA) and the General Data Protection Regulation (GDPR). Jim, Marina and Gabriella are all based in Perkins Coie's Palo Alto office.

adequate notice consistent with the CCPA, in violation of Civil Code section 1798.100(b)." *Id.* at 10. While § 1798.100(b) is a requirement of the CCPA, it is expressly not one that a private litigant is permitted to enforce. *See* Civil Code 1798.150 ("The cause of action established by this section shall apply only to violations as defined in subdivision (a) and shall not be based on violations of any other section of this title.").

Similarly, in *Johnston v. Zoom Video Communications, Inc.*, No. 5:20-cv-02376 (N.D. Cal. filed Apr. 8, 2020), Johnston alleges that Zoom violated the CCPA by using personal information of the alleged class without providing the notice required by § 1798.100(b) (categories and purposes of PI collection) and § 1798.120(b) (right to opt-out of sale) and for false and deceptive behavior regarding Zoom's sale of data. Johnston claims that Zoom's affirmative statements in its privacy policy that it **does not** sell user data are inaccurate. Again, these claims are not permitted to be brought by a private plaintiff under the CCPA, but were nevertheless made.

In addition to limiting private CCPA claims to a narrow category of data breach cases, the CCPA also explicitly limits a claimant's rights to assert CCPA violations under other claims. Specifically, the CCPA provides that "nothing in this title shall be interpreted to serve as the basis for a private right of action under any other law." 1798.150(C). Despite this restriction, plaintiffs have filed cases alleging CCPA violations as the basis for other claims, including, for example, violations of the California Unfair Competition Law (UCL) (codified at Cal. Bus & Prof. Code §§ 17200 et seq.). *See, e.g.*, Complaint, *Cullen*, No. 5:20-cv-02155 SVK; Complaint, *Burke v. Clearview AI, Inc.*, No. 3:20-cv-00370 BAS MSB (S.D. Cal. filed Feb. 27, 2020); Complaint, *Almeida v. Slickwraps Inc.*, No. 2:20-cv-00559-TNL-CKD (E.D. Cal. filed Mar. 12, 2020); Complaint, *Dennis v. First Am. Title Co.*, No. 8:19-cv-01305 (C.D. Cal. filed July 1, 2019) (class action suit alluding to defendant's reference to the CCPA in its 10-K filings in support of plaintiff's UCL claim). CCPA violations have also been asserted under intrusion upon seclusion and constitutional privacy grounds.

The lesson from these cases is that plaintiffs are asserting the CCPA broadly and that it will be the job of the courts to apply the law in the narrow way the private right of action was intended. Given how recently such cases have been filed, and given the court delays resulting from the Covid-19 pandemic, we have yet to see orders from courts narrowing these cases. While we await such orders, companies should expect and plan for broad allegations to be brought forth, and consider steps to minimize such claims from private plaintiffs.

Modifications to the Draft Regulations

In addition to the evolving private CCPA litigation, the California Attorney General has been preparing for CCPA enforcement which will begin on July 1, 2020, and has also recently finalized the CCPA regulations. The Attorney General has also made clear in a press release that he will not postpone enforcement of the statute despite the global COVID-19 pandemic, reminding California residents that it is "more important than ever for Californians to know their privacy rights." As such, companies should expect Attorney General enforcement to begin July 1, 2020.

The Attorney General has also been working on CCPA regulations for more than a year, and submitted final regulations to the Office of Administrative Law (OAL) on June 1, 2020. The

OAL has 30 working days plus an additional 60 calendar days to determine whether the regulations satisfy the procedural requirements of the Administrative Procedure Act, though the Attorney General has asked that review be expedited so the regulations can be final by the July 1, 2020 enforcement date. Once approved, the final regulations will be filed with the Secretary of State and become enforceable. One question is how the Attorney General will view enforcement of the CCPA before the regulations are effective, and we speculate that initial enforcement may focus on the language of the CCPA itself, with enforcement of the regulations to follow, after they become effective. In the meantime, companies are working to incorporate the recently finalized regulations into their compliance programs.

The final regulations contain additional important detail, including a provision that states that a violation of the regulations shall be deemed a violation of the CCPA. While a careful review should be made of the regulations, we highlight four areas below that companies should consider.

Privacy Policy Requirements

The final regulations contain additional guidance and requirements for what should be included in a business's privacy policy. Among other things, the regulations require that "[f]or each category of personal information" disclosed or sold, the business also list "the categories of third parties to whom the information was disclosed or sold." Cal. Code Regs. tit. 11, § 999.308(c)(1)(g)(2). The regulations also require that a privacy policy contain the date it was last updated as well as information whereby a consumer with a question or concern about the business's privacy policy and practices can contact the business "using a method reflecting the manner in which the business primarily interacts with the consumer." Id. § 999.308(c)(6)-(7).

A privacy policy must also include instructions on how an authorized agent can make a request on behalf of a consumer, as well as details regarding how the business will verify the consumer's request.

Methods for *Verification*

One of the biggest challenges businesses are facing with the sudden influx of consumer rights requests is how to verify requests. The final regulations provide some guidance as to how businesses should approach verifying consumer requests. For example, the regulations would give businesses the explicit ability to deny requests that cannot be verified within 45 days. Id. § 999.313(b). The regulations also state that a business "shall not" respond to access requests for specific pieces of information that cannot be verified. *Id.* §999.325(f). Two examples are provided for how a business could comply with this requirement. One example specifies that a retailer that maintains purchase history information may require the consumer to identify his or her recent purchases, or the dollar amount of his or her most recent purchase. *Id.* §999.325(e)(1). However, if a business is not engaged in retail but maintains a mobile app, a suggested verification method is to ask consumers to provide information that only the person using the mobile app would know, or require they respond to a notification sent to their device. Id. § 999.325(e)(2). The Attorney General also addressed verification requirements for consumers submitting rights requests through authorized agents by stating that businesses may "directly confirm" with the consumer that he or she did in fact grant the agent signed permission to submit

a request on his or her behalf. This addition may help businesses avoid phony authorized agent requests that have not in fact been authorized by a consumer.

