University of Mississippi

eGrove

Honors Theses

Honors College (Sally McDonnell Barksdale Honors College)

2019

Who's liable? The Intersection of Free Speech and Content Regulation on Social Media Platforms

Hayden Benge
University of Mississippi

Follow this and additional works at: https://egrove.olemiss.edu/hon_thesis

Part of the Communication Commons

Recommended Citation

Benge, Hayden, "Who's liable? The Intersection of Free Speech and Content Regulation on Social Media Platforms" (2019). *Honors Theses*. 1049.
https://egrove.olemiss.edu/hon_thesis/1049

This Undergraduate Thesis is brought to you for free and open access by the Honors College (Sally McDonnell Barksdale Honors College) at eGrove. It has been accepted for inclusion in Honors Theses by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.

WHO'S LIABLE? THE INTERSECTION OF FREE SPEECH AND CONTENT REGULATION ON SOCIAL MEDIA PLATFORMS

by
Hayden Benge

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of the requirements of the Sally McDonnell Barksdale Honors College.

Oxford
May 2019

Approved by

____

Advisor: Associate Professor Cynthia Joyce

____

Reader: Associate Professor Charlie Mitchell

____

Reader: Assistant Dean Scott Fiene

© 2019
Hayden Kelsey Benge
ALL RIGHTS RESERVED

ACKNOWLEDGEMENTS

I would first like to thank my mom and dad, for their never-ending encouragement and support of my endeavors, for instilling in me the value of working hard to achieve my goals and for providing me with opportunities to continuously learn.

I would also like to thank my advisor, Cynthia Joyce, for her abundant ideas, helpful guidance and encouragement to ensure I completed this project to the best of my ability. Thanks to my other readers, Charlie Mitchell and Scott Fiene, for agreeing to read and help make this work better.

I cannot thank the honors college enough for always pushing me outside of my comfort zone and engraining it in me to become a better citizen and a better scholar; for allowing me the opportunity to grow, and for allowing my curiosity to run wild the last four years. And, for offering the course that sparked my interest in free speech, which ultimately led to this thesis. Thank you, Sally McDonnell Barksdale Honors College, and all those who were a part of my education there.

ABSTRACT

Who’s liable? The intersection of free speech and content regulation on social media platforms

(Under the direction of Cynthia Joyce)

This thesis explores the developing legal environment surrounding speech liability, and the extent of free speech that goes with it, on social media platforms. As this new media has grown exponentially in the last decade, the legal questions facing the platforms have also expanded in range, from privacy to security to speech.

By looking at the guiding statute, Section 230 of the Communications Decency Act, as well as the case law involving online intermediary liability, this project uncovers where the law currently stands and what critics point to as its greatest flaws. The current protection given to social media under Section 230 shapes daily interactions online. This thesis addresses what specific areas of the digital world could be impacted by changing Section 230, including the content moderation process and free speech online, as well as how it shapes public discussion and flow of information.

As the issue evolves every day, the findings of this thesis are in no way concrete. Rather, the conclusion looks at a variety of ways that different parties view this area of law, and how they would like to see it develop. Politicians are calling for change to Section 230; free speech advocacy groups calling for it to remain the same; scholars suggesting new theories that challenge and shift the traditional way of viewing the

dynamics of free speech online. While there is no definite answer in 2019, the development of this law has the potential to change the way users on social media interact every day.

PREFACE

This project all began with a class titled “The First Amendment in 2017.” That was two years ago, and little did I know it would spark my interest in the First Amendment and lead me to choose the topic for this paper.

This topic, though not an easy one to tackle, addresses an important area of the law that needs to be discussed, because it has important ramifications in the way our online communities will be shaped in the future.

Social media has changed the landscape of public discourse, dialogues and the interactions we have with each other on a daily basis. By understanding how liability impacts online expression, we can ensure the decisions we make are molding the environments we wish to see in our online communities. The law always lags behind technology, but with the pace at which our world is becoming more interconnected through social media, exploring topics like this one are necessary for the creation of policies that are conducive to the values we wish to see implemented.

So, as our online world continues to grow and more of the population participates in this global community, I hope questions such as the ones posed in this paper invite insightful discussion on how to maintain public discourse and foster positive interactions online, while still promoting the free speech principles championed by the free world.

TABLE OF CONTENTS

I. INTRODUCTION………………………………………………………….. 1

II. CURRENT LEGAL ENVIRONMENT……………………………………. 4

III. INTERNET FREE SPEECH, CONTENT MODERATION AND SECTION 230 CONCERNS………………………………………… 16

IV. A NEW MEDIA: DEFINING ONLINE SERVICES' ROLE……………… 29

V. MAINTAINING PUBLIC DISCOURSE AND INTERNET FREE SPEECH: SUGGESTIONS AND WHERE WE ARE HEADED………………………………………………. 36

VI. CONCLUSION……………………………………………………………46

I. INTRODUCTION

Instances of big technology companies allegedly infringing on the First Amendment rights of their users are occurring at an increasingly rapid pace–this is no secret. Within the past year, Facebook, Twitter and Google have all been summoned before the Senate Judiciary Committee in Washington for various reasons, some of which include First Amendment concerns.

When Alex Jones, an American radio host and widely-known conspiracy theory propagator, got banned from Twitter for violating their abusive behavior policy in early September 2018, the discussion surrounded his First Amendment rights; when the Center for Immigration Studies was banned from using "illegal alien" on Twitter, they too turned to the First Amendment to make their case.[1] Left and right, situations like these are causing users to plead the First against tech companies, claiming that the platforms are actually public forums and therefore, users have the right to post whatever they choose. This, however, is not the argument they should be making.

While speech is protected from limits imposed by bodies of government under the First Amendment, it does not necessarily protect speech posted onto platforms owned by

---

[1] Coll, Steve. "Alex Jones, the First Amendment, and the Digital Public Square." The New Yorker. The New Yorker, April 24, 2019. https://www.newyorker.com/magazine/2018/08/20/alex-jones-the-first-amendment-and-the-digital-public-square.

private companies. If speech violates the policies and guidelines set forth by these platforms, it is within the platforms' power to remove it without being held liable for that speech.[2]

All of these companies–Facebook, Twitter, Instagram– are privately owned, which ultimately gives them the right to regulate their platforms and moderate content however they choose, regardless of the perception that they are free public forums where people can say anything. Up to this point, no case law in the United States has set the precedent for these platforms to be legally considered public forums where speech is broadly protected in the same way that is protected from state actors.

However, there is still friction for social media companies at the intersection of moderating content, allowing certain speech on their platforms and claiming liability for the consequences of certain speech or action. Their content moderation process often gets criticized by the public, but more importantly has gotten them into legal battles. This paper will look at where the line gets drawn–legally speaking–for claiming liability of user-generated content on their platforms and how this, along with other societal factors, play a role in shaping their company content-moderating policies.

How can tech companies, moving forward, moderate user-generated content while themselves avoiding government regulation? It will encourage industry solutions in which social media tech companies can be proactive, in addition to reactive, in the way they regulate content on their platforms. Whether it is ads, hate speech, terrorism or interference with elections, the companies have consistently been playing defense, and

[2] "Section 230 of the Communications Decency Act." Electronic Frontier Foundation, n.d. https://www.eff.org/issues/cda230.

have been trying to catch up with issues in user speech as they happen real time. As the "move fast and break things" mentality credited to tech CEOs like Mark Zuckerberg[3] is beginning to catch up with them, can they implement policies that will help stabilize the volatile environment they have created?

[3] Osnos, Evan. "Can Mark Zuckerberg Fix Facebook Before It Breaks Democracy?" The New Yorker, The New Yorker, 14 Sept. 2018, www.newyorker.com/magazine/2018/09/17/can-mark-zuckerberg-fix-facebook-before-it-breaks-democracy.

II. CURRENT LEGAL ENVIRONMENT: INTERMEDIARY LIABILITY, SECTION 230 OF COMMUNICATIONS DECENCY ACT AND COMPARATIVE CASE LAW

A. Intermediary liability

Social media platforms fall under the protection of intermediary liability models. Defined, this means they are legally akin to messengers who, quite naturally, do not bear any legal responsibility for messages they deliver. This differs from how other media sources such as print publications, broadcast or radio, are held liable under United States law. Because these media sources are considered publishers since they edit and vet their content, they can be held responsible for libelous or defamatory language that is published on their platform. Up to this point, social media companies have been protected under a model of intermediary liability called broad immunity that exists in the United States.[4]

Intermediary liability is the term used to refer to the set of guidelines in a given country that regulate the relationship between what users post to platforms and what the platform can be held legally responsible for. An example of a question these guidelines would answer is: “Should the intermediary service be responsible for individuals posting

[4] “Libel, Slander, Defamation.” Communication Law and Ethics, May 11, 2017. https://revolutionsincommunication.com/law/?page_id=34.

something illegal on their platform?" According to UNESCO, these intermediary liability provisions "formalize government expectations for how an intermediary must handle 'third-party' content or communications," and they vary from country to country. There are three broad categories of intermediary liability models that exist: strict liability, conditional liability and broad immunity.[5]

Strict liability is when the intermediary service is held liable for third-party, or user, content even if the service is not aware of the content being illegal or is unaware the content even exists. This means intermediary services in such countries where this model exists must be extremely proactive in filtering, removing and sifting through user posts that could be considered illegal by the country's government. It also does not matter the size of the intermediary service. All of them are held responsible for monitoring and filtering content submitted by their users to ensure that unacceptable content never gets posted. China and Thailand are both examples of countries that have a strict liability model in place. Repercussions for violating strict liability in these countries include: "fines, criminal liability, and revocation of business or media licenses." In China, intermediaries can be held liable for any unlawful content even if sites are unaware of content and fail to remove it in a timely manner. A 2014 case including Sina.com

