12-2022

Hitting Refresh: Regulating internet speech in the 21st century

Jack Charles Valentino
Bates College, jvalenti@bates.edu

Hitting Refresh:
Regulating internet speech in the 21st century

An Honors Thesis

Presented to
The Faculty of the Department of Politics
Bates College
In partial fulfillment of the requirements for the
Degree of Bachelor of Arts

Jack Charles Valentino

Lewiston, ME
March 30, 2022

Acknowledgments

First and foremost, I would like to thank my advisor, Professor Alyssa Maraj Grahame, for her thoughtful interventions, moral support and overall partnership throughout this project. Amidst my frenzy of questions and ideas, her guidance at our Monday morning meetings made this project possible. I also owe a debt of gratitude to the many other professors in the department over my time at Bates who have strengthened my research abilities, my writing, the way I think about politics, and my ongoing interest in the law.

To my Mom, I cannot thank you enough for inspiring this love of policy, your ever-ready ears and eyes throughout this endeavor, and your moral support. I joke about the clippings of the New York Times Technology section you would snail-mail to my campus address (yes, the irony is unmistakable), but I am forever grateful for your dedication to this project. To my Dad, thank you for inspiring me to undertake this year-long endeavor in the first place and reminding me that big projects like this happen one step at a time.

Last but not least, I would like to thank my roommate, Dylan DiSunno, for his never-ending role as our suite's morale officer amidst our late nights and early mornings working on thesis together. Your comedic interjections and indulgence in my many thesis and non-thesis-related thought experiments kept me going.

Table of Contents

Table of Figures

Abstract

By the mid-1990s, the internet had taken new form outside of its original military applications and became commercially available at an unprecedented rate. Western democracies recognized that such a new frontier, therefore, necessitated regulation. Their shared goal was to restrict objectionable content while simultaneously creating a pathway for this nascent industry to blossom. With this in mind, both the U.S. and European Union enacted linearly different and mutually exclusive regulatory regimes to govern "online intermediaries," or sites like Facebook and Twitter that merely host the speech of their users. The E.U. enacted aggressive content removal statutes, while the U.S. offered nearly blanket immunity to these sites in the hope that the marketplace of ideas would dilute objectionable content. Using the U.S. and Germany as case studies, this thesis argues that, twenty years later, neither pathway emerged particularly victorious in their quest to curb the dissemination of radicalizing content. I find that the failure under the German *preemptive framework* derives from a contradictory monitoring obligation and lack of oversight by the European Commission on the state. Conversely, I find that the failure under the American *deregulatory framework* is rooted in a contradictory allocation of jurisdiction and a lack of oversight by the state upon intermediaries. By scrutinizing the incentive structures of both countries' regulatory regimes, this thesis challenges the way Western democracies conceptualized and continue to conceptualize the internet and points out how neither extreme has responsively moderated internet speech.

Introduction

Addiction to social media and negative perceptions of one's self are pervasive issues amongst my generation and one's that certainly deserve attention from a variety of fields. There are volumes of books being written about the effect that childhood social media usage has on later personal development. But such is out of the scope of what I am seeking to investigate. This thesis targets the extent to which the reality around us, not so much our conceptions of ourselves, is distorted by the content we consume online. There is an important subset of social media studies emerging as a result of our entrance into what many pundits are calling the resurgence of a "post-truth era." The validity of the world around us is up for debate and the internet offers people an interactive opportunity to distort their reality. Our ability to develop and comprehend ideas and think for ourselves is jeopardized by our overreliance on fast news, curated content, and inflammatory responses.

I've seen both ends of this spectrum firsthand. Classmates at my all-boys, Catholic high school became consumed by the radicalizing content of Tucker Carlson, Alex Jones, and the late Rush Limbaugh. It's still tough to see, almost four years later, the social media content my high school classmates continue to repost: misinformation about vaccinations or voter fraud in addition to generally xenophobic, racist, and sexist content. They give power to this content every time they retweet, reshare, or comment. With all the love I have for Bates, I can say at times it is no different. I've been disheartened by the sight of my fellow Batesies shouting down an invited speaker in the spring of 2019, targeting specific right-leaning students by name in hurtful online campaigns in the fall of 2021, and largely shutting down avenues of healthy discourse in our recently reopened Commons dining hall that don't align with the talking points of the New England liberal left.

None of this is unique to the internet and none of this should come as a shock. Political radicalization is a widely discussed concept in the field of political science. These instances of left and right-wing radicalization are problems, but not the crux of my investigation – they are symptoms of a more deadly issue. Events like January 6th in the United States, the attempted Day X in Germany, and unfortunately many others are all the result of people whose reality had been distorted to an even more harmful degree, who were empowered to believe that their fringe beliefs were mainstream maxims, and who felt obliged to act on that empowerment. The inspiration behind this thesis was to investigate radicalization that begets violence and particularly how attempts to mitigate that pipeline have uniquely failed.

The most recent comprehensive internet regulations in the E.U. and the U.S. took place in 2000 and 1996, respectively. Over the past 20-25 years, we've seen a dramatic increase in the incidence of domestic terror. In fairness, this is the result of a variety of factors. But social media has played a not insignificant role in that phenomenon. As a result, Section 230, the E-Commerce Directive, and other major pieces of legislation have become household names, or at least have entered mainstream debate. Further, the majority of articles on content moderation speak to its need for reform. To me, the question became: then why hasn't it happened in the last two decades? Why has objectionable content continued to radicalize the polity? It's certainly not out of a lack of effort or legislation, so why have we not seen results?

Both the U.S. and Germany enacted linearly different and mutually exclusive regulatory regimes to solve the problem of the internet being increasingly used to host and amplify objectionable content. Further, by regulatory regimes I mean the whole body of the regulation and not just the way the bills themselves. Rather regulatory regimes are meant to refer to the interaction of the various branches of government (legislative, executive, judicial, state-level,

Federal-level, etc.) to execute the philosophical regulatory framework each body set out to employ. However, neither direction has proven to be demonstrably more efficient in reducing the radicalization of people to the fringes.

In researching this topic, I hypothesized that it's because both bodies are metaphorically throwing ladders into quicksand. Western democracies are using a shaky foundation to support their attempts to escape from sinking. Both the U.S. and Germany have been using exogenous solutions to solve their endogenous problems. If Western democracies want to meaningfully improve the quality of online speech and reduce the influence of radicalizing content, then they need to bite the bullet and reflect on the foundations of their regulatory regimes. The puzzle then becomes: how did both linearly different pathways uniquely fail?

The literature review establishes where the conversation about internet regulation began and speaks about how that has either changed or stagnated. Chapter 1 opens with a review of the origins of the *preemptive* and *deregulatory* frameworks that I define. The chapter couches these frameworks in existing national principles and values. Next, it provides background on the mutually exclusive nature of these frameworks by briefly citing the failed attempts of both sides to pass legislation that would align them more with the opposing framework. This is done to reify the unique nature of the puzzle: why did two mutually exclusive and linearly different pathways yield the same result? Next, the literature review offers historical evidence of radicalizing content abusing technology even before 1996. This demonstrates that the mutual failure of both bodies is not easily explained by a lack of foresight. The framers of both bills knew clearly that radicalizing content would be a problem after their regulation like it was before it. Chapter 1 then moves away from the scholarly work about the bill's passage and moves into discussing the extremism that abuses the bill. The literature is clear that there is a demonstrable

rise in violent extremism across many metrics. But the literature is split over whether there is a causal link between extremism online and reciprocally manifested violence. Finally, I offer scholarly background on the parts of each framework that I hypothesized were the reason for each framework's failure.

The research design I employed in answering this question needed to rule out certain externalities. The implementation or statutory language of their attempts alone cannot be to blame. Further, not knowing there was a problem that would need to be solved is not evidence of failure. Therefore my evidence needed to prove that both bodies not only (1) suspected this radicalization was a problem on the horizon but (2) drafted their original legislation in a way that was wholly incongruous with their given policy priorities. At a basic level, I compared the trajectory of their regulatory frameworks over time and the incentive structures created as a result. At a higher level, I then offer ways that their frameworks (*preemptive* or *deregulatory*) didn't fit with the regulatory regime (the accompanying cast of political institutions or systems used to support that framework) that surrounded it. This incongruous relationship between their framework and regulatory regime allowed objectionable and radicalizing content to take root in these gaps.

I argue in Chapter 2 that the E.U. made clear their intent to harmonize national laws and recognized the transnationality of the internet; and yet, still chose to pass their hallmark piece of legislation as a *directive* and not a *regulation*. Regulations align all applicable, underlying laws and replace the existing law of each member state word-for-word. Directives, on the other hand, can be implemented by member states with an unnecessarily generous amount of deference to the will of the national legislature of each member state. Allowing for each E.U. member state to maintain its own underlying criminal codes and set its own enforcement mechanisms makes

regulating a transnational industry difficult. As a result of the lack of oversight by the E.U. upon Germany's transposition of the Directive to national law, Germany exacerbated the existing issue of heckler's vetoes. Heckler's vetoes allow people to report counterspeech (well-intentioned content that brings the debate closer to the mainstream and away from the fringe) and get it taken down thereby allowing objectionable content to rise as a proportion of all content.

I argue in the U.S. chapter (Chapter 3) that Congress made clear its intent to prioritize free speech considerations and to vest in the courts the power to ameliorate complaints. This becomes a problem, however, because courts are reactive, not proactive. Without a regulatory body to issue guidelines, this "court order standard" leaves no pathway for removing radicalizing content until the harm has already been committed. The U.S. desire to have this industry be free from proactive regulation but also clean enough for people to want to use social media and grow the industry is incongruous. As a result of the blanket legal immunity and lack of a proactive regulatory agent like the FCC to establish guidelines, ISPs are under no incentive to remove objectionable content. Such content, therefore, increases in quantity even if it represents a smaller proportion of all content.

Failing to do anything and continuing along with the status quo can only exacerbate both the frequency with which objectionable content makes its way to impressionable eyes and the power of platforms to resist future regulation by the state. But by critiquing the incongruous regulatory regimes of two of the world's preeminent beacons of democracy, I don't mean to signal that this is either an unsolvable problem or that the only solution is tighter government control. In Chapter 4, I'll briefly lay out the censorship of modern authoritarian countries to demonstrate the other end of the spectrum. This chapter is meant to illustrate (1) the costs of

having an open internet and (2) the fact that fully eliminating objectionable content and maintaining democratic ideals are mutually exclusive.