Service Providers Rights and Restrictions

The Regulations also clarify some of the rights and restrictions related to service providers. Specifically, the regulations identify five ways service providers may use, retain, or disclose personal information, including: (1) "[t]o process or maintain personal information on behalf of the business that provided the personal information, or that directed the service provider to collect the personal information, and in compliance with the written contract for services required by the CCPA"; (2) "[t]o retain or employ another service provider as a subcontractor"; (3) "[f]or internal use by the service provider to build or improve the quality of its services," provided, however, that the "use does not include building or modifying household or consumer profiles to use in providing services to another business, or correcting or augmenting data acquired from another source." *Id.* § 999.314(c)(1)-(3).

Notice Requirements

The Regulations impose additional notice requirements that should be reviewed carefully. The regulations also attempt to clarify some ambiguity in the CCPA as to who needs to provide notice. For example, the Regulations confirm that data brokers who may not have a means of providing notice "at or before the point of collection," are not required to provide such notice so long as they include in their registration with the Attorney General a link to their online privacy policy that includes instructions on how a consumer can submit an opt-out request. *Id.* § 999.305(e).

Financial Incentives

The Regulations include detailed requirements regarding financial incentives and price or service differences related to the collection, retention, or sale of personal information. These provisions require disclosures that include opt-in and opt-out rights, specific notice requirements, as well as anti-discrimination provisions. These provisions should be carefully reviewed by any business who may be providing financial incentives, and we recommend that such review take place under privilege given the ambiguity that exists in the CCPA and the Regulations regarding financial incentives.

Potential New Legislation: The California Privacy Rights Act

There is also another piece of legislation potentially on the horizon in 2021. The California Privacy Rights Act ("CPRA") is a ballot initiative that is authored by the same individual whose efforts in 2018 resulted in the California legislature enacting the CCPA. At the time of writing this article the future of the CPRA remains unclear, although it may be likely that this initiative will appear on the November 2020 ballot. Although the signatures gathered to place this initiative on the ballot have yet to be verified by the Secretary of State, one thing that is for certain is that if enacted, the CPRA would amend the CCPA in several significant ways, some of which we discuss below.

Creation of a New State Agency

The CPRA would create the California Privacy Protection Agency (the "Agency") which could potentially dramatically impact regulatory enforcement of the CCPA in California. The Agency would be responsible for enforcing the CPRA through administrative enforcement actions. Decisions resulting from such actions would be subject to judicial review in an action brought by an interested party and subject to an abuse of discretion standard.

Rulemaking power would also be transferred from the California Attorney General to the Agency on the later of July 1, 2021, or 6 months after the Agency provided notice to the Attorney General that it was prepared to begin rulemaking.

New Right to Correction

The CPRA would also give consumers new rights, such as the right to request that a business that maintains their personal information correct such inaccurate personal information. Businesses would also be required to disclose to consumers the right to make such a request, and upon receiving a request use commercially reasonable efforts to correct the inaccurate personal information at issue.

New Requirements for Sensitive Personal Information

The CPRA would also create the newly defined term "sensitive personal information," which would include personal information that reveals a consumer's social security number, driver's license or state identification card number, account login information, financial account information, precise geolocation, racial or ethnic origin or religious beliefs, in addition to other information. Consumers would also be given new rights with respect to their sensitive personal information, including the right to limit the use and disclosure of their sensitive personal information and the right to be informed at or before the point of collection as to the categories of sensitive personal information to be collected and the purposes for which the categories will be used and whether such information is sold or shared, in addition to the length of time the business intends to retain each category of sensitive personal information.

Broader Time Frame for Access Rights

The CPRA would also go beyond what is currently required under the CCPA by extending the look-back period for a consumer's right to request access to their personal information. With respect to personal information collected on or after January 1, 2022, the CPRA provides that a consumer can request that "the business disclose the required information beyond the 12-month period and the business shall be required to provide such information unless doing so proves impossible or would involve a disproportionate effort."

Added Clarity

In addition to imposing new obligations on businesses, the CPRA would also provide some additional clarity. For example, the CPRA would make clear that a business's requirement to disclose certain information upon request will not require a business to disclose trade secret information. The CPRA would also make clear that the right to be free from discrimination is not

absolute, and that a business is not prohibited from "offering loyalty, rewards, premium features, discounts, or club card programs consistent with this title."

Purpose Limitations

The CPRA would also limit a business's collection, use, retention, and sharing of a consumer's personal information to that which is "reasonably necessary and proportionate to achieve the purposes for which the personal information was collected or processed, or for another disclosed purpose that is compatible with the context in which the personal information was collected, and not further processed in a manner that is incompatible with those purposes." Therefore, a business would not be allowed to retain a consumer's personal information or sensitive personal information for longer than is reasonably necessary for the purpose for which the business disclosed that the personal information was being collected.

Additional Opt-Out Mechanisms

The CPRA would broaden the opt-out rights that California consumers are currently afforded by extending such opt-out rights beyond personal information "sold," to include the right to opt-out of the "sharing" of personal information and the right to opt-out of the disclosure and use of "sensitive personal information." "Sharing" is defined under the CPRA to include "sharing, renting, releasing, disclosing, disseminating, making available, transferring, or otherwise communicating orally, in writing, or by electronic or other means, a consumer's personal information by the business to a third party for cross-context behavioral advertising, whether or not for monetary or other valuable consideration, including transactions between a business and a third party for cross-context behavioral advertising for the benefit of a business in which no money is exchanged."

Does the First Amendment Include a Right to Scrape Photographs from Public Websites?

