Cecily Gao

Professor Joan Feigenbaum

CPSC 610 – Topics in Computer Science and Law

Spring 2021

Section 230: Past, Present, and Future

Introduction

"The Twenty-Six Words That Created the Internet" is no hyperbole when it comes to Section 230 of the Communications Decency Act (CDA) of 1996.[1] By every measure, § 230 is objectively the single most influential piece of Internet legislation ever passed. CDA § 230(c)(1) grants online service providers immunity from publisher liability with the following language:

> 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.[2]

CDA § 230(c)(2)(A) also grants safe harbor from civil liability for "Good Samaritan" moderation of third-party and user-generated content:

> No provider or user of an interactive computer service shall be held liable on account of—
> (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.[3]

In recent years, as technology giants, like Google and Facebook, have solidified their foothold in nearly every aspect of American online life and the consequences—both positive and negative—

[1] Kosseff, Jeff. *The Twenty-Six Words That Created the Internet*. Ithaca; London: Cornell University Press, 2019. Accessed May 10, 2021. http://www.jstor.org/stable/10.7591/j.ctvr7fcrd.
[2] Communications Decency Act of 1996, Pub. L. No. 104-104, 110 Stat. 137.
[3] Ibid.

of unregulated, instantaneous communication platforms have crystallized, the calls to revoke broad-stroke immunities granted to online service providers by § 230 and expanded by the courts have only gained momentum and bipartisan support. In this paper, we will survey the legislative and legal history of § 230 as well as the current reform bills, and I will make a case that none of the extant proposals are appropriate—technically, legally, or practically—to sufficiently and properly address the issues created by § 230. Instead, § 230 reform requires a more nuanced approach that utilizes the "Protection for 'Good Samaritan' blocking and screening of offensive material" provision of the original text to scale back the fortress of unequivocal protection that the courts have thus far granted to online service providers.

Historical Context of § 230

In late April of 1993, CERN made the World Wide Web freely available to the public.[4] Previously only accessible to select institutions and companies, the Internet experienced a boom in traffic and innovation in the short years following its public debut, as individuals and companies around the world realized the vast potential of the Web for instantaneous access to information and resources. With the concurrent advent of compression methods around this time (e.g., JPEG, MPEG, PDF), online service providers were able to expand their offerings from standard, text-only webpages to include multimedia content. The "Big Three" online information services of the time were America Online (AOL), Prodigy, and CompuServe, which each provided a selection of services that included news, weather, market quotes, games, and email.[5] As is readily apparent, the majority of online engagement during this time was of consumptive

[4] "A Short History of the Web," CERN, accessed May 10, 2021, https://home.cern/science/computing/birth-web/short-history-web.
[5] Lewis, Peter H. "PERSONAL COMPUTERS; The CompuServe Edge: Delicate Data Balance." *The New York Times*, November 29, 1994.

behavior, i.e., the vast majority of Internet users were not generating online content themselves but rather consuming the content made and provided by others. However, the Internet quickly shifts towards a more participatory nature, centered around user-generated content began in the form of online discussion boards—precursors to the social media platforms of today—hosted by service providers, and Congress realizes that the Internet had the potential to democratize speech at an unprecedented scale like no medium before.

Two major developments during this era of the Internet caught the full attention of Congress and prompted forceful legislative efforts to regulate the budding Internet: (1) the astronomical rise in popularity of online pornography, and (2) the divergent court decisions in the libel cases of *Cubby, Inc. v. CompuServe Inc.* (1991) and *Stratton Oakmont, Inc. v. Prodigy Services Co.* (1995).[6] The product of these two concerns were two amendments to the Telecommunications Act of 1996—the CDA and § 230.

The introduction of multimedia formats to the Internet was inevitably accompanied by pornographic materials and other obscene content. Bolstered by Christian activists, then-Senator James Exon (D-NE) believed that the Internet made obscene and pornographic materials too easily accessible to minors and proposed the CDA, which would criminalize making available to minors any material considered "patently offensive as measured by contemporary community standards," including material that described or depicted "sexual or excretory activities or organs."[7] During deliberations on the Senate floor, Sen. Exon infamously brought a blue binder full of extremely pornographic content collected from the Internet and encouraged his colleagues to view the obscene material for themselves. As one can probably imagine, the blue binder,

---

[6] *Cubby, Inc. v. CompuServe Inc.,* 776 F. Supp. 135 (S.D.N.Y. 1991); *Stratton Oakmont, Inc. v. Prodigy Services Co.,* 23 Media L. Rep. 1794 (N.Y. Sup. Ct. 1995).
[7] Telecommunications Act of 1996, Pub. L. No. 104-104 (1966), https://www.congress.gov/bill/104th-congress/senate-bill/652

accompanied by Sen. Exon's impassioned speech, was sufficient motivation for the majority of his colleagues to vote for the bill.

For those familiar with the First Amendment of the United States Constitution, the gut reaction to the CDA should be that it is unconstitutional on its face, which is the same conclusion the Supreme Court came just over a year after its passage in *Reno v. ACLU* (1997).[8] There is no legal precedent for suppressing speech for the entire populace to protect minors. Two of Exon's detractors in the House of Representatives—Reps. Christopher Cox (R-CA) and Ron Wyden (D-OR)—were concerned about what the CDA would mean for the trajectory of the Internet on top of the concerning precedents set in *Cubby* and *Stratton Oakmont*.