Chapter 1: Literature Review

The "internet" has undoubtedly defined life in the late-20th and early-21st centuries. Its regulation affects the way we transact business, conduct the functions of government, pass down traditions, and connect with our communities during times of isolation. Suffice to say, the internet has become integrated with arguably all previously analog interaction humans had with the world around them. What largely differentiates the internet from other forms of communication is that it is quickly translatable, instantly available, interactive, transnational without additional infrastructure and exponentially evolving.

Although this combination has made the internet one of the most potent tools for change (as evidenced by movements like the 2011 Arab Spring and #MeToo), the internet also has proven to be a tool for radicalizing, organizing terror, and disseminating other generally objectionable material. Intermediaries, like Twitter and Facebook, took center stage after it was discovered that more and more extremist ideologies were using their sites to organize and radicalize. Like any other industry, the fine line between the costs and benefits of this tool lay in its regulation and oversight. Once the internet started to become commercially more accessible in the late 1990s, Western governments conceptualized the regulation of online intermediaries in two linear and mutually exclusive ways. On one hand, the United States passed the Communications Decency Act (CDA) of 1996 and endorsed a *deregulatory framewor*k (Kuczerawy 2015). Section 230 of the CDA immunizes internet service providers (ISPs) from being responsible for monitoring or liable for hosting the content posted by their users. On the other hand, the European Union passed the E-Commerce Directive (ECD) of 2000 and endorsed a *preemptive framework.* Article 14 of the ECD holds intermediaries liable for the content

created by their users once they are made aware that such content is illegal or objectionable under certain guidelines (Keller 2018).

Academics in the field often use the term *internet exceptionalism* in conversations about online intermediary liability and the regulation of speech on the internet. Interestingly, the term has become co-opted by both frameworks and construed to fit their side. Representatives Chris Cox (R-CA) and Ron Wyden (D-OR), in 1995, used the term as a justification for passing their "26 words that created the internet," Section 230. They believed the untapped potential of the internet would become a blossoming industry and a tool for change that justified the blanket liability protections they were offering to ISPs (Kosseff 2019). To them, the internet didn't need to abide by the regulations that currently governed analog companies that provided the same services (ie: newspapers, magazines, and books) (Bechtold 2020). On the other hand, the European Commission used the same principle of *internet exceptionalism* to put the onus on ISPs to police content. They recognized early that the internet was too complex and constantly evolving for governments to keep up with and that only ISPs were able to police themselves and therefore the liability should be on them to set enforceable standards and remove objectionable content (Bechtold 2020).

Matthew Collins was an early supporter of a different argument all together with his 2001 book *The Law of Defamation and the Internet*. He argued that the medium shouldn't make a difference; to Collins, libel was still libel whether it was on a screen or in a newspaper. These sites, to Collins, should be regulated commensurate with how their analog counterpart would be with no additional onus on the government or the ISP. He was an early rejector of internet exceptionalism, which is a stance that Neil Fried (2021) argues is becoming the modern norm. Companies like Twitter, Facebook and Google are increasingly being called before Congress to

answer questions about what they are doing to make the internet less toxic and objectionable. Even Facebook could read the writing on the wall and acknowledged in a recent advertisement campaign that their heyday of internet exceptionalism may be coming to a close. Facebook's new advertisements[1] show pictures of technology in the 1990s and then a timeline with how that technology has evolved and concludes with the title "The internet has changed a lot since 1996… Facebook supports updated regulations." Suffice to say, Collins' refutation of internet exceptionalism was foreshadowing the position the literature is in now – reconsidering fixes to the 25-year old regulations because the content on the internet has become increasingly objectionable.

Twenty-five years ago these ordinarily allied bodies chose two completely different frameworks and that should rightfully come as a shock, but scholars explain this schism by tracing a line back to the underlying principles of each side. Eliza Bechtold notes that such a deregulatory framework is rooted in the "American tendency to value liberty over equality, a staunch commitment to individualism, and an enduring tradition of negative freedom" (2020, 41). In stark contrast, she notes that Europe has historically preferred "the attachment of rights and responsibilities to the enjoyment of individual rights, the imposition of positive obligations on the government, and the balancing of the rights of an individual speaker with the rights of others, including the targets of virulent speech" (42). To Damian Tambini, this American conception of negative rights arose as a result of British colonial crackdowns on "speech, media, and debate through censorship and particularly through taxation" and continues today as a result of America's incentive to be a hub of competition (2021, 301). Europe, on the other hand, is scarred by the images of the private monopolization of the media that occurred during its quarrel with fascism during World War II and their preemptive framework has continued today as a

[1] https://www.ispot.tv/ad/OHcg/facebook-internet-regulations-remember-the-internet-in-96

result of Europe's interest in maintaining its status as a bastion of international human rights enforcement (Tambini 2021). As a result of these disparate conceptualizations of protecting liberty, Europe is proactive in putting the onus on intermediaries while the U.S. is more deferential to platform expression (Bechtold 2020) (Julia-Barcelo and Koelman 2000).

Because of these fundamental disagreements over how a government should ensure the protection of liberty, each side has been historically unable to bridge the divide and find a middle ground. Each side has attempted to incorporate aspects of the other framework, but both times failed. There will be a more comprehensive discussion in future chapters about these case studies. But there is a body of literature that speaks to the way the American SESTA/FOSTA Act of 2018 and the European Commission review of the E-Commerce Directive in 2010 represent attempts to assimilate some of each other's frameworks. However, they both uniquely failed to achieve each of their desired policy goals as a result of institutional limitations within the U.S. and European Union (Kuczerawy 2015) (Weisner 2020).

This kind of mutual exclusivity would lead to the belief that one side got it right and one got it wrong. But in reality, both could be seen as a failure. Quantifying this failure has been a meteoric rise of objectionable, radicalizing, and extremist content on the internet. The internet was no stranger to this type of content before the enactment of the CDA and ECD, but neither side has seen a reduction in extremist content as a result of their respective frameworks. Dating all the way back to the 1980s, there existed groups like the Aryan Nation Liberty Net who used dial-up technology to connect people to White Nationalist resources in their area (Conway et al. 3). Then in the early 90s, *Stormfront* came to prominence as one of the original platforms for unadulterated xenophobic and racist extremism (Walther and McCoy 2021). The literature agrees

that the internet has always had elements of being a breeding ground for radicalization and extremism but not to the same extent as it exists today.

In the past five years, right-wing extremism has increased worldwide by almost 320% (Institute for Economics and Peace 2019, 44). In an analysis of almost 70 years of extremism in the United States, PIRUS data (Profiles of Individual Radicalization in the United States) revealed that it took extremists an average of 18 months in 2005 to come across their first radicalizing content but in 2016, they were coming across that same content an average of five months faster (Jensen et al. 2018, 1). Daniel Koehler corroborates this sentiment with his research on formerly radicalized extremists in Germany. His interviews illuminated the scope of this issue. Not only was the internet effective in radicalizing users but it was considered the most effective tool to do so. All of the formerly radicalized extremists interviewed labeled the internet as the most frequent and most valuable tool they used to amass an army of supporters (Koehler 2014).

Despite the internet being a transnational apparatus, researchers have remarked that this culture of digital hate, both past and present, is something uniquely Western (Ganesh 2018) (Institute for Economics and Peace 2019). This is likely a result of the fact that two of the major Eastern countries in the world (Russia and China) are synonymous with complete internet censorship, an extreme that's likely to reduce radical organizing but, in turn, runs wholly opposed to the guiding principles of liberal democracy. The question still remains as to why America and Germany both struggle uniquely with radicalization and extremism despite both of them purportedly being some of the world's leaders in internet regulation. Is there a link between the regulation of the internet and the frequency of extremist terror? The literature is split but leans in an unexpected direction.

Bechtold explains that a country's approach to governing intermediaries is based on a "framework of causation and harm." Essentially she posits that Europe believes there is a translation between extremist content online and instances of domestic terror, while America believes that such a link is either nonexistent or outweighed by a greater factor (2020, 36). Personally, she repeatedly sides with the American conceptualization and argues that there is no strong causal link between online activity and domestic terror (Bechtold 2020). Walther and McCoy (2021) support this; their research clarifies that while there was certainly an empirical rise in the size of extremist groups on Telegraph after the Christchurch attacks in 2019, there was still no causal link to say that the growth of these online groups was the cause of future attacks. There is anecdotal evidence of things like copy-cats or other "inspirational" attacks, and there is clear evidence of echo chambers being created when RWE (right-wing extremist) content is reshared to similar people through the intermediary algorithms (Conway et al. 2019, 8).

But the literature has struggled to find a provable causality between the prevalence of extremist language on the internet and occurrences of terror activity. Ganesh acknowledges that such a causal link is yet to be proven but presents evidence that online speech is substantially related "in limited cases" to the rise of terror activity (2018, 42). There has been overwhelming evidence, like Guhl and Davey's (2020) research of 208 white supremacy chat rooms, to prove that online extremists often advocate for violence. But their research admittedly never yielded any conclusive proof to suggest that these incitements materialized into actual violence. While there is disagreement as to the totality of this lack of a causal link, the literature does seem to largely support the American conceptualization of causation and harm which feels there is a tenuous, if any, link between online extremist content and real-life terror activity.

Despite this tenuous link, we would still expect that governments would feel obliged to address it. Although scholars seem to feel the likelihood of online speech inciting violence is low, the impact of such violence is unmistakably high. So why have both of the current models failed to adequately address the exponentially growing problem of digital hate? It's been discussed how both are linearly different and mutually exclusive from each other, so we would expect one extreme would fare better, even if just marginally, than the other. But that is empirically not what scholars or citizens have noticed. As mentioned, extremism is at an all-time high and rising in both the U.S. and the European Union. Each model has failed for three unique reasons; acknowledging that both extremes have been a failure has become instrumental in the way some scholars have begun to propose solutions.

One of the earliest failures of the deregulatory framework is how naively tailored it is. Section 230 was originally imagined to govern one-on-one interactions. Representatives Wyden and Cox drafted this section as a response to *Stratton Oakmont v. Prodigy* (1995). The case held that Prodigy was liable for the defamatory speech posted by its users against Mr. Belfort and his company. Congress took issue with the effect this would have on the growth of a new industry and essentially immunized the "middle man." They wanted to steer litigation away from the host of the speech and back between the one who said it and the one requesting redress for the speech. Neither Cox, nor Wyden, nor any of the other early drafters imagined that this section would eventually be cited in the proliferation of extremists and terror content. The scholarly work published on intermediary liability between the years of 1997 and 2000 refers to Section 230 and "defamation law" as one and the same (Friedman and Buono 2020) (Middendorf 1998) (Sheridan 1997). Take as a counterpoint Professor Eric Goldman's recent list of the top ten most important Section 230 cases since then. Although anecdotal, the list demonstrates how expansive

this issue has become. Of the ten cases, only three of them are concerned with person-to-person defamation, and only one of those cases took place before 1999. The other seven cases concern: copyright and trademark infringement, violent extremism, sexual assault, sex trafficking, software contracts, online rental services, and search engines (Goldman 2017). The main section only needed to be twenty-six words long because the framers of the section only imagined it being used for cases similar to *Prodigy*. The overly broad language has since been construed to protect dozens of sites hosting objectionable content.