By Jeff Hermes[1]

On February 5, 2020, Hoan Ton-That, the CEO of controversial facial recognition company Clearview AI, claimed on *CBS This Morning* that "there is also a First Amendment right to public information" when questioned about his company's practice of scraping public-facing websites for photographs to add to its database of more than 3 billion images. *Google, YouTube, Venmo and LinkedIn send cease-and-desist letters to facial recognition app that helps law enforcement*, CBS This Morning (Feb. 5, 2020), https://www.cbsnews.com/news/clearview-ai-google-youtube-send-cease-and-desist-letter-to-facial-recognition-app/. However, the assertion that the First Amendment protects this conduct has been met with skepticism from both privacy advocates and First Amendment lawyers. *See* Margot E. Kaminski and Scott Skinner-Thompson, *Free Speech Isn't a Free Pass for Privacy Violations*, Slate (Mar. 9, 2020), https://slate.com/technology/2020/03/free-speech-privacy-clearview-ai-maine-isps.html; Alfred Ng, *Clearview AI says the First Amendment lets it scrape the internet. Lawyers disagree*, CNET (Feb. 6, 2020), https://www.cnet.com/news/clearview-says-first-amendment-lets-it-scrape-the-internet-lawyers-disagree/.

Prompted by this discussion, this article considers whether the First Amendment prohibits a ban on scraping of publicly available photographs, herein defined as the automated collection of photographs in bulk (as opposed to one-by-one access to and collection of photographs at the direction of a human being).

Background

While this article does not analyze the legality of Clearview's business model, some additional background on Clearview's tool is helpful in order to understand why questions about facial recognition relate to scraping of data from the internet. The company describes its tool as a "search engine for publicly available images," Rebecca Heilweil, *The world's scariest facial recognition company, explained*, Recode (May 8, 2020), https://www.vox.com/recode/2020/2/11/21131991/clearview-ai-facial-recognition-database-law-enforcement. Clearview's tool allows a user to upload a photo of a person and receive in response a collection of other publicly-available photos of that person with links to where they appear on the internet. Kashmir Hill, *The Secretive Company That Might End Privacy as We Know It*, New York Times (Feb. 10, 2020), https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html.

There are few public details about how the tool works (other than that it uses a "state-of-the-art neural net"), and data regarding the accuracy of the tool is inconsistent. Heilweil, *The world's scariest facial recognition company*. The primary distinguishing factor between Clearview and other facial recognition databases is the size of the database against which users can run a search:

[1] Jeff Hermes is a Deputy Director of MLRC.

3 billion images scraped from public-facing websites versus approximately 400 million in the FBI's database, and many fewer in state and local databases. Hill, *The Secretive Company*.

It is the scope of Clearview's data collection that has raised eyebrows and triggered a host of other privacy-related questions, particularly as the site has apparently collected at least some of its data in violation of website terms of service. Critics have called out the potential abuse of such a system, which is now used by more than 600 law enforcement agencies, because of its pervasiveness, its ability to link individuals to their social media activity, and the fact that with larger databases there are more likely to be "doppelgängers" who could be mistaken for one another. *Id.* But as discussed at the start of this article, this is information that Clearview's CEO believes he has a First Amendment right to collect.

Does a Ban on Photo Scraping Raise First Amendment Issues?

The analysis in this article will start by asking whether a ban on scraping of photographs implicates the First Amendment at all. From there, it examines a ban on scraping as a "manner" restriction on speech, albeit an atypical one. Under that rubric, the analysis will consider potential content-based and content-neutral justifications for a restriction on scraping, and whether such restrictions allow ample alternative channels for the receipt and analysis of the information at issue. With the answer to those questions determining whether a scraping ban is subject to strict or lesser scrutiny, the analysis will conclude by examining the strength of the government interests at stake and whether a ban on scraping can be adequately tailored to meet the appropriate constitutional standard.

To focus more clearly on the constitutional question, I will for the most part put aside issues related to website terms of service that prohibit automated scraping. I will similarly bypass the question of online material that requires specific permission from the poster to access (for example, material that can only be viewed with an invitation to a private group or after acceptance of a friend request). And I will save until the end of the article the question of whether the constitutional issues discussed may be addressed by limiting a ban to scraping photos for the purpose of developing a facial recognition system.

As a general principle, the Supreme Court has recognized that a "[l]aw enacted to control or suppress speech may operate at different points in the speech process," *Citizens United v. FEC*, 558 U.S. 310, 340 (2010), and that "whether government regulation applies to creating, distributing, or consuming speech makes no difference," *Brown v. Entm't Merchants Ass'n*, 564 U.S. 786, 792 n.1 (2011). This only makes sense; a speaker's First Amendment rights have little value if restrictions prevent an audience from being in the time or place to hear the speaker's message, or from using the tools necessary to do so.

Moreover, the Supreme Court has repeatedly held that the First Amendment protects both the transmission and receipt of information. *See Packingham v. North Carolina*, 137 S.Ct. 1730, 1737 (2017) (ban on use of social media unconstitutionally "bars access to what for many are the principal sources for knowing current events, checking ads for employment, speaking and listening in the modern public square, and otherwise exploring the vast realms of human thought and knowledge"); *see also Kleindienst v. Mandel*, 408 U.S. 753 , 762-65 (1972) (First Amendment interests of academics who wished to hear in person from foreign speaker were

implicated by speaker's exclusion from United States); *Lamont v. Postmaster General*, 381 U.S. 301, 305 (1965) (postal service's conditioning of receipt of "communist political propaganda" on addressee's submission of written request violates addressee's First Amendment rights).