---

[5] MacKinnon, Rebecca, Elonnai Hickok, Allon Bar, and Hae-in Lim. "Fostering Freedom Online: The Role of Internet Intermediaries." UNESCO. UNESCO, 2014. https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p%3A%3Ausmarcdef_0000231162&file=%2Fin%2Frest%2FannotationSVC%2FDownloadWatermarkedAttachment%2Fattach_import_24ae9827-6f29-4edd-a133-673ac25384d9%3F_%3D231162eng.pdf&locale=en&multi=true&ark=%2Fark%3A%2F48223%2Fpf0000231162%2FPDF%2F231162eng.pdf#2739_14_CI_EN_int_WEB.indd%3A.130902%3A5399, 39.

required it to lose part of its publishing license because of pornographic material on the network.[6]

Conditional liability is when the intermediary service can be exempt from liability if it meets certain requirements, including removing content upon notice, notifying, upon notice, the user who posted the possibly infringing content, or removing repeat offenders upon notice. If, and only if, the service fails to complete such actions, then it can be held liable for the content. Conditional liability requires little to no proactive intervention of user content that is illegal or infringing. However, this model has been criticized for being susceptible to censorship and little due process for users who wish to appeal the removal of their content. The EU E-Commerce Directive is an example of conditional liability. It allows intermediaries to request immunity from liability if they meet certain criteria.[7]

The last model, broad immunity, gives the intermediary service exemption from a wide range of user-generated content. Broad immunity is often recognized as allowing the most free-flowing dissemination of communication online, and is supported by groups who wish to see principles such as transparency, due process and accountability in the digital world. An example of a broad immunity provision is Section 230 of the Communications Decency Act in the United States.[8]

[6] MacKinnon, 40.
[7] MacKinnon, 41.
[8] See note 7 above.

B. Communications Decency Act

As it currently stands, Section 230 of the Communications Decency Act (CDA) is the applicable legal standard for online intermediary liability cases in the United States. The act was passed in 1996, which predated platforms such as Facebook or Twitter, so opponents to the CDA and Section 230 argue that it is outdated and not suited for prosecuting such companies. Congress, when writing this legislation, stated that the rapid growth of the internet in the early 1990s led to "interactive computer services" as platforms through which Americans were "relying on for a variety of political, educational, cultural, and entertainment services." Due to these findings and the belief that such platforms would enhance the ability to attain education and information, Congress created the CDA to protect online intermediaries from government intervention. In hopes of promoting this continued development, encouraging the further development of similar technologies, and preserving the competitive marketplace online, the CDA was enacted in 1996 immediately upon passage.[9]

Section 230 of the Communications Decency Act, on the issue of content regulation, states:

> "(1) 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.

[9] "47 U.S. Code § 230 - Protection for Private Blocking and Screening of Offensive Material." Legal Information Institute. Legal Information Institute. Accessed April 28, 2019. https://www.law.cornell.edu/uscode/text/47/230.

(2) 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; or

b. Any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1)."[10]

This act, because it was conceived before social media existed, never explicitly states what category social media platforms fall under, which is partly due to the inability to assign an exact definition of what social media does.[11] However, social media companies have typically fallen under the term "internet computer service" in Section 230 and are referred to as such in modern case law. Internet computer service is defined by the CDA as: "Any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including

---

[10] "47 U.S. Code § 230 - Protection for Private Blocking and Screening of Offensive Material." Legal Information Institute.

[11] Selyukh, Alina. "Section 230: A Key Legal Shield For Facebook, Google Is About To Change." NPR. NPR, March 21, 2018. http://www.npr.org/sections/alltechconsidered/2018/03/21/591622450/section-230-a-key-legal-shield-for-facebook-google-is-about-to-change.

specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions."[12]

Under Section 230 of the CDA, social media platforms cannot be treated as publishers, such as newspapers or broadcasters; therefore, they are immune to civil actions for libel or defamation and cannot be held liable for content their users post to their platforms. It also gives them the authority to slightly alter or edit content posted to their platform without revoking their "non-publisher" status. When they begin consistently moderating and actively editing the content posted by users, however, the line gets blurred concerning what their legal responsibilities are. Should they be considered publishers in such instances or should they be given the freedom to moderate harmful and obscene posts and remain immune from liability?

Although they are categorized as an "interactive computer service" in coordination with the language of the CDA, this law does not necessarily cover tech companies' legal responsibilities in the best way. Section 230 is increasingly controversial, and key stakeholders have recently suggested a variety of changes.

Origins

The internet in 1996, in relation to intermediary services, was a much different world than the social media giants who rule the web today. Original intermediary services included websites such as CompuServe, Prodigy and AOL, which offered their subscribers a platform for chats and discussions online. CompuServe, launched in 1979, was the original version of an online platform offering news, chat rooms and file sharing.

---

[12] "47 U.S. Code § 230 - Protection for Private Blocking and Screening of Offensive Material." Legal Information Institute.

These early websites offered their users forums and bulletin board services in which the third-party users could hold discussions.[13] A relic version of it still exists.[14]

Case law leading into CDA

In 1991, one of the first cases that brought online intermediary liability to the legal scene was *Cubby, Inc. v. Compuserv, Inc.* and dealt with the liability of companies concerning individual posts on third-party bulletin boards and forums. In this case, Cubby, Inc., claimed that a public forum on CompuServe included defamatory statements against Cubby and attempted to sue CompuServe, stating that CompuServe, as the publisher, was liable for the post made by a user of the service. The Southern District of New York ruled against Cubby, with a decision that said: "CompuServe has no more editorial control over such a publication than does a public library, book store, or newsstand, and it would be no more feasible for CompuServe to examine every publication it carries for potentially defamatory statements than it would be for any other distributor to do so." This case ruling distinguished CompuServe as a distributor, rather than a publisher, of the user-generated content on its bulletin boards and forums, and could therefore not be held liable for possible defamatory language in each forum.[15]

---

[13] Selyukh, Alina. "The Big Internet Brands Of The '90s - Where Are They Now?" NPR. NPR, July 25, 2016. https://www.npr.org/sections/alltechconsidered/2016/07/25/487097344/the-big-internet-brands-of-the-90s-where-are-they-now.

[14] "Home." CompuServe.com, n.d. https://webcenters.netscape.compuserve.com/home/.

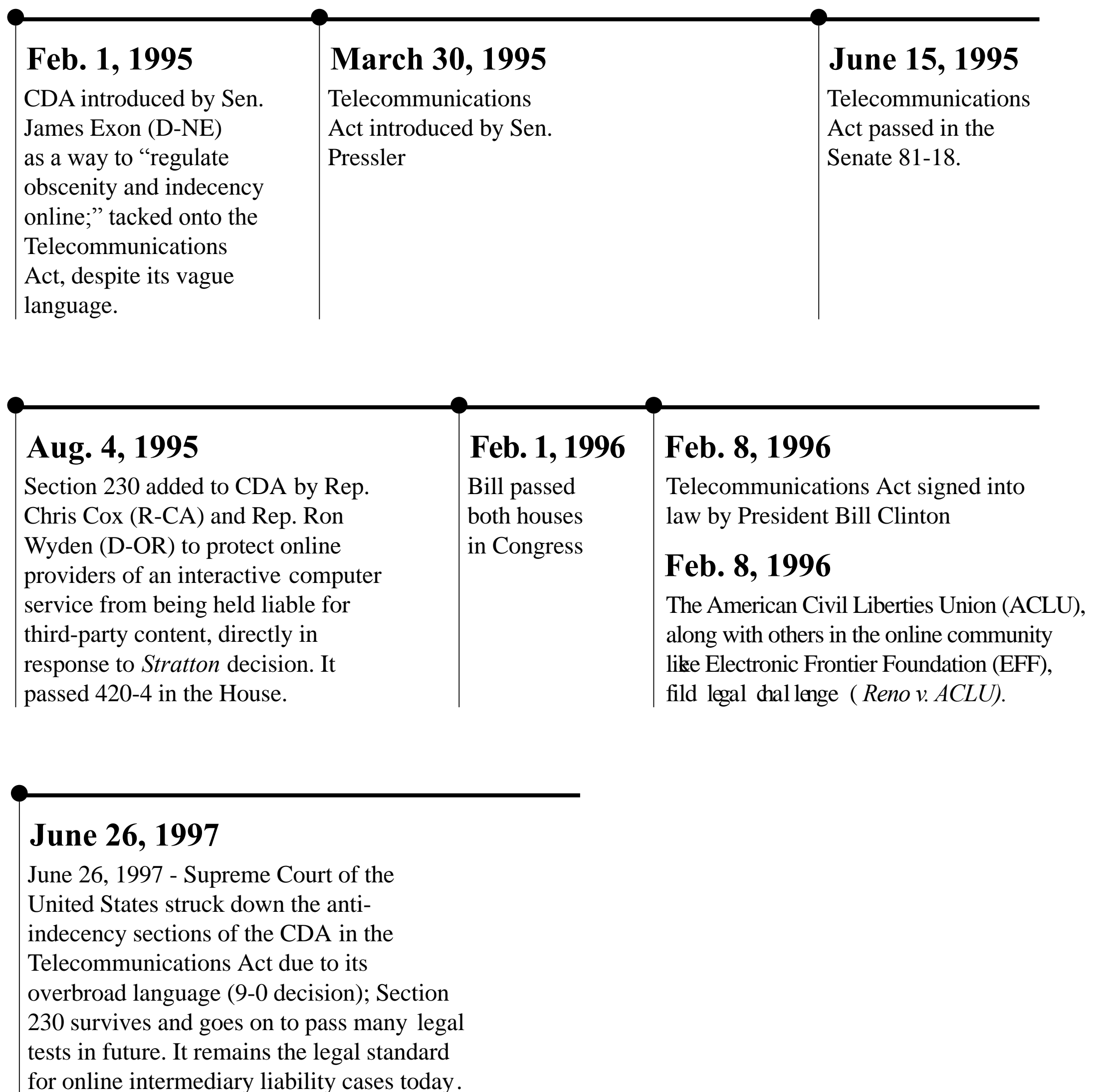
[15] Digital Media Law Project staff. "Digital Media Law Project." Cubby v. Compuserve | Digital Media Law Project. Berkman Klein Center for Internet & Society, 2007. http://www.dmlp.org/threats/cubby-v-compuserve.