In 1995, the case of *Stratton Oakmont v. Prodigy* set an important precedent for how online service providers can be held liable for speech in a court of law. For the facts of the case: in the 1990s, Prodigy offered "bulletin boards" that hosted user-generated content. To keep their websites family-oriented, Prodigy screened submitted content and reserved the right to take down or refuse to post submissions based on offensiveness or "bad taste." In 1994, an anonymous user posted accusations on Prodigy's *Money Talks* board that Stratton Oakmont, a financial firm, was guilty of fraud. In response, Stratton Oakmont sued Prodigy (and the anonymous user) for defamation. Their legal team argued that by moderating submissions, Prodigy stepped out of its role as a distributor of third-party content and instead filled the role of an active publisher of the material posted on the platform. For reference, there are three classes of tort liability for defamatory speech in defamation law—publisher liability, distributor liability, and common carrier liability. A publisher is defined as:

> …an entity that exercises some degree of editorial control over the dissemination of the defamatory material will be generally liable for its publication…A newspaper, for

[8] "Reno v. ACLU." Oyez. Accessed May 10, 2021. https://www.oyez.org/cases/1996/96-511.

> example, may be liable for defamation if a letter to the editor that it publishes contains false and defamatory statements.[9]

A distributor is defined as:

> …an entity that distributes but does not exercise editorial control over defamatory material may only be liable if such entity knew or had reason to know of the defamation…News vendors, bookstores, and libraries generally qualify for this standard of liability.[10]

And a common carrier is defined as:

> …an entity that merely acts as a passive conduit for the transmission of defamatory material, such as a telephone company, is not subject to defamation liability, even if such entity knew or had reason to know of the defamation. Furthermore, in the event that the conduit service could be characterized as a publisher, it is entitled to a qualified immunity from liability subject to the common law exception for malice.[11]

Guided by such presiding legal definitions, the New York Supreme Court sided with Stratton Oakmont and determined that Prodigy's enforcement of content guidelines, employment of "Board Leaders" (moderators), and usage of automatic screening algorithms categorized it as a publisher rather than a distributor. The court's interpretation of precedent was consistent with the prior court decision of *Cubby v. CompuServe*, in which Cubby accused CompuServe of hosting defamatory content produced by a third party on their sites. What differed in the *Cubby* case is the fact that CompuServe performed no moderation on their platforms, and thus, by demonstrating they had no knowledge of the defamatory content posted on their platform, they were shielded by distributor liability in court. These two decisions in *Cubby* sent a very strong message to online service providers: if you perform any moderation of third-party content on your service, you will be held liable as a publisher. Conversely, if you refuse to perform any moderation of third-party content, you can enjoy distributor immunity.

---

[9] Friedman, Jonathan A., and Francis M. Buono. "Limiting tort liability for online third-party content under section 230 of the communications act." Fed. Comm. LJ 52 (1999): 650.
[10] Ibid.
[11] Ibid.

Reps. Cox and Wyden aptly realized that the Internet was shaping up to be unlike any other communication medium to date, and while the court decisions made in *Cubby* and *Stratton Oakmont* followed prior precedent, the traditional, real-life analogs used in both cases were not sufficient to prepare for a participatory Internet. Unlike the more established communication mediums, like public broadcasting or newspapers, the Internet posed two unique challenges: (1) the number of speakers on the Internet could theoretically equal the number of users, and (2) content can instantaneously be posted onto the Internet. If content moderation were unequivocally de-incentivized, several members of Congress feared that the Internet would not be able to grow to its fullest potential as a powerful resource for the dissemination of information and a tool for democracy.

Between 1991 and 1995, the number of websites and worldwide users on the Internet grew from 1 to 257,601 and 2.6 million to 44 million, respectively.[12] As a matter of technical feasibility, Reps. Cox and Wyden and the more tech-savvy members of Congress realized that the enforcement of the CDA's anti-indecency provisions was neither practical nor possible. The legislative rationale then followed that if online service providers were punished for voluntary moderation—to remove the very things that the CDA wanted to target—simply because they could not catch everything, there would be two inevitable consequences: (1) the complete de-incentivization of performing any type of moderation, which would then (2) either completely de-incentivize the creation of participatory platforms or lead to the proliferation of more obscene, pornographic, and indecent content on participatory platforms due to lack of moderation.