The U.S. District Court for the District of Columbia directly considered the application of the First Amendment to scraping in *Sandvig v. Sessions*, 315 F. Supp. 3d 1 (D.D.C. 2018) ("*Sandvig I*"). The case involved researchers who claimed that they needed to violate websites' terms of service in order to scrape sufficient information to determine whether those websites engaged in discriminatory behavior; fearful of prosecution under the federal Computer Fraud and Abuse Act, they filed a pre-enforcement declaratory judgment action asserting that the First Amendment protected their activity. On the federal government's motion to dismiss for lack of standing, the court held that a prohibition on scraping plausibly implicated the plaintiffs' First Amendment rights, stating:

> "[T]he First Amendment goes beyond protection of the press and the self-expression of individuals to prohibit government from limiting the stock of information from which members of the public may draw." *First Nat. Bank of Boston v. Bellotti*, 435 U.S. 765, 783, 98 S.Ct. 1407, 55 L.Ed.2d 707 (1978). … [S]ix courts of appeals have found that individuals have a First Amendment right to record at least some matters of public interest, in order to preserve and disseminate ideas. That plaintiffs wish to scrape data from websites rather than manually record information does not change the analysis. Scraping is merely a technological advance that makes information collection easier; it is not meaningfully different from using a tape recorder instead of taking written notes, or using the panorama function on a smartphone instead of taking a series of photos from different positions. And, as already discussed, the information plaintiffs seek is located in a public forum.

Sandvig I at 15-16.

But there is a lurking issue here, namely that the First Amendment rights of a recipient of information are premised on the existence of a willing speaker. *Va. State Bd. of Pharm. v. Va. Citizens Consumer Council*, 425 U.S. 748, 756 (1976) ("Freedom of speech presupposes a willing speaker. But where a speaker exists, . . . the protection afforded is to the communication, to its source and to its recipients both."). The court in *Sandvig I* found specifically that the plaintiffs were not seeking special access to information denied to others, thus distinguishing the plaintiffs' claims from cases where the Supreme Court had held that the First Amendment does not compel others to supply information. *Sandvig I* at 17-18.

Many if not most people who post photos online would object that in doing so they did not intend to disclose their photos for processing by companies like Clearview AI (or for that matter, by Google for its image search function or by other companies for any number of purposes). Thus, one might argue, the First Amendment should not come into play in the first place, because the posters of the photos were not willingly conveying the information to serve a scraper's purposes. But that argument, while perhaps facially appealing, is dangerous because it would grant speakers a radical degree of control over their audiences.

It is fundamental in any communication that a speaker's intended meaning and an audience's received meaning can be different. Indeed, it might be rare (or actually impossible) that these two quantities are identical because information received by a listener will always be filtered through the listener's own perspective and synthesized with the listener's own prior knowledge. When we speak, we often convey more information than we think we are conveying, whether through unconscious expressions such as body language or because an audience knows something we do not and can therefore draw conclusions of which we are unaware.

For that reason, the "willingness to speak" on which a recipient's First Amendment rights depend must necessarily be the speaker's willingness to engage in the act of communication, and not be defined by the speaker's subjective understanding of the meaning being conveyed or how the recipient will use that information. Imagine the consequences, for example, if politicians could negate the First Amendment rights of the press by claiming that the press was deriving information from their statements that the politicians did not mean to disclose. This does not mean that individuals' interests in how their photos are used are irrelevant; it just means that those interests should be considered as a factor in First Amendment scrutiny rather than as a basis to declare that no scrutiny is required.

Scrutiny of a Scraping Ban as a "Manner" Regulation

Presuming that the First Amendment applies, a direct ban on scraping photos is properly parsed as a restriction on communications technology; it thus falls into the category of time, place or manner (specifically, "manner") restrictions on communication. *See Sandvig I*, 315 F.Supp.3d at 16 (comparing scraping to use of tools such as tape recorders or smartphones). It is somewhat unusual because a typical manner regulation is focused on the methods and/or tools used by the speaker, while a ban on scraping places restrictions on the tools used by a would-be recipient of the information. But concomitant with the recognition of a First Amendment right to receive information is the fact that such a right can be burdened through time, place and manner restrictions just like the right of a speaker.

As a manner restriction, a scraping ban will be considered constitutional if it is "narrowly tailored to serve a significant government interest," provided that it also "leave[s] open ample alternative channels for communication of the information" and is "justified without reference to the content of the regulated speech." *Ward v. Rock Against Racism*, 491 U.S. 781, 791 (1989). If a manner regulation fails either of the latter qualifications, it would instead be subject to strict scrutiny. *See Reed v. Town of Gilbert*, 135 S.Ct. 2218, 2227 (2015) (laws that cannot be justified without reference to their content are subject to strict scrutiny); *United States v. Grace*, 461 U.S. 171, 177 (1983) (restrictions that function as a complete prohibition of a particular type of expression must be "narrowly drawn to accomplish a compelling governmental interest"); *Linmark Assocs., Inc. v. Willingboro*, 431 U.S. 85, 93-94 (1977) (municipal sign ordinance subject to strict scrutiny, not time/place/manner scrutiny, where it was doubtful that it left open alternative channels and was not genuinely concerned with place or manner as opposed to content).

Is a Ban on Scraping Content-Neutral?

It is easy to imagine that a ban on photo scraping might be intended specifically to prevent data processing that is perceived to threaten personal privacy. Clearly, a ban on photo scraping that is explicitly based on a desire to prevent the collection or interpretation of the information contained in the photos would not be content-neutral. But what if the legislature enacting such a ban is more subtle, and voices other concerns such as interference by automated data collection with the operation of websites and databases?

As a starting point, it is possible to argue that a statute does not become content-based just because it focuses on digital image files to the exclusion of other material. The argument would run that an image is a form in which information can be presented, and is not equivalent to the information itself. That is to say, the image is the package, not the "content." Moreover, because image files tend to be larger than text files, it is conceivable that the automated accessing of such data could have a greater impact on a website's ability to manage and to fulfill users' data requests within the capacity of a given computer system. Parsed as a limitation on file size, this starts to sound more like a content-neutral restriction than a content-based restriction.

However, this would beg the question of why it does not include other large files (such as video files), which are distinguishable solely based on the nature of the information in the file. And this argument truly falls apart if a ban focuses specifically on photographs, which are distinguishable from other image files based solely on what they depict.