In addition to institutional limitations, political limitations exist that have hindered the U.S. deregulatory model. There has been a shift in the Overton Window that has come as a result of Trump's presidency. To Bechtold, these conceptions of "causation and harm" that differentiate the two frameworks don't necessarily have to be real. Her research speaks to the risk inflation that is commonplace in the rhetorical toolbox of both the EU and the US. Both have incentives to over-hype the risk of terror threats to suit their own policy agendas. Former President Trump has routinely inflated international threats while downplaying the role of domestic terrorism (Bechtold 2020). For example, his rhetoric simultaneously downplayed the risk of violence from the Proud Boys while overinflating the risk of terrorism from the South in an effort to secure more funding for his border wall (Walther and McCoy 2021). This rhetoric is blurring the line for these groups; when the President of the United States downplays the role of the Proud Boys as a terrorist organization, their speech moves farther from the margins of society, is considered more acceptable in public discourse, and is tougher for intermediaries to police as extremist or radicalizing (Conway 2020). Furthermore, intermediaries that choose to call out right-wing politicians like Trump for underinflating the risk of certain domestic groups get labeled as censoring conservative voices.

It is well-documented by scholars that American conservatives encourage their supporters to transition to more conservative-friendly alternative intermediaries, like Gab.ai or Parler, when there is suspected censorship and these sites gradually become increasingly lax about content standards with each migration and iteration (Ganesh 2018) (Conway et al. 2019) (Conway 2020) (Walther and McCoy 2021). Therefore, as Ganesh argues, a potentially healthy discussion about free speech on the internet is being held hostage by far-right populists that control the flow of users that are "red-pilled" into believing that there is rampant censorship on mainstream intermediaries (2018, 43). These reactionary sites being created, like Gab in 2017 as a response to the December 2017 #TwitterPurge of some conservative commentators, are less interested in working towards a healthier online discourse and unfortunately gain legitimacy with every right-wing politician that joins the site (Walther and McCoy 2021). This creates a dangerous race to the bottom that one could refer to as "intermediary Darwinism." The sites that gain the most users are the ones that offer the least stringent community guidelines and the least oversight, therefore driving extremism farther and farther underground and out of the hands of more mainstream sites that have a vested interest in appearing and being more socially responsible (Conway et al. 2019).

A final barrier to the success of the deregulatory framework that arose in the literature is that the American approach creates a slippery slope for technology firms to write their own regulations. Contrary to the original purpose of the ECD which was to harmonize the laws amongst member countries, the purpose of Section 230 in the U.S. was to bolster a nascent industry (Kightlinger 2003) (Kosseff 2019). Because of the broad immunity these intermediaries are afforded in the name of internet exceptionalism, they are able to simultaneously maintain more of a user base and spend less on moderating content. The American deregulatory system is

inherently capitalist and social responsibility has taken a back seat to profit maximization. The American intermediary industry was able to blossom with the protections of Section 230 and their wealth and influence exponentially grew by both providing a quasi-public resource and by emphasizing the complexity of their industry. Despite the sector's immense wealth, they are so influential in setting their own regulation because they provide an invaluable service to the government. Coordination with the U.S. intelligence community and other surveillance-related national interests are often cited as a reason to overlook the sector's monopolistic tendencies, economically inefficient copyright protections, and undemocratic commercialization of essential digital services (Popiel 2018). Cusumano et al agree adding that the government has progressed down this slippery slope so far that they've largely lost control of the content on the internet (2021). Especially with the Federal government being preoccupied with the recent pandemic, "big tech" has been given even more fuel to grow beyond the regulatory reach of national or international governments (Jacobides 2020).

Second, the industry contains a bias towards institutional knowledge. Due to the complex nature of the technology sector, there exists a pervasive "revolving door" where regulators often come from the industry and vice versa (Gilens and Page 2014, 567; Popiel 2018, 575). This is another tool used by the industry to regulate themselves, placing their former employees within the regulatory agencies and recruiting former regulators to work in the private sector. Interestingly, one check on this power offered in the literature is that these firms are still beholden to the court of public opinion. Mehra and Trimble (2014) argue that intermediaries like Facebook and Twitter ended up voluntarily complying with a variety of the provisions in the bills they've lobbied against out of an interest to keep a positive appearance with their user base but only the provisions that suit their needs and demonstrate a willingness to 'take action' on the

problem. That being said, there is still a large disconnect between the entitlement felt by intermediaries and the dwindling power of the government that allows these digital hate cultures to grow (Ganesh 2018).

Although the European model imposes more constraining (and therefore controlling) obligations upon intermediaries, the preemptive framework isn't without its flaws. Underlying principles of "sovereignty and subsidiarity" drive the European Union to enact many of its internet regulations as voluntary initiatives or frameworks whose implementation is up to individual member states (Bechtold 2020). One of the most well-known clarifications in the E-Commerce Directive is the "country of origin" clause that states every intermediary is subject to the level of implementation in their domiciled country (Kightlinger 2003). This means if Twitter.eu was domiciled in France, the content hosted on their site is accessible anywhere in Europe but would be subject to the moderation requirements that France implemented whether that be equal to or more stringent than the baseline standards set in the Directive. Inherently, no country wants to become the host for objectionable sites. This creates another dangerous race to the bottom where countries are incentivized to each develop gradually more and more stringent moderation requirements in excess of the minimums set by the Directive to avoid becoming the haven for objectionable sites (Bechtold 2020). This constantly tightening regulatory regime incentivizes platforms to flee to the US and more content to be driven underground through an increasing number of alternative platforms like Gab and Parler.

Another reason why the European model has failed to keep objectionable discourse from permeating their sites is that their preemptive framework creates a perverse incentive for "heckler's vetoes." Heckler's vetoes are a byproduct of the "notice and takedown" clause within the E-Commerce Directive. Sites are only immunized insofar as they remove material that is

made known to them to be illegal (Keller 2018). This framework establishes a vague standard for what is considered a "notice," however, and intermediaries are incentivized to "automatically and systematically" remove content relating to any notice filed. These act as "heckler's vetoes" that greatly restrict the growth of intermediaries, the flow of verified information, and the functioning of democracy (Julia-Barcelo and Koelman 2000). But more importantly to the original design of the ECD, these heckler's vetoes incentivize extremist groups to "report" credible information and then inundate the platform's marketplace of ideas with their own conceptions of reality (Ganesh and Bright 2020). This "anticipatory obedience" deprives the government of needing to pass actual laws to justify the content being taken down and users that post credible information often don't have an opportunity to contest the removal creating an oversaturation of incorrect/objectionable content with an undersaturation of peer-reviewed content (Bechtold 2020).

The final barrier that scholars offer to the European preemptive framework's success is that the European Union doesn't have clearly defined criteria for "terrorism" to equally apply. Ever since 9/11, the UN and international human rights NGOs have repeatedly issued warnings that the overly broad definition of terrorism is a danger to the freedom of expression (Bechtold 2020). The extreme right is "composed of a fast-changing and complex overlapping of individuals, groups, movements, political parties and media organs" making them particularly hard to root out as a group (Conway 2020, 1). Although not a domestic terror organization, the case study of ISIS and the Taliban in the mid-2010s proves this point. ISIS was a much more defined group that was agreed upon as a terrorist group by the international community. On the other hand, the Taliban was comparatively less organized and less agreed upon as a terrorist group at the time and was able to flourish on social media as a result. Because of this disparity,

Facebook was able to curtail reportedly 99% of ISIS recruitment content by late-2017 and Twitter was able to report similar while Taliban recruitment content was able to flourish as a result of the fact that they don't neatly fit into the organized parameters of objectionable content as defined in the EU (Conway 2020) (Walther and McCoy 2021). This lack of a clear definition for terrorism in Europe stems from the same rationale for why the US might inflate or deflate threats – to malleably suit their given policy priorities.

Evidently, the models have had their problems, but why is this important? Scholars have come to agree that the largest impact of these framework failures is on democracy. They forewarn that bolstering extremist content and simultaneously driving it underground will have deleterious effects on our ability to interact with basic democratic functions. Ganesh (2018) warns that in a general sense, not dealing with this culture of digital will alienate populations and accelerate support for autarkic and isolationist policies that will curtail the soft power of global liberal democracy. Furthermore, Walther and McCoy (2021) warn that letting ideological, alternative sites like Gab and Parler go unchecked leaves voters susceptible to information manipulation by both domestic political interests and foreign actors. The biggest problems for Europe given the cracks arising in their framework is that their "notice and takedown" clause will result in heckler's vetoes that lead to censorship and that the race to the bottom created through the "nation of origin rule" will lead to more and more sites being driven underground. Similarly, the most pressing impact on American democracy is that ideological alternative sites will continue to pop up as a reaction to intermediaries challenging politicized risk deflation and that the internet will generally become a repository for objectionable material as companies try to maximize profits.

In summary, the internet has developed into an exceedingly powerful tool. Intermediaries have an immense ability to disseminate information and change the way we interact with the world around us. Previously, we conceptualized the internet as exceptional and requiring a unique set of rules different from its analog counterparts, but that school of thought has disappeared as conservatives allege censorship among mainstream platforms and progressive liberals demand more action be taken to reduce the amount of objectionable content hosted online. The intermediaries themselves are even seeing a change in the tide and are beginning to adapt to it. So the question arises, "how did we get to this point?" Despite their mutually exclusive and linearly different approaches to regulating online intermediaries, why have both the U.S. and the E.U. largely failed to curb the rise of extremism on social media? In describing the way that each side has (1) failed to find a middle ground and (2) actually created perverse incentives for intermediaries and users to act irresponsibly, this thesis seeks to challenge the way the two frameworks actually fit into the larger regimes (courts, jurisdictions, structures, etc) around them. Maybe it was adequate at slowing extremism's rise when the scope of the internet was smaller, but it no longer is now that the internet is more complex and covers more than just defamation like originally assumed. At risk is the fate of democracy and our ability as a polity to know what is true and what is propaganda.

