common sense kids action®

The Internet's "Safe Harbor" is Not Safe for Kids

Section 230's purpose, current interpretation, and path forward.

By Danielle Citron

Danielle Keats Citron is a professor of law at Boston University School of Law, where she teaches and writes about privacy, free speech, and civil rights. She is the vice president of the Cyber Civil Rights Initiative and the author of Hate Crimes in Cyberspace (Harvard University Press, 2014).

Each day, millions of children go online to view and interact with friends, fans, and strangers. But the anything-goes nature of the internet puts kids at risk on sites like Omegle, 4chan, and 8chan, where predators find victims and where harassers threaten and bully the vulnerable. Online interactions can manifest into physical violence and other forms of harm.

Parents and policymakers rely on tech companies to moderate and respond to predation, threats, bullying, and other troubling activity. Some platforms address online abuse because their users and advertisers demand it. Some platforms, however, make abuse their *raison d'être*. One might say, well, those platforms will surely face liability if they encourage and profit from illegality.

Not so, and that is the problem.

Reforming Section 230 -- which shields these companies -- is key to creating a safe and supportive online environment for our children and families. Kids benefit enormously from engaging online. They can publish their own content, discover new ideas, and connect with communities. But kids face clear dangers with harmful consequences.

Many do not realize how much kids are using online tools. The risk that kids might find themselves on sites that do no moderating and tolerate illegality is real. Research shows the popularity of streaming video and social media alongside their pitfalls:

- Kids are watching.
    - Eighty-one percent of parents with children 11 and younger let their kids watch videos on YouTube.[1] Sixty-one percent of these parents say their child has encountered content on YouTube that they felt was unsuitable for children.[2]

[1] Smith, A., Toor, S., & Van Kessel, P. (2018, November 7). *Many turn to YouTube for children's content, news, how-to lessons*. Retrieved from https://www.pewinternet.org/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons/.

[2] *Id*.

    - Eighty-five percent of teens say they use YouTube.[3]
    - Teens use Instagram (61%), Snapchat (63%), and Facebook (43%).[4]
- Though the overall amount of media use is about the same as in past years, how children are using media has shifted considerably.
    - The average amount of time spent with mobile devices each day has tripled (again), going from five minutes a day in 2011 to 15 minutes a day in 2013 to 48 minutes a day in 2017 for children age 0 to 8.
    - Children age 0 to 8 spend an average of 17 minutes a day (17% of all TV-/video-viewing time) watching online videos from a source such as YouTube.[5]
    - Eighty-three percent of kids age 0 to 16 say they watch YouTube for an average of one hour and 25 minutes per day during the week and one hour and 49 minutes on weekends.[6]
- Vulnerable teens feel more vulnerable online.
    - Teens with low social-emotional well-being experience more of the negative effects of social media than kids with high social-emotional well-being.[7]
- Teens are on social media more than ever.
    - A total of 81% of teens use social media, roughly the same as the percentage who "ever" used it in 2012. But it is the frequency of social media use that has changed most dramatically. The proportion of teens who use social media multiple times a day has doubled over the past six years: In 2012, 34% of teens used social media more than once a day; today, 70% do.[8]
- Kids use social media as a source for news.
    - Among children age 10 to 18 who use social media, 76% get news from a social networking site. Of those: [9]
        - Forty-one percent of tweens choose YouTube as their preferred social media site for news.[10]
        - Forty-seven percent of teens choose Facebook as their preferred social media site for news.[11]
- Social media is cause for concern.

---

[3] Pew Research Center (2018, May). *Teens, social media & technology*. Retrieved from https://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/

[4] https://www.commonsensemedia.org/research/social-media-social-life-2018

[5] Rideout, V. (2017). *The Common Sense census: Media use by kids age zero to eight*. San Francisco, CA: Common Sense Media.

[6] https://www.stac-study.org/downloads/STAC_Full_Report.pdf

[7] Rideout, V., & Robb, M. B. (2018). *Social media, social life: Teens reveal their experiences*. San Francisco, CA: Common Sense Media.

[8] Rideout, V., & Robb, M. B. (2018). *Social media, social life: Teens reveal their experiences*. San Francisco, CA: Common Sense Media.

[9] https://www.commonsensemedia.org/research/news-and-americas-kids-infographic

[10] *Id.*

[11] *Id.*

- Fifty-four percent of teens say that if parents knew what actually happened on social media, they'd be a lot more worried about it.

In Section 230 of the Communications Decency Act, lawmakers in 1996 thought they were devising a safe harbor[12] for online providers engaged in self-regulation. The goal was to encourage platforms to clean up offensive material online. Yet Section 230's immunity provision has actually incentivized harmful content on platforms.