The second case came in 1995, and taken in combination with the first, set the stage for Congress to pass the CDA in 1996. *Stratton Oakmont, Inc. v. Prodigy Servs. Co.* was a similar case with a different outcome, in which an anonymous user posted defamatory comments about Stratton Oakmont on a Prodigy web bulletin board. Stratton Oakmont sued Prodigy and the anonymous poster, arguing that the company was acting as a publisher instead of a distributor in this case. The difference between this case and Cubby, in the court's eyes, was that Prodigy employed board members to serve as moderators of content that enforced Prodigy's content guidelines. Stratton Oakmont also pointed to Prodigy's own claims that it had editorial control over content on its servers as evidence that it was acting as a publisher, not a mere distributor, of information. The court sided with Stratton Oakmont, agreeing that Prodigy was acting as a publisher and thus was liable for defamatory comments posted on its bulletin boards. This decision was based mostly on Prodigy's involvement in appointing editorial teams that monitored content posted.[16] According to the Electronic Frontier Foundation (EFF), this decision meant that "just for attempting to moderate some posts, Prodigy took on liability for *all* posts. To avoid liability, the company would have to give up moderating all together and simply act as a blind host, like CompuServe."[17] Within a year of this decision, Congress passed the CDA in response to the strict liability precedent set by this case, and in trepidation of future consequences that would stifle technological progress and freedom.

[16] Digital Media Law Project staff. "Digital Media Law Project." Stratton Oakmont v. Prodigy | Digital Media Law Project. Berkman Klein Center for Internet & Society, October 15, 2007. http://www.dmlp.org/threats/stratton-oakmont-v-prodigy#node_legal_threat_full_group_description.

[17] EFF. "CDA 230: Key Legal Cases." Electronic Frontier Foundation. Electronic Frontier Foundation, January 25, 2018. https://www.eff.org/issues/cda230/legal.

Figure 1: A timeline of the passage of Section 230.



1. Source: EFF. "Legislative Timeline." Electronic Frontier Foundation. Electronic Frontier Foundation, June 5, 2017. https://www.eff.org/issues/cda230/legislative-history/timeline.

The CDA goes to court

Immediately upon passage into law, the CDA came under scrutiny and faced legal challenge in *Reno v. ACLU*. The part of the act that was dedicated to protecting minors from inappropriate content was written in language that was overbroad and vague, using phrases like "obscene or indecent" and criminalizing information depicting or describing "sexual or excretory activities or organs" in an "offensive" manner.[18] The question before the Supreme Court was:

> "Did certain provisions of the 1996 Communications Decency Act violate the First and Fifth Amendments by being overly broad and vague in their definitions of the types of internet communications which they criminalized?"

The Court, in its 9-0 ruling, said yes and struck down the CDA for using overly broad language in its restrictions, leaving only Section 230 of it as the remnant. It was struck down with the reasoning that "governmental regulation of the content of speech is more likely to interfere with the free exchange of ideas than to encourage it," and "the interest in encouraging freedom of expression in a democratic society outweighs any theoretical but unproven benefit of censorship." These statements would be important in setting an ideological precedent for how to handle speech online and avoiding censorship of expression and ideas.

---

[18] "Reno v. ACLU." Oyez. Accessed April 28, 2019. https://www.oyez.org/cases/1996/96-511.

The Court in this case also recognized that the internet was an unprecedented medium for speech, and called it a "vast platform from which to address and hear from a worldwide audience of millions of readers, viewers, researchers, and buyers."[19]

Since the *Reno* decision, Section 230 has withstood many court cases before the judiciary, the first of which was *Zeran v America Online, Inc*. in 1997.[20] This case was the first to apply Section 230 to protect an online service provider, AOL, from being held liable for information posted by a third-party source. In this case, a false advertisement was posted on the website, which resulted in harassment of users. However, the Fourth Circuit Court applied Section 230 in its ruling, stating that AOL could not be held liable for this user-generated content and enforcing broad immunity to online intermediary services. In its decision the Fourth Circuit said:

> "It would be impossible for service providers to screen each of their millions of postings for possible problems. Faced with potential liability for each message republished by their services, interactive computer service providers might choose to severely restrict the number and type of messages posted." And later, "Thus, like strict liability, liability upon notice has a **chilling effect** on the freedom of Internet speech."

Using the language from this court case, the Supreme Court and other federal courts have ruled similarly in subsequent cases asking related questions. The *Zeran*

---

[19] White, Lauren, and Brian Willen. "Amicus Brief in Woodhull v. US." Center for Democracy and Technology. Center for Democracy and Technology, 2018. https://cdt.org/files/2019/02/CDT-amicus-brief-in-Woodhull-v-US-DC-Circuit.pdf.

[20] EFF. "CDA 230: Key Legal Cases."

decision set precedent for how online intermediaries would be treated in United States courts, following the direction of Section 230.[21] .

[21] EFF. "Zeran v. America Online, Inc., 129 F.3d 327 (4th Cir. 1997)." Electronic Frontier Foundation. Electronic Frontier Foundation, November 9, 2012. https://www.eff.org/issues/cda230/cases/zeran-v-america-online-inc.

III. INTERNET FREE SPEECH, THE CONTENT-MODERATION PROCESS AND CONCERNS WITH SECTION 230'S BROAD PROTECTION OF INTERMEDIARIES

The Court's rulings on Section 230 of the CDA have better defined the relationship between online intermediaries, how they moderate content and how it affects individuals' free speech online. Free speech culture enters the picture here in a new, unfamiliar argument in which the effects on speech are directly correlated to the model of intermediary liability in place. If held liable for what their users post, intermediary services will be more sensitive to and strict on moderating content posted to their platforms, causing a "chilling effect" on speech, as reasoned by the Court in *Zeran*. Companies, in order to avoid being sued or getting their privileges revoked, would err on the side of caution and censor more individual speech. For example, as mentioned earlier, the model of strict liability in China causes greater censorship of individuals' speech online than the broad immunity model in place in the United States.

When given this broad immunity to allow a more open public discourse on these websites, it is important that the responsibility of moderating content is still taken seriously. However, when platforms are not performing content moderation in a way that's beneficial to the public, things that would be considered hate speech and fighting

words[22] in the physical world can slip through the cracks of algorithms and human judgment in the digital world. This necessitates a discussion of the dynamics of online speech, how speech and content are moderated online, censorship concerns and why Section 230 protects the cultural tradition of American free speech. In other words, *has online speech gone rogue because of the lack of liability of platforms?*

A. *Speech governed by the state vs. speech governed online by private platforms*

Freedom of speech and expression are not absolute under the United States Constitution. It is true that the freedom of speech clause in the First Amendment in the physical world is not an absolute right of the people. For example, the government can limit free speech on grounds of (1) Libel or obscenity; (2) Threat of violence; (3) Property damage; (4) Criminal speech; (5) Infringing other rights; (6) Burdens on government function; (7) Trespassing; and (8) Time, place, manner restrictions. These limitations work the other way around, too, concerning state actors. The government is not allowed to put limitations on individuals' speech if it is (1) prior restraint; (2) content and view discrimination; (3) overbroad; (4) vague; or (5) has a chilling effect on free speech.[23]

---

[22] Words which by their very utterance are likely to inflict harm on or provoke a breach of the peace by the average person to whom they are directed.
"Fighting Words Legal Definition." Merriam-Webster. Merriam-Webster. Accessed April 28, 2019. https://www.merriam-webster.com/legal/fighting%20words.

[23] Armaly, Miles. "Unprotected Speech." *Constitutional Law*. Lecture presented at the Constitutional Law II, 2018.

However, the way communication is governed by state actors does not necessarily apply to private companies' platforms. Social media companies write the guidelines that outline how they will govern their platforms, and inform users on how they decide to moderate content and speech that is posted from users. Often, the guidelines concerning speech follow similar principles as the First Amendment of the U.S. Constitution. One thing that heavily influences their content moderation guidelines is the American free speech culture in which the companies were established.[24]

Under Section 230, companies have the freedom to choose what speech they limit based on their community standards documents without being held liable for what content or speech they choose not to limit or remove. With this protection from liability under Section 230, companies could potentially moderate speech on their platforms in a stricter manner than the state actors can under the First Amendment. However, in the past, it has ended with opposite results; examples of speech online that could be considered fighting words or hate speech in real life can make it through the moderation process of social media platforms. Identification and anonymity of online profiles make enforcing the restrictions on speech more difficult on digital platforms and can result in more unlimited speech. This speech is also less vetted and unsupported than speech that would appear in a typical "publisher" context.

---

[24] Klonick, Kate. "The New Governors: The People, Rules, and Processes Governing Online Speech." Harvard Law Review. Harvard Law Review, 2018. https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf, 1621.

Figure 2: Facebook hate speech policy as of spring 2019.