As the CDA made its way through Congress, two schools of thought began to rise to popularity regarding how—or if—the Internet should be regulated. The first is internet

[12] Max Roser, Hannah Ritchie, and Esteban Ortiz-Ospina, "Internet," *Our World in Data*, July 14, 2015, https://ourworldindata.org/internet.

exceptionalism, which argues that the Internet is unique and has no traditional, real-life analogs, and thus cannot be regulated as with real-life counterparts. The second school of thought—which some categorize as a sub-school of the former—is cyber-libertarianism, which purports that the Internet transcends physical borders and thus should be free of all governmental regulation and censorship. Instead, the web should be allowed to function as a free market, with the popularity of websites determined by the quality or demand of the services offered.[13]

Knowingly or not, Reps. Cox and Wyden managed to synthesize both schools of thought into the passing of § 230. In their effort to save the developing Internet and maintain its trajectory, Reps. Cox and Wyden turned the onus of moderation to the online service providers themselves in hopes that companies would moderate freely—even the content protected under the First Amendment, like obscenity—without fear of publisher liability. As they called it, they granted service providers a "sword and shield."[14] And thus delivers us the takeaway of this section—understanding that the *intent* of the legislators behind § 230 was to empower online service providers the power and legal immunity to do what the government could not in the Internet sphere and push ownership and consequences of online speech onto the original speakers themselves.

The Modern Internet and § 230

Before we discuss the legal challenges to § 230, let us take a brief interlude to acknowledge the significant impact the law has had on the multi-trillion-dollar Internet industry

---

[13] H. Brian Holland, "In Defense on Online Intermediary Immunity: Facilitating Communities of Modified Exceptionalism," *University of Kansas Law Review* 56, no. 2 (January 2008): 369-404

[14] "Opinion: I Wrote This Law to Protect Free Speech. Now Trump Wants to Revoke It," CNN (Cable News Network, June 9, 2020), https://www.cnn.com/2020/06/09/perspectives/ron-wyden-section-230/index.html; Christopher Cox, "Policing the Internet: A Bad Idea in 1996 -- and Today," *RealClearPolitics*, June 25, 2020, https://www.realclearpolitics.com/articles/2020/06/25/policing_the_internet_a_bad_idea_in_1996_--_and_today.html.

as we know it today. After the passing of § 230 in 1996, the shift towards an online sphere in which users generate the content accelerated. Without legal disincentives to discourage platforms from being created, social media sites like Sixdegrees, Friendster, LinkedIn, and MySpace rose at the turn of the millennia, followed by—and replaced, in some cases—the likes of Facebook, Twitter, and Tumblr. This era of the participatory Internet has been termed Web 2.0.[15]

In recent years, the rapid spread of information on these social media platforms has proved to work in both social positives and negatives. A notable example is the murder of George Floyd in May of 2020, which was filmed by a bystander and uploaded to Facebook, where it instantly went viral and spread onto multiple social media platforms, gaining national and international attention, sparking the far-reaching Black Lives Matter protests during the summer of 2020, and ultimately impacting substantial policy changes from the local to national levels.[16] Conversely, during the same summer, unfounded conspiracies regarding COVID-19 and then-presidential candidate, Joe Biden, and his son, Hunter Biden, flourished on the same platforms, resulting in substantial consequences to the COVID-19 response and the general election.[17] I will pose a question here for the reader to ponder—can we bear witness to one social phenomenon and not the other?

While the protections offered by § 230 to these social media platforms may be obvious, something that often escapes notice is the protections offered by § 230 extend to *all* interactive service providers, including websites like Amazon, eBay, Etsy, Airbnb, and Grindr. For e-commerce sites, § 230 shields them from being sued due to hosted sellers misrepresenting their

---

[15] Christopher McFadden, "A Chronological History of Social Media," *Interesting Engineering*, July 2, 2020, https://interestingengineering.com/a-chronological-history-of-social-media.
[16] Bogel-Burroughs, Nicholas and Fazio, Marie. "Darnella Frazier captured George Floyd's death on her cellphone. The teenager's video shaped the Chauvin trial." *The New York Times*, April 20, 2021. https://www.nytimes.com/2021/04/20/us/darnella-frazier-video.html.
[17] Amarasingam, Amarnath, and Marc-André Argentino. "The QAnon conspiracy theory: A security threat in the making." *CTC Sentinel* 13, no. 7 (2020): 37-44.

products. For listing sites like Airbnb, § 230 protects them from local laws that punish unlicensed short-term rentals, as the listings qualify as user-generated content. For dating sites, like Grindr, user profiles are similarly user-generated content, and this can even be used to protect against separate product liability suits (as we will see in *Herrick v. Grindr* in the next section).

Just by looking at the wide range of sites and services made available in the twenty-five years since the passing of § 230, we can irrefutably assert that § 230 has allowed Internet companies to innovate and to take risks in their offerings and business models. Although § 230 does not mandate content moderation, many of the aforementioned online service providers have an intrinsic incentive to moderate content. For example, to maximize its user base, Facebook wants to avoid hosting content that most users would find distasteful or offensive (e.g., pornography or gore). If a comparable social media site provides a similar service but can filter out more "bad" content, Facebook is at risk of losing market share. Thus, they have found their motivation to moderate. Similarly, Amazon also has a motivation to regulate its third-party sellers. If consumers increasingly receive defective or inaccurate products due to an influx of low-quality sellers, Amazon is at risk of losing customer loyalty.[18] Along a similar vein, if a single Airbnb customer has a nightmarish experience due to a lack of screening hosts, Airbnb can suffer major reputational damages, which can drive users to other listing services.[19] All of these examples of intrinsic motivators for platform moderation are arguably better than any single piece of legislation, as they are driven by the market demands, which the companies

[18] Sarah Perez, "To Fight Fraud, Amazon Now Screens Third-Party Sellers through Video Calls," TechCrunch (TechCrunch, April 27, 2020), https://techcrunch.com/2020/04/27/to-fight-fraud-amazon-now-screens-third-party-sellers-through-video-calls.
[19] "Airbnb Host Admits Manslaughter after Killing Guest over Unpaid Bill," BBC News (BBC, March 4, 2019), https://www.bbc.com/news/world-australia-47446073.

themselves are most familiar with. On the surface, those of the cyberlibertarian-internet-exceptionalism school of thought seem to have gotten it right—the Internet seems to be doing alright regulating itself.