Chapter 2: Germany

To Piter and his friends, trips to the Arena Bar and Cafe in Hanau, Germany were a staple of their post-workout routine. As he recounts, he brought a slice of pizza from around the corner on the night of February 19, 2020 into the bar and searched for his friends. Shortly after sitting down, the group was jolted by a singular shot which he, Momo, and many other bar patrons all report to believe was a blank from a fake gun or part of a prank going on outside. Unbeknownst to them, this was neither harmless nor concluded. Seconds later, an unmasked man burst into the Arena Bar and Cafe with four handguns and opened fire. He recounts that the kiosk and host area was the first target and the gunman made his way towards the back of the restaurant. Three of Piter's friends were killed at the front of the restaurant and his other friend Momo, was shot in the shoulder next to Piter. Piter jumped on top of Momo and another patron jumped on top of Piter to hide behind a bar table. The gunman continued shooting what was later reported to be four handguns and, per eyewitness testimony, he emptied the magazines before walking out calmly. When the rampage was over, Piter was relieved to hear Momo. However, he was distraught in discovering that the patron on top of him had died as a result of shielding Piter from the spray of bullets. Of Piter's group, only Piter and Momo survived and Piter was the only patron in the busy cafe to come out unharmed.

The gunman was later identified to be 43-year-old German, Tobias Rathjen. Rathjen arrived at the Arena Bar and Cafe after already killing a taxi driver and opening fire at another predominantly-Turkish bar across the river. Rathjen's story ends when he is discovered at home, dead alongside his mother in an apparent murder-suicide. But the story for the Hanau community and the fourteen victims begins much earlier when he posted his xenophobic and conspiratorial manifesto online and a YouTube video alleging the existence of a secret surveillance operation in

Germany that he purports to have discovered online. Not only was he radicalized by the content posted by others before him, but his own content (YouTube channel and personal website) would remain online for days after the attack. It was eventually removed, but not before the dark corners of the internet cached his content and disseminated it out of the mainstream.[2]

In his manifesto, he alleges that there are ethnic groups that threaten Germanic purity but that they cannot be expelled legally (alluding to Germany's ongoing acceptance of millions of refugees). He lists off many Southeast Asian Middle Eastern nations explaining that they will have to be expelled manually and that it is his job to maintain the rights of ethnic Germans even if it meant not only killing millions of immigrants but also the Germans who supported the policies that allowed their arrival. Rathjen was a known far-right extremist, whose ideas closely align with groups like the Identitarian Movement that have been known to infiltrate and coopt mainstream hashtags to slip xenophobic content by ordinary detection.[3] Unfortunately, Rathjen's radicalization and ability to find similarly xenophobic Germans to support his aggression is not unique. Right-wing extremism is an issue so pervasive to German society that Interior Minister Horst Seehofer proclaimed it to be not only at its highest level since 2001 but more alarmingly "the greatest threat to security in our country."[4] Rathjen was not a singular "bad apple," but rather a momentary window into the larger, tainted farm.

---

[2] Keys, Matthew. 2020. "Frankfurt Attack Suspect Posted Manifestos, YouTube Videos before Shooting Spree." *The Desk*. https://thedesk.net/2020/02/tobias-rathjek-manifesto-attack-terrorism-germany-frankfurt-hanau/ (March 24, 2022).

[3] Scholz, Kay-Alexander. 2020. "How the Internet Fosters Far-Right Radicalization." *Deutsche Welle*. https://www.dw.com/en/how-the-internet-fosters-far-right-radicalization/a-52471852 (March 24, 2022).

[4] Gehrke, Laurenz. 2021. "Germany Records Highest Level of Right-Wing Extremist Crime in 20 Years." *POLITICO*. https://www.politico.eu/article/germany-records-highest-level-of-right-wing-extremist-crimes-in-20-years/ (March 24, 2022).

History of Domestic Radicalization

In response, a critical reader could argue that Rathjen was a "lone wolf" and that his radicalization is (1) not evidence of a pervasive or systematic blunder in online content moderation and (2) equally possible under a traditional medium like radio, television, newspapers, or books. However, such an argument fails to account for the unique features of the internet – those features which incidentally also make the internet so popular and valuable to society, the medium's ability to amplify and reproduce content. Contrary to radio or television, social media allows for the amplification of material beyond their original transmission. In contrast to traditional information mediums, social media allows unvetted users to not only share with large audiences but repackage the information and disguise it in new ways. One user's allegation of "conspiracy" within the government could lead five more users to confirm that 'belief,' three more to add onto it with additional and typically false information, six more to reshare it to their circles, and a handful who will migrate that content to different platforms.

The process gets out of hand, a movement is given credence, and well-intentioned content (called 'counterspeech') is unable to dilute the inflammatory, outlandish rhetoric. As a result, an unquantifiable number of larger plots exist representing systematic radicalization opportunities that thrive outside of the public eye. Rathjen was a lone wolf in this attack, but the only reason we know about him is because his plot was unfortunately successful. The amplifying and reproductive characteristics of the internet not only creates pathways for the rise of lone wolves like existed pre-internet, but now also connects lone wolves to each other in a clandestine ecosystem (Hamm and Spaaij, 2017). In 2017, Germany saw the actualization of such a phenomenon previously inapplicable to traditional mediums (ie: radio, television, newspapers).

Starting in 2015, a user on Telegram named *Hannibal* began a forum that invited police and military officers to connect. Innocuous as it may sound, it would soon take a turn for the worst. The conversation gradually shifted away from job-related bantor when Hannibal started provoking xenophobic, anti-immigrant rhetoric within the forum. Eventually, the group grew, the well-intentioned officers left the forum, and the group became highly concentrated with xenophobic, anti-immigrant members of the German defense forces. By early 2016, this group had reached the upper echelons of the KSK (a German equivalent to the Navy Seals) and other exclusive branches of the German military. Eventually a rumor is started that there is going to be an overthrow of the government on a certain day, called *Day X*, in which the government will be completely overwhelmed at the mercy of immigrants and pro-immigrant politicians who had allegedly been coordinating this from within the German government for years. This is a motif to what Q'Anon dubs "the storm." These former and current military members begin stockpiling weapons and communicating with each other about how they will rendezvous on the day of the event and kidnap the pro-immigrant politicians to prevent them from executing their alleged plan.

By late 2016, plans had been codified and rehearsed for kidnapping these politicians, routing out immigrants, and protecting Germany from the ethnic pollution that they alleged would come from allowing non-natives to settle in Germany. Members of the conspiracy established other branches elsewhere in Germany and began stashing weapons in strategic locations. Later, court documents would reveal that members of this group were going so far as to case the houses of targeted politicians. The group started off as an outlet for military members and police officers to converse and has now, with the help of Telegram, has reached exponentially more people than a singular radio broadcast or newspaper ever could. The forum

became a nexus for previously unaffiliated lone wolves to collaborate and give credence to each other's beliefs with an efficiency nonexistent in the days before the internet. From here, Franco A enters the picture. Hannibal got the group started, but Franco was Hannibal's "boots on the ground" coordinating logistics and resources for the impending Day X. It wasn't until a handgun was found in the wall of the Vienna International Airport men's bathroom in 2017 that this plot would begin to unravel. German government institutions investigated the group and uncovered how unsettlingly deep it went. High ranking military members in many of Germany's elite branches were discovered to have been involved in this plot against left-wing politicians.

Although this plot was foiled, Day X and its de facto leaders Hannibal and Franco A serve to prove the point that the internet is a place that can uniquely bring together average citizens to do terrible things. Previously, large scale efforts of terror and disruption required an organizing force like a terrorist group or other non-state actor, but (as I'll further elaborate in Chapter 3) this is no longer the case. Average people are finding comfort and validation amongst others who confirm their unorthodox beliefs. In many ways, Day X is analogous to the insurrection that occurred on January 6th at the U.S. Capitol. The insurrection was also a congregation of predominantly unaffiliated, lone wolf rioters that were inspired by the rhetoric on sites like Gab to assemble at a location and form ad hoc affiliations with each other based on this shared goal. The internet is uniquely creating an ecosystem for people to not only reach a larger number of people than ever before, but more unfortunately receive feedback from other fringe users that confirms their conspiratorial ideas.

	Date	Incident	Casualties	Platforms Used
2020	June 8	**Christchurch Copycat** A 21-year-old from Hildesheim posted in an online chat forum that he intended to kill as many Muslims as possible and referenced recordings of the live-streamed Christchurch Mosque Massacre in 2019.	His plot was foiled and he currently stands trial for "preparing a state-endangering act of violence"	Undisclosed chat forum, Facebook
	February 19	**Hanau Shootings** *See above.*	Nine (9) dead, at least five (5) wounded	Personal web domain, YouTube
2019	October 9	**Halle Synagogue Attack** Stephan Balliet attempted an attack on a synagogue on Yom Kippur.	Two (2) dead, three (3) injured	Twitch, AWS, Kohlchan
	June 2	**Death of Council President Walter Lubcke** Lubcke was killed at close range by an unaffiliated, lone wolf gunman who had previously been radicalized by and posted radicalizing videos to YouTube.	One (1) dead	YouTube
2018	August 26	**Stabbing in Chemnitz** Suspect later identified as "Alaa S." stabs a man in Chemnitz and (before the police could issue their report) social media alleges that the crime was a result of immigrants. Rumors spread quickly online and inspired neo-Nazi's and far-right extremists to take to the streets in protest. AfD politicians churned up unsupported rumors and anti-Immigrant sentiment and called the attack a "knife migration." Hundreds of right-wing protestors take to the streets, riot, damage properties, enact Nazi salutes, and chase immigrants around the city.	One (1) dead, twenty-three (23) injured	Various platforms
	December 27	**Stabbing in Kandel** Afghan refugee stabs his ex-girlfriend and extremist sites exploit the incident on social media	One (1) dead, but unknown numbers of refugees injured	Various platforms

2017		to fuel xenophobic sentiments, denounce Germany's mirgration policy, and instigate violence against refugees.		
	August 25	**Indymedia Ban** The Ministry of the Interior bans linksunten.indymedia.org, a known far-left organization with a history of inspiring violence and inflammatory riots. The site took to various other platforms to plead that it was being censored. They riled up their base and asked that 500 show up to protest the decision. Police showed up to keep order at the protest and 1,600 protestors showed up and threw bottles and rocks at police.	Thirteen (13) injured	Personal web domain, various platforms
	July 28	**Hamburg supermarket stabbing** A man went on a stabbing rampage in a supermarket in Hamburg. He became absorbed in ISIS propaganda online (mainly videos).	One (1) dead, six (6) wounded	YouTube

Figure 1: Selected domestic terrorism attacks (or foiled plans) perpetrated with the help of radicalizing content online *(Source: German Counter Extremism Project[5])*

A Defensive Democratic Model

Philosophically, the attack in Hanau, the near success of Franco A and these other domestic terror attacks should come as a shock to scholars of Germany and other Germanophiles who revere the *defensive democratic model* that the country has boasted since the introduction of its post-WWII constitution in 1949. "Defensive democracies" ensure the protection of the state by limiting certain rights and freedoms. Gross (2003) summarizes the sentiment of defensive democracies in writing, "Democracy places human rights at the heart of its philosophy, however, without security there can be no rights" (122). Things like criminalizing Holocaust denial are

[5]"Germany: Extremism and Terrorism." Counter Extremism Project. https://www.counterextremism.com/countries/germany (March 16, 2022).

hallmarks of defensive democracies. On the other hand, liberal democracies like the United States would more affirmatively enshrine one's right to their own ideas and set the bar very high before it allows a necessary government intervention.[6] Liberal democracies do this in the interest of having the free marketplace of speech dilute extremism and unwanted content in civil discourse. The bar is lower in defensive democracies, like Germany. Defensive democracies prioritize rooting out that speech rather than trusting the free marketplace of ideas to dilute it.