Facebook's policy on hate speech[1]

> "We do not allow hate speech on Facebook because it creates an environment of intimidation and exclusion and in some cases may promote real-world violence.
>
> We define hate speech as a direct attack on people based on what we call protected characteristics — race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disease or disability. We also provide some protections for immigration status. We define attack as violent or dehumanizing speech, statements of inferiority, or calls for exclusion or segregation. We separate attacks into three tiers of severity, as described below."

Tier 1

Target a person or group of people who share one of the above-listed characteristics or immigration status, where attack is defined as:

1. Any violent speech or support in written or visual form
2. Dehumanizing speech such as reference or comparison to:
   - Insects
   - Animals that are culturally perceived as intellectually or physically inferior
   - Filth, bacteria, disease and feces
   - Sexual predator
   - Subhumanity
   - Violent and sexual criminals
   - Other criminals (including but not limited to "thieves," "bank robbers," or saying "all [protected characteristic or quasi-protected characteristic] are 'criminals'")
3. Mocking the concept, events or victims of hate crimes even if no real person is depicted in an image
4. Designated dehumanizing comparisons in both written and visual form

Tier 2

Target a person or group of people who share one of the above-listed characteristics or immigration status, where attack is defined as:

1. Statements of inferiority or an image implying a person's or a group's physical, mental, or moral deficiency
   - Physical (including but not limited to "deformed," "undeveloped," "hideous," "ugly")
   - Mental (including but not limited to "retarded," "cretin," "low IQ," "stupid," "idiot")
   - Moral (including but not limited to "slutty," "fraud," "cheap," "free riders")
2. Expressions of contempt or their visual equivalent, including (but not limited to)
   - "I hate"
   - "I don't like"
   - "X are the worst"
3. Expressions of disgust or their visual equivalent, including (but not limited to)
   - "Gross"
   - "Vile"
   - "Disgusting"
4. Cursing at a person or group of people who share protected characteristics

Tier 3

- Calls to exclude or segregate a person or group of people based on the above-listed characteristics. We do allow criticism of immigration policies and arguments for restricting those policies.
- Content that describes or negatively targets people with slurs, where slurs are defined as words commonly used as insulting labels for the above-listed characteristics.

1. Source: "Community Standards." Facebook. Accessed April 29, 2019. https://www.facebook.com/communitystandards/hate_speech.

Some argue that Section 230 ought to be changed so that these companies will be held responsible for eliminating harmful speech. Elie Mystal, executive editor at *Above the Law,* argued on *More Perfect "Twitter and the Law"* that these sites are not constrained by the First Amendment, and therefore, are legally free to regulate speech as strictly as they want. He believes they've allowed Nazis, and similar groups, to organize much more efficiently and they could prevent it if they were to create higher standards of speech. "Twitter trolls want inconsequential free speech," he said. He argues that there is no reason for these groups to exist on such platforms and banning their speech is entirely within the companies' power. Mystal also points to the fact that the speech in posts is already moderated and platforms already choose where they want to draw the line, and how this line could be extended to ban hate speech. If the platforms do not do these things then Mystal and those with similar views would like Section 230 changed so that it requires more strict moderation on unprotected and harmful speech.[25]

The opposing side to this view is the belief that Section 230 should be left alone because public opinion will take care of unwanted speech online through reporting or flagging harmful and offensive content. This side relies on Justice Oliver Wendell Holmes' "marketplace of ideas" theory, based in John Stuart Mills philosophy, which argues that competition of ideas will result in the acceptance of the best and rejection of

[25] Abumrad, Jad, Julia Longoria, Alexander Overington, and Suzie Lechtenberg. "Twitter and the Law." Episode. *More Perfect*. New York City, New York: WNYC Studios, November 6, 2017.

the worst, so all unwanted speech will get discarded by the community at large.[26] Corynne McSherry, legal director at the Electronic Frontier Foundation, argued on *More Perfect "Twitter and the Law"* that changing Section 230 would also lead to more censorship of speech and ideas, and that it could be used against acceptable and valuable speech. In fact, McSherry said that it is the First Amendment right of the companies to allow whatever speech they want to on their platforms. She believes that content moderation by social media sites should be better executed but that Section 230 should be left alone.[27]

These ideas can be categorized as two sides of a larger debate about free speech in general. Those who want to see Section 230 changed often believe that more should be done by the government and those in power to limit or regulate unwanted speech, such as hate speech and fighting words; those who want to see Section 230 remain the same believe in the ability of public opinion to discard and reject hate speech and fighting words without government intervention. They are also concerned with too much intervention leading to censorship and suppression of ideas.

B. The content moderation process and why it is necessary

It is important to understand why such companies ought to have the ability to regulate content and why this is a difficult task to accomplish. Some situations that have caused the public to expect tech companies to take responsibility for content regulation

---

[26] Schultz, David, and David L. Hudson. "Marketplace of Ideas." Marketplace of Ideas. Accessed April 28, 2019. https://www.mtsu.edu/first-amendment/article/999/marketplace-of-ideas.

[27] Abumrad, Jad. "Twitter and the Law."

include: election interference, terrorism, hate speech and flagging misinformation. The platforms have implemented community guidelines to follow when approaching the content moderation process but have to update them often when new problems arise from posts that may not fall under the current guideline standards.[28] The task of content moderation is a difficult one and is scrutinized often. McSherry said a few issues with current guidelines are that they protect certain groups over others, there is not a very good digital due process for those who got reported, and anonymity and speaking online without retaliation present problems with enforcing standards.

With an ever-growing global population participating in social media communities and increasingly posting content to the platforms, the demand for moderating content becomes a more difficult task to accomplish. As of the end of 2018, there were 2.3 billion monthly active users on Facebook and nearly 1.5 billion daily active users.[29] In *We are Social's 2015 Digital Statshot* report, they found that on average, 6 new Facebook accounts are created per second.[30] This growth adds to the

---

[28] "Harmful Content: The Role of Internet Platform Companies in Fighting Terrorist Incitement and Politically Motivated Disinformation." Squarespace, NYU Center for Business and Human Rights, Nov. 2017, static1.squarespace.com/static/547df270e4b0ba184dfc490e/t/59fb7efc692670f7c69b0c8d/1509654285461/Final.Harmful+Content.+The+Role+of+Internet+Platform+Companies+in+Fighting+Terrorist+Incitement+and+Politically+Motivated+Propaganda.pdf

[29] "Facebook Users Worldwide 2018." Statista. Accessed April 28, 2019. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/.

[30] "Global Digital Statshot: August 2015." We Are Social UK - Global Socially-Led Creative Agency, n.d. https://wearesocial.com/uk/special-reports/global-statshot-august-2015.

already difficult task of creating community standards for such a large global platform and moderating content based on those standards.

I. What factors influence this process?

America's deeply-ingrained cultural notion of free speech plays an invisible role in influencing how content moderation takes place on social media. These companies were established in the United States so standards concerning speech are more likely to reflect the culture of those in charge of creating it. This cultural influence has resulted in social media taking a lax approach to moderating or regulating any type of speech posted to their platforms by users. This background for companies' community standards can be tricky, however, since they are used globally. The United States' First Amendment does not apply in other countries, so companies run into issues with their moderating guidelines that are based in this ideology. For example, Thailand threatened to block Youtube from users in its country because of videos that featured Photoshopped images of the king with feet on his head. While these would be considered political cartoons in the United States and protected under the First Amendment, in Thailand it is illegal to insult the king and at the time was punishable by 15 years in prison. Navigating such global, cultural differences makes creating applicable guidelines a tedious process.[31]

Controversies that have caused danger or harm to users have also influenced the way social media monitors user activity. A scenario such as the Russian interference and influence on the 2016 United States presidential election is a case that justifies why content moderation and control is necessary for these platforms. The threat to security

[31] Klonick, Kate, 1623.

that this type of content posed made a lot of users uneasy about the future of elections in the United States and sent Facebook to Capitol Hill to answer questions of concern from lawmakers.[32]

Lastly, fear of government intervention and regulation propels social media companies to moderate content. Similar to the concept of corporate social responsibility, fear of government interference encourages the creation of community guidelines that will at least mitigate some of the most harmful and offensive speech that users post. Correlated to this is the demand from the users themselves for a website that does not contain obscene and harmful content on it. These companies benefit from creating an online environment in which people wish to partake.

II. The actual process:

In 2008, Facebook began writing its first document to provide guidelines on what content the company could remove. This document saw its first complications in its classification of breastfeeding photos as nudity. A protest outside their headquarters led Facebook to adjust the protocol on nudity, but this was only the start to a convoluted process of defining what's allowed and what's prohibited.[33] Facebook's document today is over 80 pages long with only general anecdotes available to the public. These documents have gone from short documents that implement standards (open-ended,

---

[32] Osnos, Evan. "Can Mark Zuckerberg Fix Facebook Before It Breaks Democracy?"

[33] Adler, Simon. "Post No Evil." Episode. *RadioLab*. New York City, New York: WNYC Studios, August 17, 2018.

vague guidelines) to large documents that implement actual rules (specific qualifiers given).[34]

There are different ways content can be moderated. Ex-ante moderation is content that gets removed before it is posted, while ex-post moderation is content that gets removed after it is posted. Automatic moderation is done by algorithms and manual moderation is done by human workers. Reactive moderating is when something is flagged or brought to the attention of moderators, while proactive moderating is when employees seek out the content to remove. More specifically, most ex-ante moderation is done automatically by algorithms, as it gets run through the system while uploading to make sure it does not violate the rules. Ex-post moderation is where the human content moderators become involved in the process.