Legal Challenges to § 230

Immediately after the passing of § 230, both the legal and Internet communities were interested in knowing exactly how far the protections of the new law stretched. The first test to the limits of § 230 happened just a year later in *Zeran v. America Online* (1997), which was another defamation case against an Internet company.[20] The plaintiff, Zeran, was repeatedly harassed due to defamatory ads placed on AOL and he alleged that the company took no actions even after being made aware of the harassment. With § 230 as their shield, AOL did not attempt to claim that they were unaware of the materials posted on their discussion boards, as CompuServe had done. Despite the plaintiff arguing that distributor liability (i.e., notice-based takedown) applied in this case and was not protected under § 230, the 4th Circuit ruled in favor of AOL. In the court opinion, the bench went a step further than the text of § 230 and stated that platforms like AOL were also exempt from distributor liability, in addition to publisher liability.[21] More specifically, even if a provider were made aware of harassing behavior on its platform, it would not need to remove the content. The concern of the 4th Circuit was that holding providers liable as distributors could lead to immediate takedowns or equivalent consequences for any user-generated content that was found "objectionable" by other users, and thus would effectively chill speech.

[20] *Zeran v. America Online, Inc.,* 129 F.3d 327 (4th Cir. 1997).
[21] Jonathan A. Friedman; Francis M. Buono, "Limiting Tort Liability for Online Third-Party Content under Section 230 of the Communications Act," *Federal Communications Law Journal* 52, no. 3 (May 2000): 647-666

This is undeniably an example of bench legislation, as the original text of § 230 only specifies protections for online service providers from publisher liability and not distributor liability. The predicted abuse of a hypothetical reporting feature of a service in which the service provider could be held liable as a distributor was complete conjecture. However, I will concede that this finding has some underlying logical basis. If providers were subject to distributor liability, they might lack the incentive to moderate or monitor their platforms. By definition of the distributor liability, the notice-based takedown rule could encourage an "ignorance is bliss" approach to content moderation because as long as the provider is unaware of the bad behavior occurring on their platform, they cannot be held liable for any consequences. In other words, when faced with the high likelihood that objectionable or consequential material may fall through the cracks of moderation, providers would rather air on the side of caution and may refuse to moderate at all, thus creating a paradox on the original intent of § 230 and placing us back at square one. That said, despite the intention of Congress and the theoretical rationality behind the court's decision in *Zeran*, the decision is still an example of overstepping the interpretation of the law by the courts.

The court opinion in *Zeran* set the precedent for nearly all subsequent cases involving online service providers. A notable, more recent exception to ubiquitous § 230 protection came in the form of *Fair Housing Council of San Fernando Valley v. Roommates.com* (2008).[22] The Fair Housing Council sued Roommates.com for asking users demographic questions (i.e., questions related to gender, sexual orientation, age, etc.) that were protected under the Fair Housing Act. Roommates.com argued that filling out the form resulted in user-generated content and was thus protected under § 230. However, in a first, the 9th Circuit ruled against the

[22] *Fair Housing Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157 (9th Cir. 2008).

defendant and determined that the questionnaire did not fall under § 230 protections, as Roommates.com was an active participant in the creation of the discriminatory demographic questions, which legally made them the "speaker" of the questionnaire.

As a result of the precedents set in *Zeran* and *Fair Housing Council*, courts have developed a three-pronged test to determine whether the § 230 liability shield applies to a defendant:

1. The defendant must be a provider of a user of an interactive computer service, as defined in § 230.
2. The plaintiff must argue that the defendant is the publisher or speaker of harmful content.
3. For § 230 to apply, the harmful content made available by the provider must have been generated and provided by another party other than the defendant.

Fair Housing Council was one of the very few cases in which the courts have drawn a bright-line rule regarding the limits of § 230. More often than not, courts have sided with defendants and further expanded the reach of § 230 protections. Some cases of particular concern involve real-life harassment or assault that have resulted from the failure of websites to take any sufficient action to adequately warn or protect users (e.g., *Doe v. MySpace* (2008) and *Jane Doe No. 14 v. Internet Brands* (2015)).[23]

Two notable cases have emerged from Internet-related tort cases in which § 230 has consistently superseded gray areas: *Fields v. Twitter* (2017) and *Herrick v. Grindr* (2018).[24] In *Fields,* the plaintiff's husband was shot and killed in an ISIS attack while working in Jordan. The plaintiff sued Twitter for providing a platform on which ISIS could spread its ideology and recruit new members. Under the Supreme Court ruling in *Holder v. Humanitarian Law Project*, the act of providing any type of material support to terrorists can be made illegal by the federal

[23] *Doe v. Myspace*, 528 F.3d 413 (5th Cir. 2008); *Doe No. 14 v. Internet Brands, Inc*., 778 F.3d 1095 (9th Cir. 2015).
[24] *Fields v. Twitter, Inc.,* 881 F.3d 739 (9th Cir. 2018); *Herrick v. Grindr LLC*, 18-396 (2d Cir. Mar. 27, 2019).

government, even political speech that is protected under the First Amendment.[25] Despite the plaintiff's artful arguments to circumvent § 230, the 9th Circuit ruled in favor of Twitter on the basis that there was not enough causal evidence to prove that Twitter allowing members of ISIS to use their platform directly contributed to the death of the plaintiff's husband.