In Germany, this defensive democracy is on display constitutionally and institutionally in addition, to statutorily like mentioned with the criminalization of Holocaust denials. Article 9 of the German Constitution, in contrast to the First Amendment in the U.S. Constitution, codifies the idea that freedom of assembly can be restricted from people whose aim is hostility towards the Constitution. Even more expansively, Article 18 allows the government to suspend any previously aforementioned "basic rights" (Articles 5, 8, 9, 10, 14, or 16) like freedom of speech and of the press if a German is found to be using those rights to advance disorder against the state or the Constitution. Pursuant to these ends, the government has state institutions like the Office for the Protection of the Constitution ("Bundesamt für Verfassungsschutz", or BfV) whose task is similar to that of the U.S. F.B.I and specializes in surveilling for instances of attacks against the constitution and the national cohesion it provides. Contrary to the mechanisms used for maintaining constitutional hegemony in the U.S. and other liberal democracies, Germany and other defensive democracies operate uniquely by preemptively curtailing the rights of individuals or groups that wish to disrupt the democratic order.

[6] For example, see *Brandenburg v. Ohio* (1969) ("The constitutional guarantees of free speech and free press do not permit a State to forbid or proscribe advocacy of the use of force or of law violation except where such advocacy is directed to inciting or producing *imminent* lawless action and *is likely* to incite or produce such action.")

Germany's Preemptive Regulatory Regime

Parallel to this defensive democratic model is the regulatory regime Germany uses to govern online intermediaries, or sites that host the speech of users online. This European, and especially German, regulatory structure is what I define as the *preemptive framework*; such is defined in juxtaposition to what I will define as the *deregulatory framework* that exists in the United States to govern the same sites. This chapter will explore the European *preemptive framework* using Germany as a case study and argue that haphazard oversight and vague statutory language created a perverse incentive to remove counterspeech. To do so, this chapter will critically analyze the major pieces of legislation enacted to address speech hosted by intermediaries. Throughout such analysis, attention will be paid to the ways each iteration's language or implementation could be co-opted to unintentionally allow objectionable and radicalizing content to reach its intended audience and simultaneously dilute the stock of legitimate information that could refute it.

The legislative line will start in 1997 with the Tele Services Act and point out what the European conceptualization of the internet looked like at its infancy. After that, the European Union's E-Commerce Directive (ECD) of 2000 was passed in an attempt to harmonize national attempts at regulating the internet, but I'll argue that both it's language and implementation actually muddied the waters from an early point of regulating this transnational industry. The ECD would be implemented by Germany through the Tele Media Act of 2007, substantially updated with the NetzDG of 2017 and then tweaked with a recent 2021 amendment. In the timeline, the ECD was the pivotal point and the focus in examining the post-ECD legislation is to point out how some language would get updated but that principles that underscored the Directive continue to plague future iterations of intermediary regulation that stem from it.

1997	Teledienstegesetz (The "Tele Services Act")	TSA
2000	Directive 2000/31/EC (The "E-Commerce Directive")	ECD
2002	*Deadline for Member States to transpose the ECD*	
2007	Telemediengesetz (The "Tele Media Act")	TMA
2017	Netzwerkdurchsetzungsgesetz	NetzDG

Figure 2: Acronyms and dates associated with the below discussion of the German intermediary liability regime over time

The Tele Services Act of 1997 (Teledienstegesetz)

The Tele Services Act, or TSA, was successful in showing Germans that traditional media roles *could* map to the internet, but was not particularly successful in showing them specifically how. As Professor Determann writes in 1999, Germans and the ISPs (internet service providers[7]) around the world that hosted content for them anxiously awaited how Germany would adapt their 1996 Telecommunications Act (Telekommunikationsgesetz) and earlier pieces of legislation to the internet. Not only did Germany have to respond to the monumentally-important and recently passed U.S. Communications Decency Act of 1996 but also international outrage over the case of Felix Somm.

Case of Felix Somm

Felix Somm was the managing director at CompuServe Deutschland (CSD), a subsidiary of the U.S. company, CompuServe (CS). CS was a popular American site that "newsgroups" could rent space from and publish their stories and content. CS maintained the contracts with

[7] ISSs (Information Society Services) was the original terminology in European regulation. For the sake of clarity when comparing U.S. and European regulation, I will be referring to ISSs as ISPs going forward. The terms encompass the same category of internet publishers but ISP is the term that has become more widely accepted in the modern day.

these newsgroups and handled all of the logistics of such an arrangement to get the content online and edit it at the bequest of their newsgroups. CSD merely operated the telephone line that transmitted the data from CS in America to Germans that were requesting newsgroup content. In December 1995, prosecutors in Munich handed CSD a list of 282 newsgroups that they alleged were publishing unlawful material (child pornography and extreme violence). Felix Somm told his counterpart at CS who then blocked access to these sites. CS didn't find any evidence of child pornography or violence, but blocked access to the 282 groups at the request of Munich prosecutors. American consumers were up in arms; they poured German beer into the sewers in front of CompuServe's San Francisco headquarters and threatened to terminate their membership to the service.[8] In response, CS unblocked all but five of them and implemented "Cyber Patrol" to prevent German minors from being able to see the 277 newly-unblocked groups. The German prosecutors brought charges against Somm because child pornography was still illegal in Germany even for adults, there was a likelihood that the content had been cached and republished by other newsgroups despite the blocking, and, they alleged, as executive of CSD it was his negligence that made him an accomplice. He was tried and sentenced to two years of probation.

Tele Services Act as a Remedy

In 1997, the Bundestag passed the Tele Services Act in an attempt to clarify the legal environment, as their documents reveal, to create plentiful and "long lasting jobs" with more companies staying in Germany. Clarifying the liability for online intermediaries, to them, would prevent another situation like what happened to Mr. Somm and CSD. They made sure, also in an

[8] Kossel, Axel. 1997. "CompuServe Reagiert Auf Porno-Anklage (CompuServe Responds to Porn Accusations)." *Heise Online*. https://www.heise.de/newsticker/meldung/CompuServe-reagiert-auf-Porno-Anklage-9112.html (March 15, 2022).

apparent nod to the case, to distinguish this law from their Telecommunications Act and the State Broadcasting Treaty due to the interactive nature of this new medium. Somm's liability if he was, for example, the managing director at a news station would be different than if he was the managing director at CSD. At a news station, he can cease distribution of illegal material upon receipt of its illegality. But at CSD, the content he made available was not being requested or directed by him nor was it under his sole control; anyone could request it at any time even if he didn't want them to and the material could be copied ad infinitum. The legislature recognized this in affording him and CSD more protections.

Section 5 of the TSA codifies this fact. ISPs that "make available" content are more liable than those who strictly "provide access." In the scope of this case, CS "makes available" content and CSD strictly "provides access." CS was the only one that could shut down access at the root and CSD was merely a pipeline facilitating the access. Upon appeal, Somm was acquitted under the retroactive provisions of the TSA and his ISP (CompuServe Deutschland) was immunized against the content of the users to which they "provide access." If CS was based in Germany and Somm was the leader, as Professor Determann explains, it would have been a different story. Providers that "make available" content are immunized under the TSA from civil liability as long as they don't have knowledge of the content, are technically unable to remove it, or cannot be "reasonably expected" to block the use of certain content. The latter two are seemingly a final nod to the Somm case where CSD argued at trial that they neither could remove nor block access to the content because of caching and CS-ownership of the content, respectively. The case gets more complex in such a thought experiment where CS is based in Germany and CSD didn't exist, but the central premise here remains: the TSA recognized the fact that the internet was

different from other previously governed mediums like broadcasts and therefore necessitated room to grow and less liability.

Although the TSA did provide value in clarifying the application of traditional media roles to this new frontier, it still represents a window into Europe's early conceptualization of the internet. First, the act defines teleservices as any communication services that uses "combinations of characters, images, or sounds" – certainly not the most nuanced or scalable definition. In defining "teleservices," the act conceptualized the primary examples of such to be "tele banking" and "data exchange;" the first of which is comically representative of the portmanteauing that characterized nascent internet knowledge and the second of which referred to sharing raw data back and forth. Along the same vein, the act only defined two actors: users and providers. At this point, the internet was understood as a linear transaction where users ask to see something on their screen and providers provided it. This act serves as a primary source and a screenshot of the early understanding of the internet; the German citizens that made up the Bundestag had limited experience with the internet and their understanding of its scope reflected as such.

Tele Services Act as a Primary Source

In addition to this act being a primary source for early internet naivete, the TSA also stands to represent the starting point for intermediary liability in Germany. An important part of this act is how it reflects the starting point for future debates about hosting, caching, and, most relevantly, being a "mere conduit." These are the three legal classifications that would develop for ISPs. ISPs are considered hosts if their information is stored permanently and they strictly provide the space for it to reside. ISPs are considered eligible for caching safe harbors if it logs

user transmissions for a finite, but not insignificant amount of time in the interest of making things like loading speeds faster. “Mere conduit” sites are most relevant to the discussion of radicalizing content online. “Mere conduits” are ISPs that strictly transmit the information provided by users to facilitate the transaction to another user or users. This represents a linear scale of not only ISP capability but more importantly ISP liability under this regime.