Facebook has 3 tiers of human content moderators. Tier 3 are those that do the day-to-day content reviewing, tier 2 moderators supervise tier 3 and review prioritized or escalated content, and tier 1 moderators are typically lawyers or policymakers based at Facebook headquarters.[35]

---

[34] Klonick, Kate, 1631.
[35] Human moderators have a high-stress job due to the obscene material they view every day and it has recently been getting media attention. In an article published on *The Verge* called "The Trauma Floor," Casey Newton explores the typical day for a moderator working in Phoenix, Arizona for the company, Cognizant, that works on moderating content for Facebook. Newton found that these people make $28,800 per year, while the average Facebook yearly salary is $240,000.
They are not given many breaks and they can be fired after just a few mistakes since their job is critical in removing unwanted material. These employees have also started developing PTSD after leaving their positions and some even begin believing the radical conspiracy theories they read through because they are exposed to them so regularly. One former employee, Chloe, has developed PTSD-like symptoms that can be triggered from movie scenes that involve gun or knife violence.
These employees are necessary in maintaining the safety of online posts. Newton, Casey. "The Secret Lives of Facebook Moderators in America." The Verge. The Verge,

Moderating content is also a balancing act that sometimes requires decision-makers in the companies to act almost like publishers. When an image of an ISIS member beheading a journalist was posted online, Facebook had to decide whether this powerful image was something people needed to see or if it was inappropriate content. In instances such as this, the line between content moderation and publishing becomes blurred and the role of these companies is ambiguous. Decisions such as this are made on a case-by-case basis, especially when dealing with terror groups or conflict.[36]

C. *Protection of intermediaries hosting criminal activity*

One area in particular where Section 230 protection of online intermediaries attracts criticism is on sites through which criminal activity occurs regularly. Websites such as Backpage and Craigslist have had issues with criminal activity, especially in human trafficking.

A recent case, *Doe v. Backpage,* was not heard at the Supreme Court but raised concerns about whether Section 230 protects the owner of Backpage.com, when the website is contributing to injuries of its users.

In the spring of 2018, as a response to this and similar cases, lawmakers had to make a decision on how to adjust Section 230 so that it no longer protected sites serving as platforms for illegal activity, such as human trafficking.[37] The bill package contained a

---

February 25, 2019. https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.

[36] Klonick, Kate, 1652.

[37] Selyukh, Alina. "Section 230: A Key Legal Shield For Facebook, Google Is About To Change."

House Bill titled "Fight Online Sex Trafficking Act" and a Senate Bill titled "Stop Enabling Sex Traffickers Act," also known as the FOSTA-SESTA package.[38] This legislative measure created an exception to Section 230 in which website publishers would be held liable if third parties post ads for prostitution on their platforms. As with the original CDA, the language in this FOSTA-SESTA package was concerning because of its overbroad language. Because of this, immediately upon its passage, many internet platforms opposed this legislation and a new case– *Woodhull v. US*–is making its way through the courts.

In the most recent Supreme Court decision that deals with protection of criminal activity online, the decision in *Packingham v. North Carolina* held that registered sex offenders have the right to create social media platforms. Although a narrow decision applying to this specific case, language referring to social media platforms as a First Amendment right has future implications that could be used to eventually argue First Amendment public forum online.

North Carolina had created a law that made it a felony for a registered sex offender to use or access any social media used by minors. It was struck down by the Supreme Court in a 5-3 decision, with the citing of First Amendment protection that "all persons have access to places where they can speak and listen." They said that this principle also applied to online forums of the internet, since they "provide perhaps the most powerful mechanisms available to a private citizen to make his or her voice

[38] Romano, Aja. "A New Law Intended to Curb Sex Trafficking Threatens the Future of the Internet as We Know It." Vox. Vox, July 2, 2018. https://www.vox.com/culture/2018/4/13/17172762/fosta-sesta-backpage-230-internet-freedom.

heard. They allow a person with an Internet connection to 'become a town crier with a voice that resonates farther than it could from any soapbox.'" This ruling made it clear that the internet is becoming, if not already, the most important platform for exchanging ideas.[39]

The language in this case could have crucial future implications in linking social media platforms to the definition of modern-day public forums, which would impact how the sites are governed. But, this all depends on how the role of online services is defined.

[39] White, Lauren, and Brian Willen. "Amicus Brief in Woodhull v. US."

IV. A NEW MEDIA: DEFINING ONLINE SERVICES' ROLE

A. Problems and solutions depend on definitions

In traditional media, the guidelines for assigning liability were determined based on their status as publishers who exercise plenary content control. According to the Global Network Initiative, "Intermediary liability" describes the allocation of legal responsibility to content providers of all kinds for regulated categories of content. Because traditional media sources are publishers and editors of the information, this is a fair application of outlining who is liable for what is said on their platforms. With new social media platforms, the issue is that they have thus far claimed to be immune from the "publisher" title. This immunity combined with Section 230 allows more open, public discourse but can upset people if they feel their views are being "edited," or moderated too much by the social media company. The Global Network Initiative describes the apprehension of governments to impose liability on new media platforms as a way to encourage "user free expression, as well as platform innovation, and is often credited with facilitating the tremendous expansion of internet and mobile communications networks across the world."[40]

---

[40] "Intermediary Liability & Content Regulation." Global Network Initiative, Global Network Initiative, 2018, globalnetworkinitiative.org/policy-issues/intermediary-liability-content-regulation/.
Holland, Adam, et al. "Intermediary Liability in the United States." Page Has Moved, Berkman Center for Internet & Society at Harvard University, 16 Feb. 2015, publixphere.net/i/noc/page/OI_Case_Study_Intermediary_Liability_in_the_United_State s

Although the hesitation to charge new media with the same liability as traditional media has been conducive to the growth of the internet and social media, it does not address the question of what role social media platforms play. There is no outline of what a tech company *is*, consequently, there is no outline of what they ought to do and what rules they cannot break or lines they cannot cross. They claim to be hosts of platforms, but many users and lawmakers are unaware of a working definition on what to expect from their services. Pinpointing the role of companies such as Facebook, Twitter and Google has only become more confusing over the years and is now a question that lawmakers are asking. This was apparent when Mark Zuckerberg appeared before Congress and they questioned what these companies actually do and what their roles are. However, Zuckerberg did address the question before Congress about whether platforms still do not consider themselves publishers when he said: "When people ask us whether we're a media company or a publisher, what they're getting at is: do we feel responsible for the content on our platform? I think the answer is clearly yes." These companies are "discovering that they were not just software companies, but that they were also publishing platforms."[41]

Solutions addressing issues in online content moderation and liability models depend on how that party views the role of these online services. It is helpful to outline the roles of similar media industries and how they compare with the structure of social media, an entirely new industry.

[41] Klonick, Kate, 1618.

B. Definitions of previously similar industries

a. Print publishers

In 1974, a unanimous decision from the Supreme Court handed newspapers a definitive legal protection over what role and services of the press the Constitution protects. *Miami Herald Publishing Co. v. Tornillo* was a case in which the *Miami Herald* published two editorials that criticized Pat Tornillo, a candidate for the Florida House of Representatives. Tornillo wanted the *Miami Herald* to publish his response to the editorials but they refused, so he sued in court under a Florida statute that stated political candidates who had been criticized by a newspaper had a right to publish a response to the criticisms.

The *Herald* challenged this statute, saying it violated the free press clause of the First Amendment and the Supreme Court ruled in favor of the newspaper, 9-0. This case set precedent for publishers to have First Amendment protection over their editorial judgments, stating that statutes, such as this Florida one, were "an intrusion into the function of editors" and the press cannot be mandated or regulated by Congress. Chief Justice Burger cited the *New York Times v. Sullivan* case in his decision and argued that Florida's statute limited "the variety of public debate," and so was unconstitutional.[42] The language in this case sets up a pretty clear picture of the Constitutional protection over the press and publishers' role. The Court stated: "the choice of material to go into a newspaper, and the decisions made as to limitations on the size and content of the paper,

---

[42] "Miami Herald Publishing Company v. Tornillo." Oyez, www.oyez.org/cases/1973/73-797. Accessed 28 Apr. 2019.

and treatment of public issues and public officials—whether fair or unfair—constitute the exercise of editorial control and judgment."[43]

While on the surface, similarities can be drawn between newspapers and social media, there are glaring differences. They both offer platforms that facilitate public discourse and conversation, and they both have power to edit or moderate content. However, newspapers operate through reporters, while social media platforms are created by users' decisions to post individually without editorial approval. The function of newspapers is much more vetted, and it is legally acceptable to hold the editors of a newspaper responsible for harmful, obscene and other categories of unprotected content because they are exercising their editorial judgment. On social media, anyone can say anything–shout into the void–and the platforms cannot be held responsible since they do not go through the same editorial judgment process. While they moderate content, they have yet to be charged with the definition of publishers and this may not be a bad thing. According to Steve Coll in his *New Yorker* piece, "This is a be-careful-what-you-wish-for intersection; none of us will be happy if Silicon Valley engineers or offshore moderators start editing our ideas."[44]

b. Broadcast and radio

Social media has a more legally analogous situation to broadcast and radio. Although the Court ruled in *Reno* that broadcast and radio's invasive nature, history of regulation and scarcity of frequencies did not apply to the internet, social media has redefined the nature of online platforms enough to revisit their similarities.

---

[43] White, Lauren, and Brian Willen. "Amicus Brief in Woodhull v. US."

[44] Coll, Steve. "Alex Jones, the First Amendment, and the Digital Public Square."

One case defining broadcast was *Red Lion Broadcasting Co. v. FCC* in 1969. The fairness doctrine of the FCC requires televisions broadcasts to hold fair and balanced discussion about public issues on the airways. In response to this, Red Lion Broadcasting challenged the fairness doctrine on First Amendment claims. The question before the Court was whether the FCC's fairness doctrine regulations violated the free speech clause of the First Amendment.

In another unanimous decision, the court ruled in favor of the FCC, that the fairness doctrine did not violate the First Amendment due to the "spectrum scarcity." In the opinion, the Court actually stated that the fairness doctrine protected free speech rather than infringing it.