Herrick is one of the more recent § 230 cases that has received significant coverage in the media. In this suit, the plaintiff's ex-boyfriend made several fake profiles on multiple dating apps, including Grindr, to impersonate the plaintiff and set up meetings with strangers for sex. This led to over a thousand men approaching the plaintiff via the geolocating feature on Grindr.[26] After multiple failed attempts to get Grindr to ban the fake profiles, Herrick sued. As in *Fields*, Herrick's legal team tried to circumvent § 230 by making a separate argument that, by not having a feature to ban users, Grindr was a defective product and was subject to product liability rules. Despite this legal argument being completely independent of § 230, all the court decisions fell back on § 230 immunity and ruled in favor of Grindr. Though the case was denied a writ of certiorari at the Supreme Court, Justice Clarence Thomas cited *Herrick* in a separate brief in which he expressed concern that § 230 is being applied too broadly and used ubiquitously to shield technology companies from all legal repercussions.[27]

Criticisms and Failures of § 230

Public criticism of § 230 has risen to new levels in recent years, and the present debate surrounding § 230 reform is incredibly complex and involves political, economic, socio-cultural,

[25] *Holder v. Humanitarian Law Project*, 561 U.S. 1, 130 S. Ct. 2705, 177 L. Ed. 2d 355 (2010).
[26] Carrie Goldberg, "Winning Through Losing," American Bar Association, December 10, 2020, https://www.americanbar.org/groups/diversity/women/publications/perspectives/2021/december/winning-through-losing/.
[27] *Malwarebytes, Inc. v. Enigma Software Grp*. U.S., No. 19-1284 (U.S. Oct. 13, 2020).

and technological factors. For the sake of simplicity, let us contextualize the criticisms along a linear axis. At one end of this axis (let's call this Terminal A) is the belief that technology companies are moderating too much and are subsequently restricting free speech. At the other end (let's call this Terminal B) is the view that technology companies are not moderating enough to protect their users from harm, repress disinformation, and squash bad actors. The crux of both beliefs is that the immunity conferred by § 230 gives online service providers too much power over what content is accessible to the public. Somewhere in the middle of this spectrum lies the belief that § 230 is perfect just the way it is, and somewhere else on this spectrum lies the belief the § 230 should be modified to require platforms to perform some minimum level of content moderation to filter out the worst of the worst, like Child Sexual Abuse Imagery (CSAI).

Terminal A is generally populated by conservative-leaning individuals, who believe that social media companies are explicitly censoring "conservative" viewpoints. The underlying belief of this argument is that § 230 was not intended to give platforms the power to censor political opinions. From a legal standpoint, this is an illegitimate argument, as the theoretical ability to moderate or selectively "publish" content is granted to private entities by the First Amendment, not § 230. In the lead-up to the 2020 presidential election, Twitter explicitly labeled false tweets by prominent politicians, including then-President Donald Trump. This prompted an executive order from Trump to prompt the FCC to punish technology companies that engaged in "censorship."[28] This further initiated calls from conservative figures that Twitter, Facebook, and YouTube were acting as publishers of material with the power to sway public opinion, and should be treated as such, without the liability shield provided by § 230.

[28] Exec. Order. No. 13925, 72 Fed. Reg. 34079 (June 2, 2020).

Terminal B is generally populated by those who hold liberal beliefs. An important temporal turning point for many in this group occurred during the 2016 presidential election cycle, during which foreign bad actors (primarily originating from Russia) successfully held a powerful online disinformation campaign to disadvantage the Democratic presidential candidate, Hillary Clinton.[29] A common criticism—as vocalized by Senator Cox, one of the original proposers for § 230—is that the online service providers are not doing enough to keep the Internet clean of "slime," i.e., hate speech, extremism, election interference, falsehoods.[30] Referencing his original "sword and shield" analogy, Sen. Cox recently warned platform providers that if "you don't use the sword, there are going to be people coming for your shield."[31]

Outside of the chasmic political divide, several concerns have garnered some bipartisan support. Danielle Keats Citron, a law professor at Brown University, developed a Good Samaritan/Bad Samaritan dichotomy and argues that § 230 was intended to protect service providers that made a good-faith effort to regulate content on its platforms.[32] Many believe that granting Wikipedia and 8kun (the platform on which several mass shooters have posted their manifestos) the same legal shields does not seem to be a fair or balanced approach.