It’s easier for a hosting site to lose its liability than a site operating as a “mere conduit” because of the level of heightened involvement hosts have in amplifying the content. Modern-day ISPs are tough to lock into singular categories because of their diversification, but a simplified way to consider these distinctions is that sites like Whatsapp, Instagram, Facebook, Twitter, and Reddit are predominantly “mere conduits” or hosts (depending on the functionality of the communication) whereas services like Amazon Web Services could be classified as predominantly a caching service. The TSA doesn’t make much of a distinction between hosts and caching services (even though the two have become more delineated and clearly divided today) but does distinguish hosts/caching from “mere conduit” ISPs in its use of “provide access” and “make available,” respectively.

This first iteration of Germany’s intermediary regulation gave blanket immunity to mere conduits (Section 5.3) but gave caveats to the immunity of hosts/caching services (Section 5.2). But even then, the caveats for host/caching services are generous: “Providers shall not be responsible for any third-party content which they make available for use unless they have knowledge of such content and are technically able and can reasonably be expected to block the use of such content.” To escape liability under this early act, ISPs like CS (if they were based in Germany) would just have to prove that (1) they didn’t know it existed or (2) if they knew the objectionable content existed that they couldn’t technically block it or be expected to block it. If

the Somm case was initiated after the passage of the TSA, the prosecutors couldn't allege that the content was going to be copied ad infinitum by other newsgroups as a reason to hold Somm negligent; such an argument would actually give Somm and CSD an easy case to prove that moderating this specific content was not possible with the technology available to them at the time and therefore should be instead immunized.

As it should be abundantly clear, these laws and the act of defining stakeholders, enumerating their liabilities and choosing their exemptions within those laws becomes very complex. To boot, this process was, at the same time, playing out in fourteen other EU member states. Granted, there was likely some level of coordination between certain countries on what language to use and that might work well for the regulation of crop production or fuel-efficiency standards, but the internet is different. The internet is a constantly evolving, simultaneously occurring, instantaneously arriving, decentralized, transnational industry. The European Union saw the writing on the wall that such an industry would be economically stifled if forced to abide by potentially fifteen different regulatory regimes. The European Union took action (although not in a way that retrospectively was good for the internet).

E-Commerce Directive of 2000 (2000/31/EC)

Shortly after the passage of the TSA, the European Commission proposed to the European Council and Parliament a directive to harmonize "certain legal aspects of electronic commerce in the internal market."[9] Recital 40 is of particular note in proving the oxymoronic nature of this endeavor. Recital 40 of the E-Commerce Directive, or Recital 16 as included in the original proposal, articulates the mutually beneficial nature of rectifying a balkanized legal environment for both member states and ISPs. Through their recitals and legislative journals, the

[9] OJ C 30/4, 5.2.1999, p. 1

European Commission, Parliament and Council all repeatedly emphasize the importance of transnational legislation. They cite different reasons throughout (ie: economic gain, citizen protection, effective member state prosecution, international competitiveness, etc.) but all make note that there is a need for legislation that brings together "existing and emerging disparities in Member States' legislation and case-law."[10] Therein lies the puzzle: if unifying legal regimes was the policy priority, what was the utility in passing this governance as a *directive* rather than a *regulation*?

Directives leave the implementation of codified policy goals up to the member states whereas *regulations* directly implement binding legislation that must be enacted in their entirety by the member state. Both of which have value, but different value. It's in this choice to pass such governance as a *directive* that I argue the European Union set itself up for decades of ineffective content oversight. By opting for a directive, the European Union allowed member states to continue the asymmetric regulation between member states the governance sought to address. First, this chapter will explain the provisions of the E-Commerce Directive as it was passed and then the incongruities between it and future German national legislation. By tracing the iterations of Germany's national legislation which was passed to implement this directive, it becomes clear that the directive empowered the omission of both procedural specificity and governance structure that allowed many loopholes and vague language to arise. These textual missteps made it difficult to systematically enforce content standards.

The E-Commerce Directive has more than 60 recitals (contextual statements about the legislative intent) and 23 articles but the majority of which govern the logistics surrounding its four major, substantive articles (Article 12-15). Articles 12-14 govern the three classifications of intermediaries and Article 15 prohibits member states from requiring ISPs to actively monitor for

[10] *See* Council Directive 2000/31, 2000 O.J. (L 178), recital 40 [hereinafter *Directive*]

objectionable content. The impact of my larger argument about objectionable content relates to the accessibility of radicalizing content on social media ISPs, and not so much the impact these laws have on, for example, government action against ISPs (like Amazon Web Services) that domestic terror cells may use to share documents internally. Because of this, Articles 12, 14, and 15 of are particular importance while Article 13 about caching does not bear much relevance to the argument about radicalizing content on social media platforms

Article 12 of the E-Commerce Directive addresses the liability standards that should be imposed by member states on ISPs that act as "mere conduits," like Whatsapp or Telegram. Again, this is where information is provided by a user and not stored by the ISP in any additional way or for longer than directed by the user to facilitate that communication. It specifies that ISPs are not liable for the user content that they facilitate the transmission of so long as the ISP doesn't initiate the transmission, select the receiver of the transmission, or modify the transmission in form or content.[11] Content posted on "hosts", like Facebook or Reddit, have a larger reach and the content is visible for longer, so Article 14 imposes a more stringent standard of liability on hosts. Article 14 establishes guidelines for member states to govern sites like Facebook, Reddit, 4chan, Twitter, Parler and Gab – defined collectively as "hosts." Hosts are immune from civil liability so long as they (1) did not have "actual knowledge" of the illegal material or (2) failed to remove the material "expeditiously" after receiving notice[12]. The first part of which naturally raises two central questions of internet content moderation. First, if a site wants to remain immune, can it simply not open its eyes in search of content violations? Second, what defines "actual knowledge?" If the content violation exists on its site and the ISP is the omnipotent, omniscient, omnipresent force overseeing its users then it hypothetically could have

[11] *Directive*, Article 12 (1)
[12] *Directive*, Article 14 (1)

knowledge of any piece of content at a given time. This is precisely where Article 15 clears the air. In the interests of fostering the growth of ISPs and reducing their administrative burden, Article 15 puts the onus of content moderation on users and third-parties to report. Article 15 (1) enumerates, "Member States shall not impose a general obligation on providers... to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity."[13] As mentioned, the Directive was implemented to both harmonize legal regimes for the growth of the ISP sector and also prevent the proliferation of objectionable material. These immunities, or "safe harbors" as they are also called, insufficiently clarified the legal environment for ISPs and failed to offer codified protections for the consumer in the form of content moderation guidelines.

How was the E-Commerce Directive toothless from the start?

Both the message and the medium of the E-Commerce Directive would later present themselves as obstacles to responsible content moderation and the protection of consumers. As it relates to the message (textual failures), I argue that the text of the Directive offers a variety of protections to ISPs but lacks three reciprocal mechanisms established in the U.S. DMCA that would protect consumers. As it relates to the medium (enactment failures) of the Directive, I argue that its passage as a "directive" rather than a "regulation" left the crucial provisions that *were* enumerated up to member states to implement but that those provisions were not ones that should have been left up to member states given the transnational nature of the internet.

[13] *Directive*, Article 15 (1)

Textual Failures

The U.S. Digital Millennium Copyright Act was passed in 1998 to map analogue copyright jurisprudence to the digital age. During the initial drafting of the E-Commerce Directive by the European Commission, legislators took inspiration from many of the safe harbors and principles enshrined in this American law (Martinet and Oertli 2015). The DMCA only applies to ISPs harboring copyright infringements and not other types of objectionable content whose safe harbors are instead addressed in the Communications Decency Act (CDA) of 1996 (Baistrocchi 2002). The DMCA represents an aggressive, trademark European, regulatory regime that distinguishes it from the CDA. We would expect that Europe's adoption of such DMCA principles in their ECD would yield equally smooth results as has been seen with the DMCA. However, the ECD only took some, but not all of the DMCA's provisions. The statutory language of Articles 12 and 14 protecting ISPs are eerily similar to the DMCA but it chose not to adopt three crucial consumer protection aspects of the DMCA. Notably, the ECD lacks (1) a clear monitoring rule, (2) a "put back" procedure, and (3) a punishment for submitting fraudulent reports ("heckler's vetoes"). Each of these is at the root of future abuse by malicious, radical internet users and the future constraints upon law enforcement to stop them. These three would be passed on repeatedly through future iterations of German ISP legislation and continue to materialize as real life loopholes to content moderation efforts.

The DMCA makes clear that there is no obligation for sites to monitor, or actively seek out, copyright infringements.[14] But the ECD tried to have its cake and eat it too; sites are both told that there is no obligation to monitor in Article 15, but that they could also be held to a "[duty] of care, which can reasonably be expected from them and which are specified by national

[14] 17 U.S. Code § 512 (m) (1)

law, in order to detect and prevent certain types of illegal activities."[15] There is a vague line that European ISPs have to tow between whether they are required to constantly monitor for illegal material under their "duty of care" or are immune under Article 15. This issue came to prominence in the infamous case *Delfi AS v. Estonia* (2015). Delfi was a major, informative ISP in Estonia whose business model thrived off the comments at the bottom of news postings by pseudonymous users. At the bottom of a 2006 story about ice road commerce and a Estonian ferry company, users began to post threatening and offensive comments targeting the majority shareholder of the ferry company, identified as "L" in court documents. The comments alleged violence and were anti-Semetic. L demanded the comments come down and compensatory damages of 32,000 euros. Delfi took down the comments but refused to pay the damages. After a lengthy court battle, the case made it's way out of the Estonian appeals system, out of the Estonian Supreme Court, out of the European Court of Human Rights, out of the European Chamber, and into the 17-member European Grand Chamber. The Grand Chamber ruled that Delfi should have "been able to anticipate the potential liability that arises from operating an open commenting system" and that if they wanted to solicit comments to their postings that they needed a faster way of screening and deleting hateful, incendiary comments. The dissenting two justices note that this causes what they called "collateral censorship" (Kosseff 2019, 150). Sites are incentivized to lean on the side of caution and proactively remove anything with certain keywords in the hope of not missing a comment or a posting.

As a result of Delfi's repeated defeat in the courts and their cost-aversion to hiring a human team to do this, sites like Delfi have implemented digital screening tools that are either artificially intelligent or run on keywords. There are numerous examples of Facebook gardening

[15] *Directive*, Recital 48

groups being shut down for sharing usage with a certain sexually suggestive term[16], Indigenous accounts (ie: Shane Creepingbear) being shut down under suspicion of being a fake bot[17], and legitimate news reporting on certain topics being hidden for using certain keywords in reporting.[18] Amidst all this mainstream content being restricted, radicalizing "fake news" has co-opted "meme culture" and embraced euphemisms (similar to "Let's Go Brandon!" in the U.S.) to subvert such constant monitoring attempts. This is all to say that a textual contradiction has created the legal uncertainty that forces ISPs to overly police their platforms. Their over policing has proven ineffective and has, most importantly, led to legitimate news sources being restricted while fake news constantly adapts and flourish.