Section 230's immunity provision secured breathing space for innovative new ways to work, speak, and engage with the world. But over time the court has applied an overbroad interpretation that has been costly to expression and equality, especially for members of traditionally subordinated groups. Congress should revise Section 230's safe harbor to clearly apply only to platforms that take reasonable steps to address unlawful activity. It is time for platforms to pair their power with responsibility -- especially when it comes to protecting kids.

The early beginnings of the Decency Act

The Communications Decency Act (CDA), which was part of the Telecommunications Act of 1996, was introduced to make the internet safe for kids and to address concerns about pornography.[13]

Besides proposing criminal penalties for the distribution of sexually explicit material online, members of Congress underscored the need for private sector help in reducing the volume of noxious material online.[14] Congresspersons Christopher Cox and Ron Wyden offered an amendment to the CDA entitled "Protection for Private Blocking and Screening of Offensive Material."[15] The Cox-Wyden amendment, codified in Section 230, provided immunity from liability for "Good Samaritan" online service providers that either over- or under-filtered objectionable content.[16]

Twenty years ago, federal lawmakers could hardly have imagined the role that the internet would play in modern life. Nonetheless, Section 230's authors were prescient. In their view, "if this amazing new thing -- the Internet -- [was] going to blossom," companies should not be "punished for *trying* to keep things clean."[17] Cox noted that, "the original purpose of [Section 230] was to help clean up the Internet, not to facilitate people doing bad things on the Internet." The key to

[12] A legal safe harbor typically is an entity that gets immunity from liability if it meets a certain standard of behavior. Yet Section 230(c)(1) does not actually condition filtering too little content on any behavior (unlike Section 230[c][2], which immunizes online service providers from liability for over-filtering if they filtered content in good faith). Section 230(c)(1) offers blanket immunity, whereas 230(c)(2) conditions the immunity on good faith efforts.
[13] S. Rep. No. 104-23, at 59 (1995). Key provisions criminalized the transmission of indecent material to minors.
[14] S. Rep. No. 104-23, at 59 (1995).
[15] H. R. Rep. No. 104-223, Amendment No. 2-3 (1995) (proposed to be codified at 47 U.S.C. § 230).
[16] Pub. L. No. 104-104; *see* H. Conf. Rep. No. 104-458 (1996).
[17] Selyukh, *supra* note (quoting Cox).

Section 230, explained Wyden, was "making sure that companies in return for that protection -- that they wouldn't be sued indiscriminately -- were being responsible in terms of policing their platforms."[18]

The judiciary's interpretation of Section 230 has not squared with this vision. Courts have stretched Section 230's safe harbor beyond what its words, context, and purpose support.[19] Attributing a broad interpretation of Section 230 to "First Amendment values [that] drove the CDA,"[20] courts have extended immunity from liability to platforms that:

- republished content knowing it violated the law;[21]
- solicited illegal content while ensuring that those responsible could not be identified;[22]
- altered their user interface to ensure that criminals could not be caught;[23] and
- sold dangerous products.[24]

Granting immunity to platforms that deliberately host, encourage, or facilitate illegal activity would seem absurd to the CDA's drafters.[25] The law's overbroad interpretation means that platforms have no reason to take down illicit material and that victims have no leverage to insist that they do so.[26] As First Amendment scholar Rebecca Tushnet put it well a decade ago: Section 230 ensures that platforms enjoy "power without responsibility."[27]

---

[18] *Id.*

[19] Danielle Keats Citron, *Cyber Civil Rights*, 89 B.U. L. Rev. 61, 118 (2009). In the landmark *ACLU v. Reno* decision, the Supreme Court struck down the CDA's blanket restrictions on internet indecency under the First Amendment. Reno v. ACLU, 521 U.S. 844, 853 (1997). Online expression was too important to be limited to what government officials think is fit for children. *Id.* at 875. Section 230's immunity provision, however, was left intact.

[20] Jane Doe No. 1 v. Backpage.com LLC, 817 F.3d 12, 25 (1st Cir. 2016).

[21] Shiamili v. Real Estate Group of New York, 2011 WL 2313818 (N.Y. App Ct. June 14, 2011); Phan v. Pham, 2010 WL 658244 (Cal. App. Ct. Feb. 25, 2010).

[22] Jones v. Dirty World Entertainment Holding, 2014 WL 2694184 (6th Cir. June 16, 2014); S.C. v. The Dirty LLC, No. 11-CV-00392-DW (W.D. Mo. March 12, 2012).

[23] 817 F.3d 12 (1st Cir. 2016).

[24] *See, e.g.*, Hinton v. Amazon, 72 F. Supp. 3d 685, 687 (S.D. Miss. 2014).