The recent rise of revenge porn has also surfaced another criticism that § 230 supersedes any state regulations involving Internet platforms. The case of revenge porn is particularly interesting, as 48/50 states currently have some form of law that punishes revenge porn.[33]

---

[29] Leary, Mary Graw. "The indecency and injustice of section 230 of the Communications Decency Act." *Harv. JL & Pub.* Pol'y 41 (2018): 553.
[30] Daisuke Wakabayashi, "Legal Shield for Social Media Is Targeted by Lawmakers," *The New York Times*, May 28, 2020, https://www.nytimes.com/2020/05/28/business/section-230-internet-speech.html.
[31] Ibid.
[32] Citron, Danielle Keats, and Benjamin Wittes. "The internet will not break: Denying bad samaritans sec. 230 immunity." Fordham L. Rev. 86 (2017): 401.
[33] Zak Franklin et al., "Justice for Revenge Porn Victims: Legal Theories to Overcome Claims of Civil Immunity by Operators of Revenge Porn Websites," California Law Review, October 21, 2018,

However, due to the lack of federal regulation, none of the state courts have overruled § 230 to enforce anti-revenge porn regulations on platforms where such content aggregates.

Policy Proposals and Their Technical Feasibility

Before walking through the § 230 reform bills currently making their way through Congress, we should examine two amendments that have passed in 2018—Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA) and Stop Enabling Sex Traffickers Act (SESTA). FOSTA-SESTA defines an amendment to § 230 in which the legal liability shield cannot be "construed to impair or limit any claim in a civil action brought under" violation of a state or federal sexual assault civil law. These amendments also added the following clarification on the perceived original intent of the law:

> Section 230…was never intended to provide legal protection to websites that unlawfully promote and facilitate prostitution and websites that facilitate traffickers in advertising the sale of unlawful sex acts with sex trafficking victims…[34]

Like the original anti-pornography and anti-obscenity parts of the CDA, the intentions behind these bills were most likely of "moral" intent. The primary target of the amendments was Backpage.com, a classified advertising site and the largest online source for illegal human trafficking.[35] Despite their blatant disregard for laws prohibiting human trafficking, Backpage.com won several court cases thanks to § 230 protections. Right after FOSTA-SESTA was signed into law, the federal law enforcement agencies were empowered to raid and shut down Backpage.com. In theory, FOSTA-SESTA would encourage such platforms to more

---

https://www.californialawreview.org/print/4justice-for-revenge-porn-victims-legal-theories-to-overcome-claims-of-civil-immunity-by-operators-of-revenge-porn-websites/.

[34] Allow States and Victims to Fight Online Sex Trafficking Act of 2017, Pub. L. No. 115-164, 132 Stat. 1253.

[35] Eric Goldman, "The Complicated Story of Fosta and Section 230," *First Amendment Law Review* 17, no. Symposium (2018): 279-293

heavily restrict and control potential sex trafficking that occurred on these platforms and aid law enforcement in tracking down perpetrators.

However, the actual consequences of these acts were exactly as predicted by Sens. Wyden and Cox back in 1996. Companies now had three options if they wanted to avoid potential liability: (1) bolster their filtering algorithms to try to catch all prostitution or sex trafficking-related content, (2) turn off moderation all together and adopt the "ignorance is bliss" strategy, or the nuclear option, (3) shut down completely. Immediately, several sites—including Craigslist's Personals Ads page and Pounced.org (a Furry dating site)—chose the third option and shut down in light of these new liability exceptions.[36] Ironically, such sites shutting down makes it even harder for law enforcement to track down criminals who violate sex trafficking laws, as their content can no longer be found on the surface web. Furthermore, sex work advocates have estimated that losing safe, online tools to screen customers has subsequently resulted in more deaths and sexual assaults of sex workers.[37] Companies that could afford the first option are generally the large incumbents (e.g., Google, Facebook) who can dedicate resources to improving their algorithms, thereby further increasing the barrier to entry for new tech companies. These consequences harken back to a game theory tradeoff, in which legislative optimism always loses to the desire for companies to avoid legal liability.

The failures of FOSTA-SESTA have not deterred Congress from introducing more legislation to reform § 230. With sufficient background on § 230, let us now peruse and discuss the existing amendments currently moving through Congress. At the present, the following are some of the proposals that have been introduced in the past year, in an arbitrary order:

---

[36] Lura Chamberlain, "FOSTA: A Hostile Law with a Human Cost," Fordham Law Review 87, no. 5 (April 2019): 2171-2212

[37] Ibid.