The language in this case that legally ties it to new social media platforms is the Court's argument that "without government control, the medium would be of little use because of the cacophony of competing voices, none of which could be clearly and predictably heard." This phrase describing the environment of broadcast media at the time is a near perfect fit for the environment users of social media encounter every time they log in.[45]

C. An entirely new media

Though comparisons can be drawn between traditional media legal protections and responsibilities, social media still presents an entirely new frontier. Because the

[45] "Red Lion Broadcasting Co. v. FCC." Oyez, www.oyez.org/cases/1968/2. Accessed 28 Apr. 2019.

United States legal system operated in the framework of legal precedent, it is difficult to draw analogies from previous cases dealing with traditional media.

No industry is quite like the social media giants that have taken hold of the internet and so defining it in legal terms might require veering a little from precedent. Some want to assign them publisher status, while others would like to see them as a public utility or even a public forum, but none of these things encapsulate the entirety of social media's various roles.

This is why finding solutions to these issues is difficult, because not everyone can agree in what they think social media should be and so they cannot decide on what they want them to be. Even the companies themselves cannot define what their role is, they generally only say what they are not.

Steve Coll argues that: "Facebook and YouTube have long positioned themselves as neutral platforms, akin to eBay, open to all who are willing to abide by community standards. They've resisted the argument that they are in fact publishers—that their human moderators and algorithms function like magazine editors who select stories and photos."[46] However, if they are not categorized as anything then they cannot really be held accountable.

The solution will look different for social media since it is unlike any industry that has come before it, but looking back at the examples of industry solutions in these other areas of media provide some guiding ideas for what tech companies might possibly do.[47]

---

[46] Coll, Steve "Alex Jones, the First Amendment, and the Digital Public Square."

[47] "Understanding the Legal Issues for Social Networking Sites and Their Users." Findlaw, Find Law, 2018, technology.findlaw.com/modern-law-practice/understanding-the-legal-issues-for-social-networking-sites-and.html.

It should not be based entirely in previous definitions, as it is entirely new terrain that warrants groundbreaking language.[48] It is this culmination of definitions that have led to some ideas for how to define and regulate social media giants.

---

[48] Gillespie, Tarleton. "Governance of and by Platforms." Culture Digitally , Sage Handbook of Social Media, 2017, culturedigitally.org/wp-content/uploads/2016/06/Gillespie-Governance-ofby-Platforms-PREPRINT.pdf.

V. MAINTAINING PUBLIC DISCOURSE AND INTERNET FREE SPEECH: SUGGESTIONS THAT HAVE BEEN OFFERED AND WHERE WE ARE HEADED

A person's view of what role platforms play and how they define social media companies tends to influence what proposal they seek to address the state of Section 230 and allowing speech online. Politicians on both sides of the aisle are increasingly calling for some type of governmental regulation on the private companies that own the platforms; meanwhile, the CEOs such as Mark Zuckerberg, are searching for ways to address it on their own terms.

Some nonprofits have weighed in with proposals as well, such as *Article 19,* located in the European Union. The EU has outlined an approach, which includes the creation of a Social Media Council created at a national or international level, or both. It would deal with content regulation issues and be funded by the tech companies who would benefit from it. Their suggestion is based in their research that tech companies differ from traditional media, so their regulation and solutions to issues of liability must differ as well. In the sense of traditional media, the function is to publish and produce

content. Social media takes on an entirely different purpose and ultimately serves a combination of different factors, the main two being hosting and online distribution.[49]

Other nonprofits in the United States like the Electronic Frontier Foundation, the Center for Democracy and Technology and the Knight First Amendment Institute have all played a role in the Supreme Court cases that have upheld Section 230. They defend the protection it provides to platforms and argue that it ought to be left alone. These nonprofits have paved the pathway for many of the legal precedents that currently stand for internet governance and law.

A. Zuckerberg's move toward a private messaging platform

In a Facebook post on March 6, 2019, Mark Zuckerberg made it clear what he believes the solution is: a shift from the "town square" mentality to a "digital living room" mentality.[50] Zuckerberg envisions the future of the internet as intimate through private messaging either in a one-on-one conversation or small group setting. This shift would remedy many of the issues concerning privacy, security and content moderation that have called him to Capitol Hill for questioning from Congress.

His vision of a privacy-focused platform revolves around six principles:

1. Private interactions.

> "We plan to add more ways to interact privately with your friends, groups, and businesses. If this evolution is successful, interacting with your friends

[49] "Self-Regulation and 'Hate Speech' on Social Media Platforms." Article19, Article 19, Mar. 2018, www.article19.org/wp-content/uploads/2018/03/Self-regulation-and-%E2%80%98hate-speech%E2%80%99-on-social-media-platforms_March2018.pdf.

[50] Zuckerberg, Mark. "A Privacy-Focused Vision for Social Networking." Facebook. Facebook, March 6, 2019.

and family across the Facebook network will become a fundamentally more private experience."

This first principle would dramatically affect the "town square" feel that empowers people to speak freely. It would take pressure off of the platform to moderate and intervene with inappropriate posts, since users would be interacting on a much more personal level and communication would not be spread on a massive level. Questions of liability and free speech would, for the most part, become moot.

The current default for user posts is to be shared publicly and available for their "friends" to share to their own timelines as well. This change would shift that public sharing to a more private, close groups of friends in which the conversations are more group-centered.

2. Encryption.

This principle is focused with securing users' privacy so that governments or hackers can't collect personal, private data.

3. Reducing permanence.

Dealing with photos or information posted a long time ago, this principle would aim to reduce permanence of user information by having photos or posts expire automatically, or by allowing users to archive automatically over time. This allows users to control content that could possibly cause problems for them in the future.

4. Safety.

5. Interoperability.

6. Secure data storage.

Facebook also announced on March 27, 2019, "a ban on praise, support and representation of white nationalism and white separatism on Facebook and Instagram, which we'll start enforcing."[51] While the company's policies have long included a ban on hate targeted toward people based on race, ethnicity or religion, this ban would be even tougher in entirely prohibiting anything related to white nationalism or separatism sentiment. The company even stated that people searching terms related to these topics will be redirected to *Life After Hate,* which is "an organization founded by former violent extremists that provides crisis intervention, education, support groups and outreach."

In this announcement, the company also recognized the importance of being faster at finding and removing hate, which would take care of some speech-related issues on their platform.[52]

By making these changes to the design of Facebook's platform, Zuckerberg hopes to address the growing list of concerns users have with social media.

B. Government interference

The government is seriously considering more regulation of social media platforms. If there is one thing the United States Congress can reach across the aisle and agree on right now, it is the necessity of regulating big tech companies. Attorneys general

---

[51] "Standing Against Hate." Facebook Newsroom. Facebook, March 27, 2019. https://newsroom.fb.com/news/2019/03/standing-against-hate/.%E2%80%9D.

[52] Ingram, David, and Ben Collins. "Facebook Bans White Nationalism from Platform after Pressure from Civil Rights Groups." NBCNews.com. NBCUniversal News Group, March 27, 2019. https://www.nbcnews.com/tech/tech-news/facebook-bans-white-nationalism-after-pressure-civil-rights-groups-n987991.

from 14 states met with Jeff Sessions on September 25, 2018, to discuss what to do about them. Louisiana's attorney general suggested breaking them up just as the government did with Standard Oil and Microsoft. This drastic notion is not being taken lightly.[53]

Presidential candidates for 2020 are now dedicating sections of their platforms to big tech companies. Democratic presidential candidate Elizabeth Warren, in her post commented on platforms such as Facebook, Google and Amazon, stating that they have gained far too much power and eliminated any form of competition.[54] She said as president, she would create more competition in big tech.

Warren, as part of her platform, defined her favored methodology:

> "First, by passing legislation that requires large tech platforms to be designated as 'Platform Utilities' and broken apart from any participant on that platform."

Any company with an annual global revenue of $25 billion or more and offer some sort of "public marketplace" would be the companies designated as public utilities. And, "second, my administration would appoint regulators committed to reversing illegal and anti-competitive tech mergers."

---

[53] Selyukh, Alina. "Section 230: A Key Legal Shield For Facebook, Google Is About To Change."

[54] Warren, Elizabeth. "Here's How We Can Break up Big Tech." Medium. Medium, March 8, 2019. https://medium.com/@teamwarren/heres-how-we-can-break-up-big-tech-9ad9e0da324c.

U.S. Rep. Steve King, (R-IA) suggested converting these platforms into public utilities in April 2018 at a House Judiciary hearing.[55] This is a bipartisan issue that both sides are ready to take on and begin formulating government solutions to, as opposed to private industry solutions. King's concern came from the sentiment that social media such as Facebook are biased and the moderating process is more likely to remove content from Conservative Republicans, a growing complaint among this group.

Although he never gave any specific suggestion, former President Barack Obama has spoken out against social media's dangerous ability to spread misinformation if they are not regulating that content. In his speech at the 2018 Nelson Mandela Annual Lecture, he said, "We have to guard against the tendencies for social media to become purely a platform for spectacle and outrage and disinformation."[56]

Not only is the size and scope concerning government officials, but the freedom they have from liability is also becoming an issue. Attacks on Section 230 claiming it is too powerful in protecting the companies from punishment are increasing in number. At the 2019 CPAC, Sen. Hawley–R-Missouri–called for putting restriction on Section 230 to "protect conservative speech," claiming the legislation is outdated and needs to be revamped[57]

---

[55] Constine, Josh. "House Rep Suggests Converting Google, Facebook, Twitter into Public Utilities – TechCrunch." TechCrunch. TechCrunch, July 17, 2018. https://techcrunch.com/2018/07/17/facebook-public-utility/.