In addition to this inversely Darwinistic effect that the constant monitoring provisions have had on the information reaching consumers, well-intentioned user speech is not given recourse once being taken down. The DMCA spells out a clear "counter notification" procedure and timeline in which a user or news outlet can get their content put back up after an automatic or user-reported complaint.[19] The ECD adopted the DMCA regulatory regime that left open the possibility of "heckler's vetos" but doesn't adopt the DMCA provision that could fight them. Heckler's vetoes are when people fraudulently report content in the hope that an ISP will err on the side of instant take down. "However, whereas in the physical world one needs a court order to stop the distribution of an information product which a court will issue only if it feels the complaint (is likely to be) justified, in the virtual world it will be enough to merely claim

[16] Ortutay, Barbara. 2021. "Hoe No! Facebook Snafu Spells Trouble for Gardening Group." AP NEWS. https://apnews.com/article/lifestyle-technology-oddities-business-gardening-9c9f431f91ba450537974758de4f14d2 (March 24, 2022).
[17] Bowman, John. 2015. "Facebook Flags Aboriginal Names as Not 'Authentic'." CBCnews. https://www.cbc.ca/news/trending/facebook-flags-aboriginal-names-as-not-authentic-1.2970993 (March 24, 2022).
[18] Koetsier, John. 2021. "Facebook Deleting Coronavirus Posts, Leading to Charges of Censorship." Forbes. https://www.forbes.com/sites/johnkoetsier/2020/03/17/facebook-deleting-coronavirus-posts-leading-to-charges-of-censorship/?sh=277dea2c5962 (March 24, 2022).
[19] 17 U.S. Code § 512 (g) (2-3)

material is unlawful to stop it from being disseminated" (Julia-Barcelo and Koelman 2000, 234). Furthermore, Julia-Barcelo and Koelman (2000) point out that most ISPs have terms and conditions pages that indemnifies them from legal action against users for the removal of their content. This lack of a "put back" procedure in the text of the ECD exacerbates the inequalities created by monitoring provisions.

The final textual issue of the ECD relates to its lack of a punishment for submitting fraudulent claims (or "heckler's vetos"). The DMCA explains that any one who "knowingly materially misrepresents" a claim against an ISP is liable for the damages incurred as a result of the take down and any attorney's fees that the ISP has to pay to seek out those damages.[20] In addition, the DMCA requires that all claims against a piece of content must be made under penalty of perjury.[21] The ECD by comparison lacks both such penalties and any imposition of such a penalty is left up to member states. However, the liability laws in most member states operate on a fault-based principle where there is a very high bar to prove that the malicious actor *knowingly* misrepresented their claim against an ISP (Baistrocchi 2002). Therefore, those looking to radicalize (by diluting the pool of legitimate information) can do so with assured impunity, assuming the ISP or member state can even track them down in the first place.

Enactment Failures

The choice to enact this governance as a "directive" rather than a "regulation" is the second Achilles' heel of the ECD. Most, if not all, EU legislation seeks to harmonize the laws of its member countries to streamline its union. Little scholarly research can conclusively explain why this was passed as a directive, but textual analysis of primary sources reveals that it stems

[20] 17 U.S. Code § 512 (f)
[21] 17 U.S. Code § 512 (c) (3)

from a 1990s/2000s conceptualization of the internet as an intranational rather than transnational industry. As evidenced by a December 2000 article in The Daily Mail by James Chapman, users were turned off by the low functionality and high access costs of using the internet. The article cites experts and internet trade associations who concur that "predictions that the Internet would revolutionize the way society works has proved wildly inaccurate."[22] At the time this legislation was being debated and drafted (1998-2000), the interactive web had seemingly lost the luster it had when first unveiled in the early-90s. In addition to this limited view of the internet being represented in public sentiment, it was also represented in the legislation that lawmakers at the time drafted. In the 1997 Tele Services Act, Germany lawmakers expressly exempted all "teleservices" from having to register or get a license.[23] Flash forward ten years and the Tele Media Act changed that to expressly requiring ISPs register with the government.[24] All of this is to say that the internet was conceptualized as an industry that had potential to grow but one that would grow within its own borders, like national broadcasting or data protection.

The Television without Frontiers Directive (1989) harmonized the national laws of the member nations to allow for more streamlined transmission of "European works" but still emphasized the importance of national legislation to promote cultural programming and the growth of each member state's own broadcasting industry (Kightlinger 2003, 733). The 1995 Data Protection Directive follows a similar analogy. The European Data Protection Supervisor puts it best in their overview of the history of the GDPR: "Over the last 25 years, technology has transformed our lives in ways nobody could have imagined so a review of the rules was needed. In 2016, the EU adopted the General Data Protection Regulation (GDPR), one of its greatest

[22] Chapman, J. (2000, December 5). Internet 'may be just a passing fad as millions give up on it'. *The Daily Mail*, p. 33.
[23] Tele Services Act (1997), Section 5
[24] Tele Media Act (2007), Section 5 (1)

achievements in recent years. It replaces the 1995 Data Protection Directive which was adopted at a time when the internet was in its infancy."[25] As Hoeren astutely points out in 2000, even if transnational legislation was desired, it would remain a "mere utopia" (113). As much of a visionary as he was in that early article forecasting artificial intelligence and a more rapid digital growth than his contemporaries, he still recognizes the limitations of international cooperation (or *"wissensordnung"* as he defines it) is limited. The internet in the 1990s was still restricted to national infrastructures of telephone cable to transmit information; more cost-effective and deterritorializing satellite and wireless connection wouldn't link the internet infrastructure of different countries until after the turn of the century. This mainstream conceptualization of the internet as geographically limited likely inspired the European Commission in 2000, in the same way it did in 1989 and 1995, to take the least intrusive approach to governing and empower individual member states to legislate in the name of sovereignty.

This choice to legislate the digital world under minimally intrusive "directives," however, comes at a cost. The eventual "deterritorialization" that would happen to the internet, as Hoeren (2000) remarkably foreshadows, would reveal three things. First, the ways that a "directive" failed to harmonize national laws. Second, the ways that the internet's original conceptualization could have handled minimal transnational communication but is unprepared to keep up with the scale at which user-generated content would be produced. Finally, the way that such lackluster harmonization has led to law enforcement's inability to reduce radicalizing content online. As proven earlier, the textual safe harbors created vague standards that forced ISPs to err on the side of systematic content removal, which had a disparate impact reducing user access to legitimate information more so than reducing user access to radicalizing material. In this section, I'll

[25] "The History of the General Data Protection Regulation." *European Data Protection Supervisor*. https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_en (March 15, 2022).

elaborate on the second failure of the ECD: the way that its enactment as a directive further balkanized the legal landscape for ISPs and restricted transnational and non-EU cooperation on removing objectionable or radicalizing content.

The EU sought to harmonize jurisdiction over content moderation using a standard called the "country of origin" principle. The "country of origin" principle is not per se in the directive itself but is the scholarly way of referring to Recital 22. Recital 22, in summary, articulates that jurisdiction over an ISP's content is left up to the member state in which the ISP is domiciled. Twitter (twitter.com) is legally domiciled at its headquarters in Ireland while a site like Linksunten (www.linksunten.indymedia.org) is domiciled in Germany; therefore, objectionable and radicalizing content on Linksuten is held to the national laws of Germany that derive from the ECD while Twitter to the national laws of Ireland that are also derived from the ECD. The issue here is that the internet is accessible in all member states, but only one member state can regulate its material. The ECD states, "Member States may not… restrict the freedom to provide information society services from another Member State."[26] Therefore, the governance of objectionable content was left up to each member state whose various incentives to over- or under-police their domiciled ISPs leads to inconsistencies once the internet became more transnational and people in Poland could, for example, view content hosted by ISPs domiciled in Greece.

Pretend a Greek ISP was hosting user-posted content that glorified the Holocaust and such content was visible to citizens of Poland. If a Polish citizen reported such content, Poland is helpless to take action against this material. This being the case, the E-Commerce Directive sets out an override procedure. The reporting state must contact the state in which the ISP is domiciled and ask for action to be taken. If the glorification of the Holocaust is not a codified

[26] *Directive*, Article 3 (2)

legal offense in Greece, Greece would hardly be able to address their complaint (Kightlinger 2003). If Greece was unable to address the reported content or addressed it "inadequately," then Poland could notify the European Commission and the member state that it was invoking the Article 3 (4) override provision. Only after a review of the alleged violation, could the European Commission grant or deny the reporting state's request to take action against an ISP domiciled in another member state. For a request to be granted, the reporting state must prove the request is necessary to: "(1) public policy, in particular the prevention, investigation, detection and prosecution of criminal offences, including the protection of minors and the fight against any incitement to hatred on grounds of race, sex, religion or nationality, and violations of human dignity concerning individual persons, (2) the protection of public health, (3) public security, including the safeguarding of national security and defence, [or] (4) the protection of consumers, including investors."[27]

Such an override provision would seem to benefit international coordination on content moderation. However, a lack of underlying harmonization among national laws created high transaction costs for member states to engage with other member states. As Kightlinger (2003) notes, "Rather than giving detailed instructions to the Member States on how to harmonize national laws affecting the Internet, the E-Commerce Directive establishes a "country of origin" rule… not surprisingly, the Member States were hesitant about adopting a 'country of origin' rule covering areas of the law that had not been harmonized" (728, 733). The country of origin rule and the override procedure would work seamlessly if every member state's criminal statutes were the same. But many of the national laws that governed the internet differed from country to country. Things like the standard of proof for defamation, the age of consent, and the treatment of inflammatory free expression differed across member states and complicated the process for

[27] *Directive*, Article 3(4)(a)i

removing content whose illegality was not unanimously agreed to across member states. Because of this lack of harmonization, member states were required to consult the Commission and wait to receive a ruling before proceeding.

Such rulings, however, come as a last resort and are time consuming. Mark Kightlinger (2003) continues that his experience with the officials in the European Commission lead him to believe there "would be an informal effort to persuade" the reporting country not to require the country of origin to implement a blanket injunction against the content (740). Instead, Commission officials might encourage the reporting country to just block access to the websites that their citizens can see rather than forcing the country of origin to revoke access from everyone in the EU, even those outside the reporting country. The Commission seems to prefer the least invasive and precedent setting action whenever possible. The lack of underlying harmonization scared member states when it was proposed and was correctly foreshadowed to indeed plague member states with high transaction costs associated with international cooperation.