But suppose that instead of specifying images or photos, a scraping ban was limited with direct reference to file size (e.g., no scraping of files over 5MB). A law banning automated collection of large photos, because of the speed with which the requests are made and the corresponding burden placed on the web server to deliver large amounts of information, could plausibly be justified by reference to non-content-based considerations such as the data management and system stability issues suggested above. Moreover, such a version could be described as content-neutral on its face, even if larger files might commonly contain visual or audiovisual data. This, however, implicates the somewhat more complex question of whether a scraping ban would satisfy the "ample alternative channels" analysis.

Availability of Alternative Channels

On the question of "ample alternative channels," a restriction on automated bulk scraping does not prevent the manual gathering of photographic data made available to the public. (I use the term "manual" advisedly, here; naturally, any access to online material will be through computerized processes as browser and server communicate with one another and the server supplies the requested data without any human intervention on the server's end.) It might, however, prevent data gathering from being worthwhile to the entity performing the scraping; reduced to manual acquisition, an entity such as Clearview AI might not be able to afford the personnel necessary to gather data efficiently enough to populate a database of sufficient scope to be useful.

The Supreme Court has held that the alternative channels analysis does take considerations of cost and convenience into account with respect to a speaker's ability to reach an audience. *See*

City of Ladue v. Gilleo, 512 U.S. 43, 56-57 (1994) (ban on display of signs on private property did not leave open ample alternative channels, where alternatives such as newspaper advertisements or leafletting were more expensive and time consuming). It is logical that such considerations would extend to the audience's methods of receiving a speaker's message as well. Consider, for example, the ban on the petitioner's use of social media at issue in *Packingham v. North Carolina*. While the case does not explicitly conduct a time/place/manner analysis, the Supreme Court's discussion focuses on the unique utility of social media sites as channels to both speak and receive information. Much of the information to which the Court was concerned that the petitioner was denied access (news about current events, employment ads, and more) would have been available through other channels with enough effort, but the Court found a First Amendment violation in denying him "perhaps the most powerful mechanisms available to a private citizen." 137 S.Ct. at 1737.

But while the use of social media and the use of scraping both involve electronic tools that provide access to information at rates of efficiency previously unheard of, a social media site is a channel selected and used by both the speaker and the listener in a particular communication. In contrast, scraping tools are uniquely used by those seeking information, and denial of their use does not limit the selection of communication media available to a speaker. One could cabin the alternative channels analysis by defining a "channel" for the purposes of the analysis to include only the tools necessary for an audience member to receive information in the form that a speaker chooses to convey it, and not those which are merely convenient for the recipient. So, the use of a radio or television set to receive broadcast signals would be part of a channel of communication, but not necessarily use of a DVR to time-shift programming. A speaker who elects to publish content on social media requires their audience to use a browser or other tool to access a social media platform, but does not necessarily require the use of automated bulk collection tools.

One problem with limiting the analysis in this fashion is that the transmission and receipt of information is not the entirety of the communication process. On the speaker's end, it also involves the composition of ideas and information into a communicable form (referred to as "encoding"), while on the listener's end it involves the interpretation, analysis and synthesis of received information by the audience (referred to as "decoding"). Freedom of speech can be suppressed through restrictions on encoding and decoding just as easily as through restrictions on transmission and reception. For example, if a news organization were prohibited from publishing photographs or video, that would be a manner restriction on encoding (presenting information in the form of visual images); similarly, if a news organization were prohibited from using statistical software to interpret financial data that it has drawn from public records, that would be a manner restriction on decoding (analysis of received information). In either case, however, there is little doubt that the restriction would be found to be a manner restriction that does not allow reasonable alternatives.

What is potentially confusing about a ban on bulk scraping is that it can affect both reception (imposing a limitation on the manner in which information is accessed) and decoding (limiting the information that can be derived from gathered data). Depending on the purposes for which the scraping is being performed, reducing the overall amount of material being gathered could impair the derivation of information from that material. For example, a research study based on scraped data might need a database of particular size to produce reliable results. If a scraping ban

impairs the meaningful synthesis of public information to produce new insights, it could make the availability of ample alternatives a much closer question because there might not be other efficient methods to gather the necessary information at the necessary scale.

Note that there is a potential difference between the size of a database necessary for synthesis of new information and the size necessary to be practically worthwhile. A facial recognition system based on machine learning might be trained to distinguish faces through exposure to many photos of different people, with the system's accuracy dependent on the number of photos supplied. On the other hand, a simpler facial recognition system might be based on facial geometry data derived from individual photos; in that case, the number of photos in the database would relate to how many results are potentially returned and not necessarily the accuracy of those results. In the former case, a scraping ban would impair the actual derivation of information, while in the latter case, it would reduce the utility of the tool but not impair its underlying technology.

It is not obvious that the latter form of interference burdens First Amendment interests in the same sense as the former. Regardless, a ban on automated scraping is likely to burden a substantial enough amount of data gathering and synthesis of the former type to raise First Amendment questions under an overbreadth analysis. *See U.S. v. Stevens*, 130 S.Ct. 1577, 1587 (2010) ("[A] law may be invalidated as overbroad if a substantial number of its applications are unconstitutional, judged in relation to the statute's plainly legitimate sweep.").

Government Interests

Regardless of whether strict or time-place-manner scrutiny applies, the analysis will consider of the nature of the government interest motivating a scraping ban; the difference is whether that interest must be "compelling" or merely "significant." *Compare Reed*, 135 S.Ct. at 2231 (strict scrutiny) *with McCullen v. Coakley*, 134 S.Ct. 2518, 2534 (2014) (time-place-manner scrutiny). There are at least two potential privacy interests that could be at stake: (1) an interest in preventing the aggregation of one's photo into a database; and (2) an interest in protecting against the processing of one's photograph for biometric data. Both of these privacy interests are concerned with the acquisition and processing of information from the photographs, and so are inherently content-based. They must therefore arise to the level of a "compelling" government interest in order for a ban to survive strict scrutiny.

With respect to the first interest, our starting assumption that the individual photos in question are freely available to the public does not necessarily resolve the question of whether there can be a privacy interest implicated in the aggregation of such photos. Courts have recognized in other contexts that privacy interests can exist in compilations of information even when there might be a limited privacy interest in the discrete pieces of information compiled.