[56] Obama, Barack. "Transcript: Obama's Speech At The 2018 Nelson Mandela Annual Lecture." NPR. NPR, July 17, 2018. https://www.npr.org/2018/07/17/629862434/transcript-obamas-speech-at-the-2018-nelson-mandela-annual-lecture.

[57] Lowry, Bryan. "Hawley's CPAC Debut: a Moment in the Spotlight and a Subpoena in Missouri Lawsuit." kansascity. The Kansas City Star, March 3, 2019. https://www.kansascity.com/latest-news/article226980264.html.

Sen. Mark Warner–D-Virginia–in an interview with *The Atlantic,* expressed his interests in also changing the authority of Section 230 and its role in protecting online intermediaries. Warner is a former tech executive who worked in Silicon Valley, and he believes the framework of Section 230 from the 1990s is outdated for the growth social media platforms have experienced. He said, "by around 2016, more than half of the American people were getting their news from Facebook, let alone social media at large. Suddenly, that 1990s framework might not be exactly right." Warner believes that changing the doctrine would not "destroy the public square," but would rather update the law to be functional in the modern-day world of social media.[58]

With all of these threats of regulation coming from lawmakers, if tech companies plan on keeping their control over their platforms, they need to act quickly and present a solution that will quell the worries of government officials.[59]

C. Technological due process and social media as their own governors

Academics and law professionals have provided theories to guide this discussion that are based more in the abstract concepts of how to view these online spaces.

---

[58] Foer, Franklin. "Mark Warner Is Coming for Tech's Too-Powerful." The Atlantic. Atlantic Media Company, December 14, 2018. https://www.theatlantic.com/politics/archive/2018/10/mark-warner-is-coming-for-techs-too-powerful/572695/.

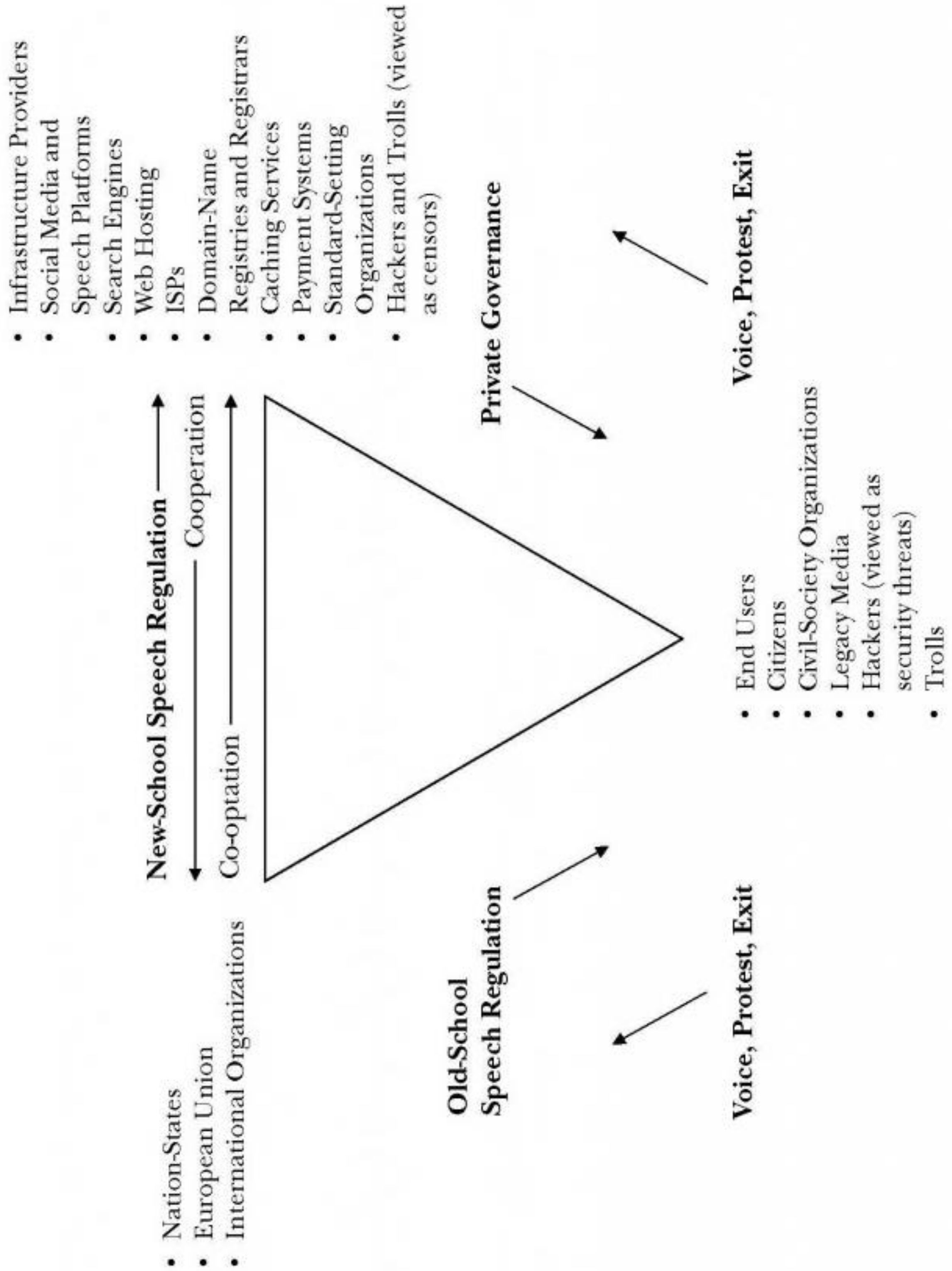
[59] Fung, Brian, and Tony Romm. "Inside the Private Justice Department Meeting That Could Lead to New Investigations of Facebook, Google and Other Tech Giants." *The Washington Post*, WP Company, 25 Sept. 2018, www.washingtonpost.com/technology/2018/09/25/inside-big-meeting-federal-state-law-enforcement-that-signaled-new-willingness-investigate-tech-giants/?noredirect=on&utm_term=.1240243fae10

One theory that frames online speech is the idea of dyadic versus pluralist models of speech governance. This idea is that in the past, the governance model has always been dyadic, or a two-way relationship. Traditionally, on one side is the state and on the other side are speakers and publishers. Since the internet was invented, suddenly there are online platforms that are their own communities, which now creates a pluralist model of speech governance. There is still a state or territorial government on one side and speakers on the other, but now in between them are social media or other forms of online platforms. The dynamic of speech governance has evolved into more of a triangular model in which all three of these participants compete for governing power.[60]

---

[60] Balkin, Jack. "Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation ." UC Davis Law Review. UC Davis Law Review, March 15, 2017. https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3_Balkin.pdf.

Figure 3: Model of pluralistic speech flow.



Another theory to consider is the idea of technological due process. This has more to do with ensuring that individuals' speech is not being censored by the content moderation process of social media platforms. This framework from Diane Citron includes: (1) securing meaningful notice so if an automated system removes or moderates content it ought to have some type of auditing trail that allows the affected user to view why content was blocked; (2) protections for hearings so that complaints will be heard; and (3) releasing source code for a system so that users can know how an automated system is working. Technological due process would ensure certain expectations of users' rights would be met.[61]

The last framework is one in which the companies owning social media platforms privately govern their own spaces as a liaison between the people and the state while still remaining autonomous outside of the territorial government. These companies are already centralized, have governing guidelines for how they moderate their platforms and must adapt based on users' demands. Klonick calls the social media companies the "New Governors of the digital era." Rather than thinking of these companies as companies, then, she suggests people look to them as their own mini-governments that govern online activities.[62]

---

[61] Citron, Danielle Keats. "Technological Due Prcocess." Washington University Law Review. Washington University Law Review, 2008. https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview.

[62] Klonick, Kate, 1662.

I. CONCLUSION: WHERE THIS ISSUE STANDS AND WHERE IT IS HEADED

There is currently a case, *Woodhull v. US,* that is challenging the new FOSTA-SESTA package passed by Congress recently. This case is concerned with the language in the bill package that alters the protection of Section 230 provided to online intermediaries. Nonprofits like the Center for Democracy and Technology and the Electronic Frontier Foundation are concerned that changes like these could alter the path of internet freedom and governance in major ways.

Section 230 and the state of online intermediaries and the future of social media platforms' governance is very fragile and susceptible to change. There is a wave a "techlash," tech backlash, that seems to be catching up with users after the rapid increase in popularity and prevalence of social media platforms in everyday life.

In conclusion to this research, I expect there to be a lot more pressure from outside forces, especially the United States Congress, in reigning in the companies that own social media platforms. Although many ideas have been thrown around, these problems will not be easily or quickly solved due to the lack of precedence in similar industries.

In a year from now, I suspect many things in this paper to be irrelevant, inaccurate or moot, but one thing I know is that this problem will still be a hot issue and politicians running for the 2020 presidential bid will be talking about how to approach big tech.

LIST OF REFERENCES

"47 U.S. Code § 230 - Protection for Private Blocking and Screening of Offensive Material." Legal Information Institute. Legal Information Institute. Accessed April 28, 2019. https://www.law.cornell.edu/uscode/text/47/230.

Abumrad, Jad, Julia Longoria, Alexander Overington, and Suzie Lechtenberg. "Twitter and the Law." Episode. *More Perfect*. New York City, New York: WNYC Studios, November 6, 2017.

Adler, Simon. "Post No Evil." Episode. *RadioLab*. New York City, New York: WNYC Studios, August 17, 2018.

Armaly, Miles. "Unprotected Speech." *Constitutional Law*. Lecture presented at the Constitutional Law II, 2018.