1. Ending Support for Internet Censorship Act, proposed by Sen. Josh Hawley (R-MO)[38]
2. Eliminating Abusive and Rampant Neglect of Interactive Technologies (EARN IT) Act, proposed by Sens. Blumenthal (D-CT) and Graham (R-SC)[39]
3. Safeguarding Against Fraud, Exploitation, Threats, Extremism and Consumer Harms (SAFE TECH) Act, proposed by Sens. Warner (D-VA), Hirono (D-HI), and Klobuchar (D-MN)[40]
4. Platform Accountability and Consumer Transparency (PACT) Act, proposed by Sens. Schatz (D-HI) and Thune (R-SD)[41]

Sen. Hawley's proposal is by far the most politically charged of the group (as indicated by his sole sponsorship). Hawley's motivation for § 230 reform lies on the extreme end of Terminal A. In the Ending Support for Internet Censorship Act, Hawley calls for large companies—as defined by some arbitrary combination of users and revenue—to ensure complete "political neutrality" across every possible "political viewpoint" in their moderation efforts to enjoy the § 230 shield. Putting aside the impossibility of a platform given an equal voice to every single political viewpoint (Neo-Nazis? Flat earthers? Anti-Vaxxers?), perhaps more egregiously, Hawley implores the FCC, an arm of the federal government, to determine whether a platform is "politically neutral," which is not only, again, arbitrary but also an incredible violation of the First Amendment. Just as it is a form of speech for a newspaper to refuse to publish an article, the FCC deciding which moderation system passes some subjective bar of political neutrality is also a form of speech—specifically government speech—that exerts control over private speech. The likelihood of Hawley's bill getting anything more than some brief press is absurd. Ending Support for Internet Censorship Act has no legislative or practical effect and is very transparently for political clout.

[38] Ending Support for Internet Censorship Act., S. 1914, 116th Congress. (2020).
[39] EARN IT Act of 2020., H.R. 8454, 116th Congress. (2020).
[40] SAFE TECH Act, S. 299, 117th Congress. (2021).
[41] PACT Act, S. 797, 117th Congress. (2021).

The EARN IT Act has thus far received the most coverage in the media, partly for its bipartisan support and intrusion on encryption. The main feature of the act is the formation of the National Commission on Online Child Sexual Exploitation Prevention, a 19-member panel that periodically determines the "best practices" technology companies must abide by to receive the § 230 liability shield. The members of the panel pull from the executive branch (the Attorney General, the Department of Homeland Security, and the Chairman of the Federal Communications Commission) and Congress-appointed individuals. As with the original anti-indecency provision of the CDA and the FOSTA-SESTA, we can choose to believe that the intentions behind the bill were moral and designed to prevent child sexual exploitation online. However, this does not mean we can turn a blind eye to the practical consequences of such legislation. First, given the recent political polarization of Washington, the panel is at a real risk of overrepresenting the interests of the executive branch or a single political party, putting the online lives of Americans under the control of unelected officials. In that possibility, the "best practices" are much more susceptible to the bend to strong governmental political interests, which may very likely run afoul of First Amendment protections against government control over speech. In addition, many security experts believe that this legislation is an attempt to install backdoors to encryption systems. Some branches of the executive branch have historically been hostile to encryption efforts, including the former Attorney General William Barr, who penned a letter to Facebook back in 2019 to deter the company from implementing end-to-end encryption on its messaging platforms.[42] The representation of such attitudes on the panel could lead to the weakened security of Americans and have the adverse effect of making American online services

[42] "Attorney General Barr Signs Letter to Facebook From US, UK, and Australian Leaders Regarding Use of End-To-End Encryption," The United States Department of Justice, October 4, 2019, https://www.justice.gov/opa/pr/attorney-general-barr-signs-letter-facebook-us-uk-and-australian-leaders-regarding-use-end.

less attractive in foreign markets, where trust in the U.S. government may be at an all-time low.[43] This bill essentially asks us to make a decision in the false dichotomy between our online privacy and online child safety. Given that the zeitgeist presently appears to point in the direction of more consumer privacy, asking Americans to unilaterally give up their privacy while introducing nebulous, to-be-determined "best practices," seems like it would have a chilling effect on speech and carry with it more harms than benefits.[44]

The SAFE TECH Act proposes some major revisions to the exceptions of § 230, building onto the changes made by FOSTA-SESTA. The sponsors of the act have all expressed concerns over how much the courts have sided with online service providers against victims in cases like *Herrick* and *Jane Doe No. 14*. The bill seeks to remove § 230 protections for (1) paid content (e.g., ads), (2) injunctive relief in which a user believes the provider's service may cause immediate and irreparable harm, (3) violations of civil rights law, (4) harassment or stalking, (5) wrongful death claims, and (6) lawsuits under the Alien Tort Claims Act, which would allow foreign users to sue American-based platforms for platform-enabled human rights violations. The amendment of these exceptions will inevitably increase the number of tort suits in court and allows victims of online activity to shift their legal blames to the platforms themselves. This has an immediate concern of creating an undue on smaller companies in the platform space who lack resources to grow and improve their product while also fending off individual lawsuits in court.[45] This is an issue of scale. Though Sen. Warner, the main co-sponsor of the act, believes that the § 230 shield has given large tech companies a "get-out-of-jail" free card that has allowed them to

---

[43] Lewis, James A., Denise E. Zheng, and William A. Carter. *The effect of encryption on lawful access to communications and data.* Rowman & Littlefield, 2017.
[44] Dwyer, Timonthy. "Evolving concepts of personal privacy: Locative media in online mobile spaces." *In Locative media*, pp. 137-151. Routledge, 2014.
[45] Gillespie, Tarleton. "Content moderation, AI, and the question of scale." *Big Data & Society* 7, no. 2 (2020): 2053951720943234.