[25] Cox recently said as much: "I'm afraid . . . the judge-made law has drifted away from the original purpose of the statute." Selyukh, *supra* note. In his view, sites that solicit unlawful materials or have a connection to unlawful activity should not enjoy Section 230 immunity. *Id.*

[26] Citron, *Cyber Civil Rights*, *supra* note, at 118; Mark Lemley, *Rationalizing ISP Safe Harbors*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=979836; Douglas Licthman & Eric A. Posner, *Holding Internet Providers Accountable*, https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1235&context=law_and_economics.

[27] Rebecca Tushnet, *Power Without Responsibility: Intermediaries and the First Amendment*, 77 GWU L. Rev. 101 (2009).

Although Section 230 has been valuable to innovation and expression,[28] it has not been a net boon for free speech.[29] The free expression calculus, as imagined by the law's supporters, fails to consider the loss of voices in the wake of destructive harassment and abuse that platforms have encouraged or deliberately tolerated. As more than 10 years of research have shown, cybermobs and individual harassers shove people offline with sexually threatening and sexually humiliating abuse.[30] Targeted individuals are more often women, women of color, lesbian and trans women, and sexual minorities.[31] Mary Anne Franks' important new book *The Cult of the Constitution* explores how a strike-oriented view of the First Amendment and Section 230 has ultimately been costly to equal protection.[32]

Section 230's broad interpretation enabled 8chan to thrive without any moderation, leaving users to incite violence against immigrant communities as we saw in El Paso, Texas.[33] The benefits Section 230's immunity has enabled likely could have been secured at a lesser price.

Now what?

Some urge Congress to maintain the immunity but to create an explicit exception from the safe harbor for certain types of behavior.[34] A recent example of that approach is the Stop Enabling Sex Traffickers Act (SESTA), which passed by an overwhelming vote in 2016.[35] The bill amends

---

[28] Jack M. Balkin, *The Future of Free Expression in a Digital Age*, 36 Pepp. L. Rev. 427, 434 (2009).
[29] Citron and Wittes, *supra* note, at 410; Danielle K. Citron & Neil Richards, *Four Principles for Digital Expression (You Won't Believe #3!)*, 2018-15 University of Maryland Legal Studies Research Paper (2018).

[30] Pew Research Center. (2017, July 11). *Online harassment study*. Retrieved from http://www.pewinternet.org/2017/07/11/online-harassment-in-focus-most-recent-experience/. (Forty-two percent of people experiencing severe harassment were "more likely to say they changed their username or deleted their profile, stopped attending offline venues or reported the incident to law enforcement"). The individual and societal costs are considerable when victims go offline; lose their jobs and cannot find new ones; and suffer extreme emotional harm in the face of online abuse. Danielle Keats Citron, *Civil Rights in the Information Age* The Offensive Internet: Speech, Privacy and Reputation, (Saul Levmore & Martha Nussbaum, eds. 2010).
[31] *See* Danielle Keats Citron, Hate Crimes in Cyberspace (2014); Danielle Keats Citron, *Online Engagement on Equal Terms*, B.U. L. Rev. Online (2015); Danielle Keats Citron & Mary Anne Franks, *Criminalizing Revenge Porn*, 49 Wake Forest L. Rev. 345 (2014); Mary Anne Franks, *Unwilling Avatars: Idealism and Discrimination in Cyberspace*, 20 Colum. Gender J. L. 220 (2011); Mary Anne Franks, *Sexual Harassment 2.0*, 71 Maryland L. Rev. 655 (2012); Danielle Keats Citron, *Law's Expressive Value in Combating Cyber Gender Harassment*, 108 Mich. L. Rev. 373 (2009); Danielle Keats Citron, *Cyber Civil Rights*, 89 B.U. L. Rev. 61 (2009); Danielle Keats Citron, *Yale ISP – Reputation Economies in Cyberspace Part 3*, YouTube (Dec. 8, 2007), https://www.youtube.com/watch?v=XVEL4RfN3uQ.
[32] Mary Anne Franks. (2019). *The cult of the constitution.*
[33] Danielle Keats Citron. *Cyber mobs, disinformation, and death videos: The internet as it is (and as it should be)*. Michigan Law Review (forthcoming). Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3435200
[34] Beyond the areas that already fall outside the immunity provision: intellectual property claims, ECPA violations, and federal criminal law.
[35] https://www.theverge.com/2018/3/21/17147688/senate-sesta-fosta-vote-anti-sex-trafficking

Section 230 by rendering websites liable for hosting sex trafficking content.[36] Other exceptions could be added, such as those related to combating nonconsensual pornography, and child sexual exploitation.[37]

Congress should avoid a piecemeal approach.[38] Carving out exceptions runs the risk of leaving out other areas of the law that should not be immunized. The law would require updating as new problems arose that would seem to demand it. Legislation requiring piece-by-piece exemptions would, most likely, not get updated.