Balkin, Jack. "Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation ." UC Davis Law Review. UC Davis Law Review, March 15, 2017. https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3_Balkin.pdf.

Citron, Danielle Keats. "Technological Due Prcocess." Washington University Law Review. Washington University Law Review, 2008. https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview.

Coll, Steve. "Alex Jones, the First Amendment, and the Digital Public Square." The New Yorker. The New Yorker, April 24, 2019. https://www.newyorker.com/magazine/2018/08/20/alex-jones-the-first-amendment-and-the-digital-public-square.

"Community Standards." Facebook. Accessed April 29, 2019. https://www.facebook.com/communitystandards/hate_speech.

Constine, Josh. "House Rep Suggests Converting Google, Facebook, Twitter into Public Utilities – TechCrunch." TechCrunch. TechCrunch, July 17, 2018. https://techcrunch.com/2018/07/17/facebook-public-utility/.

Digital Media Law Project staff. "Digital Media Law Project." Cubby v. Compuserve | Digital Media Law Project. Berkman Klein Center for Internet & Society, 2007. http://www.dmlp.org/threats/cubby-v-compuserve.

EFF. "CDA 230: Key Legal Cases." Electronic Frontier Foundation. Electronic Frontier Foundation, January 25, 2018. https://www.eff.org/issues/cda230/legal.

EFF. "Legislative Timeline." Electronic Frontier Foundation. Electronic Frontier Foundation, June 5, 2017. https://www.eff.org/issues/cda230/legislative-history/timeline.

"Facebook Users Worldwide 2018." Statista. Accessed April 28, 2019. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/.

"Fighting Words Legal Definition." Merriam-Webster. Merriam-Webster. Accessed April 28, 2019. https://www.merriam-webster.com/legal/fighting%20words.

Foer, Franklin. "Mark Warner Is Coming for Tech's Too-Powerful." The Atlantic. Atlantic Media Company, December 14, 2018. https://www.theatlantic.com/politics/archive/2018/10/mark-warner-is-coming-for-techs-too-powerful/572695/.

Fung, Brian, and Tony Romm. "Inside the Private Justice Department Meeting That Could Lead to New Investigations of Facebook, Google and Other Tech Giants." *The Washington Post*, WP Company, 25 Sept. 2018, www.washingtonpost.com/technology/2018/09/25/inside-big-meeting-federal-state-law-enforcement-that-signaled-new-willingness-investigate-tech-giants/?noredirect=on&utm_term=.1240243fae10

Gillespie, Tarleton. "Governance of and by Platforms." Culture Digitally, Sage Handbook of Social Media, 2017, culturedigitally.org/wp-content/uploads/2016/06/Gillespie-Governance-ofby-Platforms-PREPRINT.pdf.

"Global Digital Statshot: August 2015." We Are Social UK - Global Socially-Led Creative Agency, n.d. https://wearesocial.com/uk/special-reports/global-statshot-august-2015.

"Harmful Content: The Role of Internet Platform Companies in Fighting Terrorist Incitement and Politically Motivated Disinformation." Squarespace, NYU Center for Business and Human Rights, Nov. 2017, static1.squarespace.com/static/547df270e4b0ba184dfc490e/t/59fb7efc692670f7c69b0c8d/1509654285461/Final.Harmful+Content.+The+Role+of+Internet+Platform+Companies+in+Fighting+Terrorist+Incitement+and+Politically+Motivated+Propaganda.pdf

Holland, Adam, et al. "Intermediary Liability in the United States." Page Has Moved, Berkman Center for Internet & Society at Harvard University, 16 Feb. 2015,

publixphere.net/i/noc/page/OI_Case_Study_Intermediary_Liability_in_the_United_States.

"Home." CompuServe.com, n.d. https://webcenters.netscape.compuserve.com/home/.

Ingram, David, and Ben Collins. "Facebook Bans White Nationalism from Platform after Pressure from Civil Rights Groups." NBCNews.com. NBCUniversal News Group, March 27, 2019. https://www.nbcnews.com/tech/tech-news/facebook-bans-white-nationalism-after-pressure-civil-rights-groups-n987991.

"Intermediary Liability & Content Regulation." Global Network Initiative, Global Network Initiative, 2018, globalnetworkinitiative.org/policy-issues/intermediary-liability-content-regulation/.

Klonick, Kate. "The New Governors: The People, Rules, and Processes Governing Online Speech." Harvard Law Review. Harvard Law Review, 2018. https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf.

"Libel, Slander, Defamation." Communication Law and Ethics, May 11, 2017. https://revolutionsincommunication.com/law/?page_id=34.

Lowry, Bryan. "Hawley's CPAC Debut: a Moment in the Spotlight and a Subpoena in Missouri Lawsuit." kansascity. The Kansas City Star, March 3, 2019. https://www.kansascity.com/latest-news/article226980264.html.

MacKinnon, Rebecca, Elonnai Hickok, Allon Bar, and Hae-in Lim. "Fostering Freedom Online: The Role of Internet Intermediaries." UNESCO. UNESCO, 2014. https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p%3A%3Ausmarcdef_0000231162&file=%2Fin%2Frest%2FannotationSVC%2FDownloadWatermarkedAttachment%2Fattach_import_24ae9827-6f29-4edd-a133-673ac25384d9%3F_%3D231162eng.pdf&locale=en&multi=true&ark=%2Fark%3A%2F48223%2Fpf0000231162%2FPDF%2F231162eng.pdf#2739_14_CI_EN_int_WEB.indd%3A.130902%3A5399.

"Miami Herald Publishing Company v. Tornillo." Oyez, www.oyez.org/cases/1973/73-797. Accessed 28 Apr. 2019.

Newton, Casey. "The Secret Lives of Facebook Moderators in America." The Verge. The Verge, February 25, 2019. https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.

Obama, Barack. “Transcript: Obama's Speech At The 2018 Nelson Mandela Annual Lecture.” NPR. NPR, July 17, 2018. https://www.npr.org/2018/07/17/629862434/transcript-obamas-speech-at-the-2018-nelson-mandela-annual-lecture.

Osnos, Evan. “Can Mark Zuckerberg Fix Facebook Before It Breaks Democracy?” The New Yorker, The New Yorker, 14 Sept. 2018, www.newyorker.com/magazine/2018/09/17/can-mark-zuckerberg-fix-facebook-before-it-breaks-democracy.

"Red Lion Broadcasting Co. v. FCC." Oyez, www.oyez.org/cases/1968/2. Accessed 28 Apr. 2019.

"Reno v. ACLU." Oyez. Accessed April 28, 2019. https://www.oyez.org/cases/1996/96-511.

Romano, Aja. “A New Law Intended to Curb Sex Trafficking Threatens the Future of the Internet as We Know It.” Vox. Vox, July 2, 2018. https://www.vox.com/culture/2018/4/13/1717276 2/fosta-sesta-backpage-230-internet-freedom.

Schultz, David, and David L. Hudson. “Marketplace of Ideas.” Marketplace of Ideas. Accessed April 28, 2019. https://www.mtsu.edu/first-amendment/article/999/marketplace-of-ideas.

“Section 230 of the Communications Decency Act.” Electronic Frontier Foundation, n.d. https://www.eff.org/issues/cda230.

“Self-Regulation and ‘Hate Speech’ on Social Media Platforms.” Article19, Article 19, Mar. 2018, www.article19.org/wp-content/uploads/2018/03/Self-regulation-and-%E2%80%98hate-speech%E2%80%99-on-social-media-platforms_March2018.pdf.

Selyukh, Alina. “Section 230: A Key Legal Shield For Facebook, Google Is About To Change.” NPR. NPR, March 21, 2018. http://www.npr.org/sections/alltechconsidered/2018/03/21/591622450/section-230-a-key-legal-shield-for-facebook-google-is-about-to-change.

“Standing Against Hate.” Facebook Newsroom. Facebook, March 27, 2019. https://newsroom.fb.com/news/2019/03/standing-against-hate/.%E2%80%9D.

“Understanding the Legal Issues for Social Networking Sites and Their Users.” Findlaw, Find Law, 2018, technology.findlaw.com/modern-law-practice/understanding-the-legal-issues-for-social-networking-sites-and.html.

Warren, Elizabeth. “Here's How We Can Break up Big Tech.” Medium. Medium, March 8, 2019. https://medium.com/@teamwarren/heres-how-we-can-break-up-big-tech-9ad9e0da324c.

White, Lauren, and Brian Willen. “Amicus Brief in Woodhull v. US.” Center for Democracy and Technology. Center for Democracy and Technology, 2018. https://cdt.org/files/2019/02/CDT-amicus-brief-in-Woodhull-v-US-DC-Circuit.pdf.

Zuckerberg, Mark. “A Privacy-Focused Vision for Social Networking.” Facebook. Facebook, March 6, 2019. https://www.facebook.com/notes/mark-zuckerberg/a-privacy-focused-vision-for-social-networking/10156700570096634/.

———. “Digital Media Law Project.” Stratton Oakmont v. Prodigy | Digital Media Law Project. Berkman Klein Center for Internet & Society, October 15, 2007. http://www.dmlp.org/threats/stratton-oakmont-v-prodigy#node_legal_threat_full_group_description.

———. “The Big Internet Brands Of The '90s - Where Are They Now?” NPR. NPR, July 25, 2016. https://www.npr.org/sections/alltechconsidered/2016/07/25/487097344/the-big-internet-brands-of-the-90s-where-are-they-now.

———. “Zeran v. America Online, Inc., 129 F.3d 327 (4th Cir. 1997).” Electronic Frontier Foundation. Electronic Frontier Foundation, November 9, 2012. https://www.eff.org/issues/cda230/cases/zeran-v-america-online-inc.