In *U.S. Dep't of Justice v. Reporters Committee for Freedom of the Press*, 489 U.S. 749 (1989), the Court considered whether the disclosure of criminal identification records (a/k/a "rap sheets") pursuant to a Freedom of Information Act request would fall within a statutory FOIA exception, 5 U.S.C. § 552(b)(7)(C), because it "could reasonably be expected to constitute an unwarranted invasion of personal privacy." The rap sheets collected information about their subjects' criminal activity that had previously been made available to the public, and the Reporters Committee for

Freedom of the Press argued that that an individual's interest in protecting against the disclosure of a federally-created compilation of such information was trivial. 489 U.S. at 762-63.

The Supreme Court rejected that argument, calling it a "cramped notion of personal privacy." *Id.* at 763. it instead recognized a privacy value in the "practical obscurity" of various "bits of information" scattered throughout the public record: "Plainly there is a vast difference between the public records that might be found after a diligent search of courthouse files, county archives, and local police stations throughout the country and a computerized summary located in a single clearinghouse of information." *Id.* at 764. That privacy interest, said the Court, is "substantial," and "is affected by the fact that in today's society the computer can accumulate and store information that would otherwise have surely been forgotten long before a person attains age 80, when the FBI's rap sheets are discarded." *Id.* at 771.

The idea that the privacy interest in the whole can be greater than that in its parts is also the basis of the "mosaic" theory of constitutional search and seizure law, which arose in response to the capacity of a government to use technological measures to amass information about its citizens. As articulated by the Supreme Judicial Court of Massachusetts, which has been one of the courts at the forefront of accepting this theory:

> When collected for a long enough period, the cumulative nature of the information collected implicates a privacy interest on the part of the individual who is the target of the tracking. … Although these activities, taken one by one, may not give rise to a reasonable expectation of privacy, the Court aggregates their sum total for its analysis. … As the analogy goes, the color of a single stone depicts little, but by stepping back one can see a complete mosaic. … This aggregation principle or mosaic theory is wholly consistent with the statement in *Katz* [*v. United States*, 389 U.S. 347, 351 (1967)], that "[w]hat a person knowingly exposes to the public … is not a subject of Fourth Amendment protection," because the whole of one's movements, even if they are all individually public, are not knowingly exposed in the aggregate.

Commonwealth v. McCarthy, No. SJC-12750, slip op. at 19-20 (April 16, 2020) (internal citations and quotation marks omitted).

While not explicitly adopting the mosaic test, the Supreme Court of the United States addressed similar concerns in *Carpenter v. United States*, 138 S.Ct. 2206 (2018), in which it considered the Fourth Amendment implications of the warrantless search of compiled historic cell site location information ("CSLI") to determine the movements of a criminal suspect. Time-stamped CSLI records are generated whenever a cell phone connects to a cellular tower or other antenna in a service provider's network, and are collected by providers for their own purposes. Because a CSLI record indicates proximity to a particular tower at a particular time, it is possible to use historic CSLI to chart the movements of a given person (or at least their cell phone). Federal prosecutors obtained 127 days' worth of robbery suspect Timothy Carpenter's historic CSLI without a warrant and used it at trial to place him near the scene of four robberies at the time they occurred.

On Carpenter's motion to suppress, the government argued that CSLI is gathered by private companies and therefore not subject to the Fourth Amendment's warrant requirement under the third-party doctrine. *See Smith v. Maryland*, 442 U.S. 735, 743 (1979) (no expectation of privacy in phone numbers voluntarily transmitted to third party phone company by dialing one's telephone). The district court and the Sixth Circuit agreed, but the Supreme Court held that notwithstanding the third-party doctrine the prosecutors' acquisition of Carpenter's data was a Fourth Amendment search. The Court found that the CSLI compiled "over the course of 127 days provides an all-encompassing record of the holder's whereabouts. As with GPS information, the time-stamped data provides an intimate window into a person's life, revealing not only his particular movements, but through them his familial, political, professional, religious, and sexual associations." *Carpenter*, 138 S.Ct. at 2217. The potential to assemble such a comprehensive history of a cell phone holder's activities, said the Court, made CSLI qualitatively different from other records held by third parties and thus the third-party doctrine did not apply.

That said, recognizing privacy interests under FOIA or the Fourth Amendment is not the same as identifying a compelling government interest in a First Amendment analysis. *See Reporters Committee*, 489 U.S. at 762 n.13 ("The question of the statutory meaning of privacy under the FOIA is, of course, not the same as the question whether a tort action might lie for invasion of privacy or the question whether an individual's interest in privacy is protected by the Constitution."). Ruling that a privacy interest is sufficient to allow the government to withhold a compilation of information "that might be found after a diligent search" is not equivalent to finding that such an interest is sufficient to prohibit a private party from conducting that search. Nor is the aggregation of data in the case of photo scraping the same as that considered in either *Reporters Committee* or in the mosaic theory; the former aggregates individual data points regarding many different people, while the latter involve the aggregation of many data points about a single person. Thus, concerns about amassed data revealing more information about an individual than discrete data would not apply with the same force. It is thus questionable whether a privacy interest in avoiding compilation of one's photos with those of others would be considered compelling.

Concerns about the acquisition and potential misuse of biometric information are more recent in their origin; Illinois' Biometric Information Privacy Act, 740 ILCS 14/1 et seq. ("BIPA"), enacted in 2008, was the first law in the United States to regulate private entities' use of this information. BIPA prohibits the unauthorized collection and disclosure of biometric data, and flatly bans the commercial exploitation of such data. 740 ICLS 14/15. In explaining the rationale for these prohibitions, Illinois' General Assembly stated, inter alia:

> Biometrics are unlike other unique identifiers that are used to access finances or other sensitive information. For example, social security numbers, when compromised, can be changed. Biometrics, however, are biologically unique to the individual; therefore, once compromised, the individual has no recourse, is at heightened risk for identity theft, and is likely to withdraw from biometric-facilitated transactions.