There's a broader, though balanced, legislative fix. Benjamin Wittes and I have argued that platforms should enjoy immunity from liability *if* they could show that their content-moderation practices writ large are reasonable. Wittes and I offer a revision to Section 230(c)(1) as follows:

> "No provider or user of an interactive computer service that *takes reasonable steps to prevent or address unlawful uses of its services* shall be treated as the publisher or speaker of any information provided by another information content provider *in any action arising out of the publication of content provided by that information content provider*."[39]

The determination of reasonable content-moderation practices would take into account differences among online entities. Internet service providers and social networks with millions of postings a day cannot plausibly respond to complaints of abuse immediately, let alone within a day or two. On the other hand, they may be able to deploy technologies to detect content previously deemed unlawful.[40] The duty of care will evolve as technology improves.[41]

A reasonable standard of care will reduce opportunities for abuses without interfering with the further development of a vibrant internet or unintentionally turning innocent platforms into involuntary insurers for those injured through their sites. Approaching the problem as one of setting an appropriate standard of case more readily allows differentiating between different

---

[36] *Id.*

[37] I contemplated this possibility in my book *Hate Crimes in Cyberspace* (Harvard University Press, 2014). This is as an intermediate, though not ideal, step (Citron & Wittes, *supra* note, at 419).

[38] Citron & Wittes, *supra* note, at 419.

[39] *Id.* I discussed our proposal in testimony before the House Intelligence Permanent Select Committee about deep fakes and other forms of disinformation. See https://www.c-span.org/video/?c4802966/danielle-citron-explains-content-moderation; https://intelligence.house.gov/uploadedfiles/citron_testimony_for_house_committee_on_deep_fakes.pdf.

[40] What comes to mind is Facebook's effort to use hashing technology to detect and remove nonconsensual pornography that has been banned as terms-of-service violations. I serve on a small task force advising Facebook about the use of screening tools to address the problem of nonconsensually posted intimate images.

[41] Current screening technology is far more effective against some kinds of abusive material than others; progress may produce cost-effective means of defeating other attacks. With current technologies, it is difficult, if not impossible, to automate the detection of certain illegal activity. That is certainly true of threats, which require an understanding of the context to determine their objectionable nature.

kinds of online actors, setting a different rule for websites designed to facilitate mob attacks or to enable illegal discrimination from that applied to large ISPs linking millions to the internet.

The public is beginning to understand the extraordinary power that platforms wield over our lives. Social media companies are not simply publishing people's musings. Their terms-of-service agreements and content-moderation systems determine whether content is seen or heard or it is muted or blocked.[42] Their algorithms determine which advertisements are visible to job applicants and home seekers and which are not.[43] Their systems act with laser-like precision to target, score, and manipulate each and every one of us.[44]

To return to Rebecca Tushnet's framing, with power comes responsibility. Law should change to ensure that such power is wielded responsibly. Content intermediaries have a moral responsibility, and companies are beginning to recognize that. As Mark Zuckerberg told CNN, "I'm not sure we shouldn't be regulated."[45]

Of course, far more than Section 230 reform is needed to address the challenges to equality raised by platforms' design and operation, including the collection, use, manipulation, and sharing of users' personal information.[46] We must make policy for the internet and society that we actually have, not the internet and society we believed we would get 20 years ago.

[42] Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 Notre Dame L. Rev. 1035 (2018).
[43] Olivier Sylvain, *Intermediary Design Duties*, 50 Conn. L. Rev. 1 (2018).
[44] Ryan M. Calo, *Digital Market Manipulation*, GWU L. Rev. (2014).
[45] CNN. (2018, March 21). *Mark Zuckerberg in his own words*. Retrieved from http://money.cnn.com/2018/03/21/technology/mark-zuckerberg-cnn-interview-transcript/index.html.
[46] Danielle Keats Citron, *A Poor Mother's Privacy Rights: A Review*, 98 BU L. Rev. (2018); Daniel J. Solove & Danielle Keats Citron, *Risk and Anxiety: A Theory of Data Breach Harms*, 96 Tex. L. Rev. 737 (2018); Paul Ohm, *Sensitive Information*, 88 S. Cal. L. Rev. 1125 (2015); Danielle Keats Citron, *Reservoirs of Danger: The Evolution of Public and Private Law at the Dawn of the Information Age*, 80 S. Cal. L. Rev. 241 (2007).