dominate the industry, the SAFE TECH Act may inadvertently result in an even taller walled garden for large, established platforms to leverage their existing prowess in the industry via their well-established and well-staffed R&D department specifically focused to target these issues with cutting-edge technologies, like artificial intelligence.[46]

In my opinion, the PACT Act is the most interesting and nuanced of the four proposed bills, as it provides explicit protections of consumers and specific, actionable guidelines—derived from the Santa Clara principles—and timeframes to guide platforms.[47] In particular, the PACT Act requires large platforms to provide a complaint and appeals system that produces a moderation decision within 21 days and takedown court-determined illegal content within four days. Smaller platforms have more leeway in content moderation, as appropriate to scale. The act also promotes public-private partnerships in the development of open-source "best practices and guidelines," which draw from a series of transparency protocols, including biannual content moderation reports and accessible acceptable use policies. The nuanced update to § 230 via platform moderation standardization directly responds to the *Zeran* concerns that a platform would defer to immediately taking down content if the notice-based takedown liability applied. However, though the act provides more specificity, that precise specificity raises some First Amendment concerns. Each of the platform requirements (i.e., requiring moderation time guidelines, an appeals system, mandatory reporting) are textbook examples of the federal government enforcing rules on how private companies should regulate content. Further, the distinction between how "large" and "small" platforms moderate content may be an example of speaker-based discrimination, a recently articulated class of free speech in *Citizens United v FEC*

[46] Issie Lapowsky, "Mark Warner Is Ready to Fight for Section 230 Reform," Protocol, March 23, 2021, https://www.protocol.com/policy/mark-warner-section-230.
[47] "Santa Clara Principles on Transparency and Accountability in Content Moderation," Santa Clara Principles, accessed May 10, 2021, https://santaclaraprinciples.org/.

(2010), thereby running the invariable risk of going the same way of the original anti-indecency provisions of § 230—being declared facially unconstitutional.[48]

My Proposal and Conclusion

From our review of the current proposals moving through Congress, I do not believe that any of them sufficiently address the real issues that have resulted from the § 230 liability shield, as we have seen in several court cases. Instead, we can use the existing proposals to make the following wish list as to what any § 230 reform proposal should *not* contain:

1. The proposal cannot delegate the development of "best practices" to a centralized, unelected, government-appointed institution or group of individuals.
2. The proposal cannot explicitly regulate how platforms should moderate user-generated content, through a standardized set of "best practices."
3. The proposal cannot enforce "politically neutral" content moderation.
4. The proposal cannot differentiate between "large" and "small" online service providers.
5. The proposal cannot hinder competition in the market by raising the barrier of entry or placing some other undue burden on online service providers.

This long list of qualifications seemingly puts us back at square one, and for good cause—there is something inherently unappealing about holding companies responsible for the actions of their users. The *Roommates.com* decision already set a precedent for stripping § 230 protections from a provider who actively participates or knowingly facilitates and holding them illegal behavior liable in a court of law. In no reasonable world—in which the Internet survives in its current participatory state—can we reapply publisher liability to online service providers that provide platforms to over 3.4 billion users.

---

[48] Kagan, Michael. "Speaker Discrimination: The Next Frontier of Free Speech." *Fla. St. UL Rev*. 42 (2014): 765; *Citizens United v. Fed. Election Comm'n*, 558 U.S. 310, 130 S. Ct. 876, 175 L. Ed. 2d 753, 78 U.S.L.W. 4078 (2010)

However, we cannot deny that objectively harmful material, unprotected by any law, has flourished in the wake of § 230. Thus, if we must reform § 230, I want to propose that we return to the original intent of the law and re-emphasize that § 230 was meant to protect "Good Samaritans." Specially, we should ask that online service providers, regardless of size, should determine their own "best practices" for moderating the worst of the worst content in exchange for § 230 liability protection. This is a simple reform, but I believe it may be the most effective. This will not place an undue burden on smaller companies, as "best practices" are completely up to interpretation, and larger companies can proceed to moderate content as they have already been doing.

To understand how such a reform would work in the real world, let's take a hypothetical case in which some legal action is brought against some company for the presence of CSAI on their platform. If the company has deliberately refused to moderate any content, they can be held liable in a court of law. Conversely, if the company can demonstrate that they took a series of good-faith actions to moderate such content, they can enjoy § 230 protections. Even though "good faith" effort is incredibly subjective, it is still objectively better than no action. Ideally, if every company enacts some minimal moderation policies, the amount of CSAI and other comparable content can be controlled.

With the advent of the Internet, we no longer live in a society in which the only way to communicate and spread opinions is via the town square or a local newspaper. Every single one of us now has the power to trigger entire social movements or influence outcomes of a political election at scale. With great power comes great responsibility, and as a result, the modern internet is host to some real, pertinent problems that ripple through every aspect of our social life. The pure scale and reach of internet service providers can make it very difficult for us to

judge the net positive of Web 2.0. The same platforms and services that empowered the #MeToo movement and the protests police brutality are the same platforms and services on which vaccine disinformation and QAnon conspiracy theories spread rapidly. It is unlikely that any of these phenomena would have emerged if not for § 230 fostering an online culture of cyber-libertarianism and allowing service providers the freedom and the choice to moderate their platforms.