740 ICLS 14/5(c).

To date, there have been no appellate opinions considering a direct constitutional challenge to the statute, but there have been cases discussing whether a violation of the statute constitutes a substantive injury to the person whose biometric information is allegedly obtained without authorization. In *Rosenbach v. Six Flags Entertainment Corp.*, 129 N.E.3d 1197 (Ill. 2019), the Supreme Court of Illinois considered whether a plaintiff needs to plead specific injury beyond a violation of BIPA in order to be an "aggrieved" person entitled to bring a civil action under the statute. Answering "no" to this question, the court found that a violation of the statute caused cognizable harm in and of itself:

> These procedural protections are particularly crucial in our digital world because technology now permits the wholesale collection and storage of an individual's unique biometric identifiers—identifiers that cannot be changed if compromised or misused. … When a private entity fails to adhere to the statutory procedures, as defendants are alleged to have done here, the right of the individual to maintain his or her biometric privacy vanishes into thin air. The precise harm the Illinois legislature sought to prevent is then realized. … This is no mere "technicality." The injury is real and significant.

129 N.E.3d at 1206 (internal citations and quotation marks omitted). Similarly, in *Patel v. Facebook, Inc.*, 932 F.3d 1264 (9th Cir. 2019), the Ninth Circuit held that allegations of a BIPA violation were sufficient to support Article III standing without alleging some further particularized injury, stating:

> Technological advances provide access to a category of information otherwise unknowable, … and implicate privacy concerns in a manner as different from traditional intrusions as a ride on horseback is different from a flight to the moon[.] … In light of this historical background and the Supreme Court's views regarding enhanced technological intrusions on the right to privacy, we conclude that an invasion of an individual's biometric privacy rights has a close relationship to a harm that has traditionally been regarded as providing a basis for a lawsuit in English or American courts. … The facial-recognition technology at issue here can obtain information that is detailed, encyclopedic, and effortlessly compiled, which would be almost impossible without such technology. … Taking into account the future development of such technology …, it seems likely that a face-mapped individual could be identified from a surveillance photo taken on the streets or in an office building. Or a biometric face template could be used to unlock the face recognition lock on that individual's cell phone. We conclude that the development of a face template using facial-recognition technology without consent (as alleged here) invades an individual's private affairs and concrete interests. Similar conduct is actionable at common law.

932 F.3d at 1273 (internal citations and quotation marks omitted).

These opinions recognize that BIPA violations represent an impairment of privacy interests comparable to those addressed by other privacy torts. That at least suggests parity with other governmental interests in protecting individual privacy that have previously survived First

Amendment scrutiny. It seems possible that if faced with the question directly, courts would recognize that the state's interest in protecting biometric privacy is compelling.

However, these interests could also be viewed in a different light in a First Amendment analysis. Certainly, merely acquiring photographs of others in public has not been recognized to pose privacy issues. *See, e.g., Schifano v. Greene County Greyhound Park, Inc.*, 624 So.2d 178, 180 (Ala. 1993) (photograph of subjects in public park did not violate their privacy); *Gill v. Hearst Publ'g Co.*, 40 Cal.2d 224, 230 (1953) ("Consistent with their own voluntary assumption of this particular pose in a public place, plaintiffs' right to privacy as to this photographed incident ceased and it in effect became a part of the public domain."); *Mark v. Seattle Times*, 96 Wn.2d 473, 499 (1981) (photograph of subject in store visible from public space did not violate right of privacy); *see also* Restatement (Second) of Torts, § 652B cmt. c ("Nor is there liability for observing [a person] or even taking his photograph while he is walking on the public highway, since he is not then in seclusion, and his appearance is public and open to the public eye."). Similarly, there is no legitimate privacy interest in photographs that the subjects themselves have shared publicly. *See Brewer v. Hustler Magazine, Inc.*, 749 F.2d 527, 529-30 (9th Cir. 1984) (plaintiff has no privacy interest in photo that plaintiff himself had already distributed publicly); *McMann v. Doe*, 460 F.Supp.2d 259, 268 (D. Mass. 2006) ("[D]istributing a publicly available portrait photograph … resemble[s] publishing appearances made in a public place. These activities do not impinge th[e] right of privacy."). This raises the question of whether there is a legitimate privacy interest in facial geometry that has been voluntarily revealed by sharing a photograph, or only in how such data is subsequently used.

What about non-content-based justifications for a scraping ban, such as protecting against interference with website operation? As discussed above, a law premised on such concerns could be phrased in a facially content-neutral manner; whether it faces strict or time-place-manner scrutiny would therefore depend on whether the ban fails the "ample alternative channels" test. There is little doubt that protecting the operation of online websites against technical interference would be considered at least a significant state interest. *See Turner Broad. Sys. v. FCC*, 512 U.S. 622, 647 (1994) (protecting households against loss of regular television broadcasting service "is not only a permissible governmental justification, but an important and substantial federal interest"). Moreover, it could arise to the level of a compelling interest if phrased in terms of securing public access to information. *See id.* at 663 ("[A]ssuring that the public has access to a multiplicity of information sources is a governmental purpose of the highest order, for it promotes values central to the First Amendment.").

Tailoring Considerations

Perhaps the most significant challenge for a scraping ban in the constitutional analysis is the requirement that it be tailored to serve the governmental interests behind the law. Under strict scrutiny, the regulation must be the "least restrictive means" of achieving the compelling state interest at stake. *McCullen*, 134 S.Ct. at 2530. Under the lesser scrutiny accorded to neutral time, place, and manner restrictions, the regulation at issue must not "burden substantially more speech than is necessary to further the government's legitimate interests," but "will not be invalid simply because a court concludes that the government's interest could be adequately served by some less-speech-restrictive alternative." *Ward*, 491 U.S. at 799-800.

In the case of the privacy interests described above, a scraping ban protects those interests not by directly prohibiting the problematic use of the photos in question, but by making the collection of photos for processing too cumbersome to be profitable. As is almost always the case when a law regulates speech in order to prevent some other undesired conduct, the prohibition is likely to prevent activity having nothing to do with the government interests at stake. A ban on automated collection of photographs will prevent not only the assembly of facial recognition databases or the derivation of biometric data, but also the use of scraping in academic research, indexing of visual content by services such as Google and Bing, and other contexts that do not pose privacy concerns. This is not the least restrictive alternative, given that direct prohibitions on the compilation of photos or mining of biometric information would achieve the same ends without overreaching.

As far as a ban on scraping to prevent interference with website operation, even analyzed under the lower standard of time-place-manner scrutiny, such a law could not be easily tailored to achieve its objective without burdening substantial amounts of innocuous data collection. Websites are not all the same; some can handle substantially greater amounts of automated traffic without experiencing impairment of their function. It would be difficult to select a one-size-fits-all threshold for file size—or some other limiting characteristic—that does not either impose serious limitations on harmless activity by being set too low, or fail to achieve its purpose by being set so high that damaging effects on less robust sites are not prevented. It may be that there is some middle ground, but it is not clear that this is the case.

Alternatively, if a statute intended to protect against website interference is viewed through the lens of strict scrutiny because it fails the ample alternative channels test, there are plainly less restrictive alternatives to achieve this goal. A statute could explicitly prohibit knowing or willful interference with site operation, and apply only in circumstances where a scraper is notified that they are causing harm (and in fact does cause such harm). This narrower option would not, of course, address the privacy concerns discussed above, making it useless to a legislature that is seeking to engage in content-based regulation under the guise of a content-neutral justification. In fact, the failure to adopt a narrow approach to achieve a content-neutral goal could reveal that goal as a pretext for content-based concerns.

A different approach would be to defer to the choices made by website operators as to whether to allow scraping. However, this option is not without its own issues because website operators' motives for prohibiting scraping would not necessarily be limited to protecting site operation (or for that matter any other legislative motive). This problem was central to the *Sandvig* case, and allowed the plaintiff researchers to state a valid as-applied First Amendment challenge to the application of the Computer Fraud and Abuse Act. *See Sandvig I*, 315 F.Supp.3d at 30 (holding that plaintiffs stated as-applied claim based on allegation that their scraping in violation of terms of service would not cause harms that CFAA was intended to prevent). Unfortunately—at least from the perspective of resolving the First Amendment question—the researchers' argument was later determined to be moot when the district court subsequently ruled that the CFAA does not bare terms-of-service violations. *Sandvig v. Barr*, No. 16-1368 (D.D.C. Mar. 27, 2020) ("*Sandvig II*"). *See also hiQ Labs, Inc. v. LinkedIn Corp.*, 938 F.3d 985, 999-1004 (9th Cir. 2019) (CFAA does not apply to scraping of public-facing information from website), *petition for writ of certiorari filed sub nom. LinkedIn Corp. v. hiQ Labs, Inc.*, No. 19-1116 (U.S. Mar. 9, 2020).

Can a Ban on Scraping Photos be Saved by Limiting it to Scraping for Facial Recognition?

This article was inspired by questions arising out of Clearview AI's scraping of photos for its facial recognition system, and while I have discussed alternative justifications for a ban on scraping, the obvious focus will be on privacy issues. As discussed above, a privacy-based ban is content-based and will therefore be subject to strict scrutiny, and may face questions both as to whether it serves a compelling government interest and as to whether it is the least restrictive alternative.

To respond to the latter concern, a ban on scraping could be limited to prohibiting scraping of photos for use in facial recognition systems. Such a limitation would, of course, confirm that the law is content-based by focusing on privacy issues, but we are already presuming that strict scrutiny applies. Moreover, while the limitation itself would be content-based (inasmuch as it is concerned with the derivation of information from the gathered photos), the distinction focuses on the same concerns that motivate the overall concept of a ban and thus does not raise separate constitutional issues regarding content discrimination. *See R.A.V. v. St. Paul*, 505 U.S. 377, 388 (1992) (content-based limitation on scope of statute proscribing speech is permissible where "the basis for the content discrimination consists entirely of the very reason the entire class of speech at issue is proscribable."). Put simply, if privacy justifications are considered a compelling state interest for the concept of a scraping ban, there would be no additional problem in limiting that ban to scraping for purposes that are deemed especially harmful to that interest.

However, there remains the question of whether a ban limited in this fashion would be the least restrictive alternative. Image searches such as those available on major search engines have not raised major privacy concerns, despite the fact that they can be used to search for photos of individuals and depend on their own forms of image-matching algorithms. Clearview AI bills itself as a search engine, and it is not immediately apparent how a distinction could be drawn.

But even assuming that some technical or other difference could be identified, the fact that the limitation would be based on conduct subsequent to the scraping itself demonstrates that there is a less-restrictive alternative. As discussed in the prior section, a legislature could ban the processing of photos for biometric data, as in Illinois' biometric privacy law, instead of banning scraping. If there is no need to ban scraping to achieve the legislature's objective, then the ban is by definition not the least restrictive alternative.

Conclusion

While much of the outrage around the activities of companies such as Clearview AI has focused on the amount of photographic data they are collecting and the scraping technology used to accomplish that, the underlying concern is not so much with the mode or scale of data gathering as with the purposes for which that information is gathered. Running a First Amendment analysis on a ban on scraping helps to illustrate this by showing that such a ban could implicate speech interests and burden other kinds of innocuous data gathering. Moreover, the analysis helps to surface questions about the specific nature of the privacy interests that are implicated by facial recognition technology, interests that are muddled by focusing on the technology used to obtain data rather than what is done with the data post-collection. And through the tailoring analysis, it

becomes clear that such legitimate privacy interests that do exist are better addressed directly rather than using a prohibition on scraping to cut off access to information in the public sphere.