Transposition of the E-Commerce Directive

As has been shown, the naively limited conceptualization of the internet led to the creation of a European directive in 2000 that was riddled with textual and implementation-based potholes that would lay a strong foundation for the haphazard regulation of intermediaries that member states would later implement for the next two decades. Contrary to a regulation whose language superimposes the existing statutes of member states, directives need to be implemented in a timely manner by the proper legislative channels of each member state. The ECD stipulates

a January 17, 2002 deadline for national legislatures to "transpose" the directive into national legislation.[28]

According to Eur-Lex, Germany reported to the European Union on December 20, 2001 a bill (called the Elektronischer Geschäftsverkehr-Gesetz, Electronic Commerce Act, or EEG) that supposedly implemented the majority of the E-Commerce Directive. That being said, there is interestingly no such record of the EEG existing in the German federal registry of legislation. Academics seem split; a minority of which label the 2001 Electronic Commerce Act as the official transposition of the ECD and an overwhelming majority of academics labeling instead the 2007 Tele Media Act as such. The two acts are identical but the text of the EEG that was submitted to the European Union appears to be instructions for amending the 1997 Tele Services Act whereas the 2007 Tele Media Act is formatted as a finalized piece of legislation. What this implies is that Germany submitted their agreed upon intent to amend the 1997 bill as their amendment to the 1997 bill that would have satisfied their deadline. In reality, the EEG wouldn't become law until it was passed six years later as the Tele Media Act of 2007. While the bills are identical, this is important because it elucidates the lack of oversight that exists between the point at which a directive is passed by the European Union and the point at which it becomes national law.

The European Union has struggled with issues of transpositional oversight for years. Every year, it ranks its member nations on overall performance, transpositional deficit, and conformity deficit. Transpositional deficit refers to the number of months late a member state transposes a directive and conformity deficit refers to the percentage of directive language that the member state incorrectly transposed. The July 2004 and 2007 rankings note two unfortunate

[28] *Directive*, Article 22 (1)

realities. The 2004 rankings[29] judges directives whose deadlines were between March of 2001 and March 2002, like the ECD. It notes that, at the time when it was supposed to transcribe the ECD, Germany had the second highest number of outstanding directives to transcribe. The 2007 rankings[30] evaluation of conformity deficit judges directives, like the ECD, that were transcribed between March 2006 and March 2007. This report notes that even though Germany made some of the largest strides to finally clear their outstanding directives that the country still "continue[s] to accumulate a large number of infringement proceedings" related to inaccurate transposition (p. 7). Germany had the fifth highest number of open infringement proceedings (82) and took the fifth longest to resolve each one (average of 27 months). All of this is to say that transpositional delay is an issue in many EU countries, but that Germany was a particularly troubled offender in the early to mid-2000s. The inaccurate transposition of this E-Commerce Directive and lack of oversight over such transposition will continue to cause legal uncertainty and therefore hinder the ability of ISPs to formulate content moderation guidelines.

Tele Media Act of 2007 ("Telemediengesetz")

The institutional problems facing content moderation in Germany start with the haphazard textual drafting of the ECD and the improper implementation of this governance as a directive rather than a regulation. But such problems are compounded by the fact that the transposition of this directive into the Tele Media Act of 2007 was equally careless. With one 32-word sentence in Section 7 of the Tele Media Act, the Bundestag seemingly negated all the safe harbors offered to ISPs by the E-Commerce Directive. These safe harbors were the last

[29] European Commission European Union. https://ec.europa.eu/internal_market/score/docs/score13/score13-printed_en.pdf (March 15, 2022).
[30] European Commission European Union. https://ec.europa.eu/internal_market/score/docs/score16_en.pdf (March 16, 2022).

shred of protection ISPs could use to keep content up in the fight against those looking to dilute legitimate information. In what appears to be an overly expansive reading of the "duty of care" provision in Recital 48 of the ECD and a willingful ignorance to the actual legislative intent of the recital, German legislators effectively created a new type of liability that increased the likelihood of heckler's vetoes.

As mentioned earlier with the case of *Delfi AS v. Estonia* (2015), the ECD tried to have its cake and eat it too. Article 15 prohibits ISPs from being required to constantly monitor, but Recital 48 says that "duties of care" exist where ISPS can be reasonably expected to "detect and prevent certain types of illegal activities."[31] Delfi believed it was protected against constantly monitoring its comment section, but Estonia argued (and the Court ruled) it was reasonable to expect a company whose business model revolves around sollicing comments to constantly monitor the comment section. This legal uncertainty has now been codified outside the legislative intent (recitals) and instead in the textual language of the Tele Media Act. Section 7 (2) of the TMA reads: "Service providers within the meaning of Sections 8 to 10 are not required to monitor the information transmitted or stored by them or to search for circumstances indicating an illegal activity. This shall be without prejudice to obligations to remove or disable access to information under general legislation, even where the service provider does not bear responsibility pursuant to Sections 8 to 10." The first sentence is the transposition of Article 15 and the second the apparent codification of Recital 48.

Recitals are written to convey the legislative intent of the operative language in the body of a bill. They are not legally binding and are, therefore, usually free from political disagreement. Most importantly, "recitals cannot overrule a relevant operative provision: if they are

[31] *Directive*, Recital 48

irredeemably inconsistent then the text of the operative provision will take precedence."[32] Recital 48 was written as a way to keep ISPs accountable to the quality of their sites. The "duty of care" was purposefully vague and put as a recital not to be expansively read and codified as operative language but rather to be an ace in the sleeve of national governments. That provision exists to remind ISPs that the safe harbors of Articles 12-14 are not absolute, that they cannot hide behind them, and that there is still a reasonable expectation that platforms will do their part (even if not required to) to have some kind of content quality safeguards.

With this in mind, German legislators made legally binding something that was originally intended to loosely keep ISPs accountable to the unreported content on their sites. But, in doing so, legislators put the two sentences of Section 7 (2) in seeming contradiction. One exempting ISPs from monitoring if it's a host, caching service, or mere conduit and one saying that national level "obligations to remove or block unlawful content remain unaffected" by a sites status as a host, caching service, or mere conduit (Hoeren and Yankova 2012, 508). Hoeren and Yankova label this added sentence a direct contradiction to the protections offered in the E-Commerce Directive and detail the effects of unchecked transposition as more legal uncertainty. "General, unambiguous, and binding liability standards, which the Tele Media Act and Directive intended to create, are absent… In every single case, courts observe the relevant circumstances and facts before attempting to find a fair balance between the interests of the parties involved" (Hoeren and Yankova 2012, 528). This contradiction negates any ability for precedent-establishing, legal clarity upon which ISPs could draft effective and enforceable content moderation guidelines. Sites are torn between whether they need to actively monitor and err on the side of excessive "take downs" or whether Article 15 protects them from such an assault on free speech and

[32] Thomas Reuters. "Recital (EU)." Practical Law Encyclopedia. https://content.next.westlaw.com/Glossary/PracticalLaw/I13f404e5785211e79bef99c0ee06c731?transitionType=Default&contextData=(sc.Default)&firstPage=true (March 29, 2022).

expression. In 2017, German legislators seemed to acknowledge this legal uncertainty and rectified the issue by attempting to embrace the original Article 15 intent of the ECD and rejecting the burden that Recital 48 put on ISPs.

Network Enforcement Act of 2017 (NetzDG)

In 2016, three of the largest European social media ISPs (Facebook, Google, and Twitter) met to draft a Code of Conduct which represented one of the most unified opportunities for industry input into European internet regulation. This meeting and the feedback solicited from ISPs would later be the inspiration for a 2017 bill called the Network Enforcement Act (Claussen 2018). Although controversial for its constitutionality, the NetzDG passed in late 2017 was put into effect on January 1, 2018[33]. The act exclusively targets large social media ISPs (two million or more German users) who were likely paying the largest price as a result of the legal uncertainties in the TMA. NetzDG clarified not only the content that was being targeted but established procedures and time frames for moderating such content that were encouraged in the E-Commerce Directive but completely absent in the Tele Media Act.

For the first time in German intermediary liability legislation, the NetzDG clearly defined "illegal content" using references to existing analogue criminal and civil codes codes. It created a time frame under which ISPs must operate depending on the clarity of the content in question. To avoid liability, clearly illegal content had to be removed in 24 hours and more complex content that required the advice of a judge would be granted up to 7 days.[34] Especially innovative was a self-regulatory body that was established so that ISPs could turn over questionable content in the

[33] In 2021, the NetzDG was amended, but only with minor edits. Legislatively, these represent two bills. But for the intents and purposes of my argument, I will be referring to them as one item and one collective slice of the timeline.
[34] *NetzDG*, Section 3 (2) (2-3)

pursuit of receiving consistent, unbiased, quick legal judgments.[35] Finally, there was a much clearer and more accessible reporting mechanism required on ISPs. Reporting mechanisms for users were required under the TMA, but sites were incentivized to hide them in obscure sections of the website. Hiding it meant marginally less notifications from users, less content systematically taken down, more content on their site and therefore more profits. NetzDG required that all (1) reporting mechanisms be clearly visible while viewing any type of content the ISP was supporting and (2) all users of reported content be cataloged and reported to the government. All of these upgrades represent an important burden shift away from ISPs; this was a shift away from forcing sites to constantly monitor and a "duty of care." This legislation gave users more tools to be their own police for the platforms while also offering ISPs an outside self-regulatory body to offshore content moderation decisions.

That being said, the self-regulatory body should not be considered a panacea for the European Union. The self-regulatory body is not there to take all the pressure off ISPs. There is even an enacted provision that allows the regulatory body to suspect an ISPs privileges of sending content for consideration if it deems that the ISP could have handled it in-house or is otherwise misusing the regulatory body.[36] While the burden is partially shifted away from ISPs in this respect, other areas of the 2017 regulation make it so that heckler's vetoes are still powerful tools that could be used by people to dilute legitimate information. Specifically, the 24-hour timeline is what many academics and legislators feared would worsen the systematic content removal that was incentivized under the ECD and the TMA. In the short term, as Claussen explains, the 24-hour guideline was effective in taking down content (2018). But the 24-hour guideline did nothing to long term reduce the creation of this content.

[35] *NetzDG*, Section 3 (2) (3) b
[36] *NetzDG*, Section 3 (11)

The 24-hour guideline was like playing "whack-a-mole." Radicalizing, objectionable content would pop up and ISPs would (largely indiscriminately) wack it but nothing was being done to address its underlying cause. Furthermore, academics like Claussen and others even alleged that such a short timeline for removing content upon notification was a violation of the ECD. The E-Commerce Directive doesn't give a time frame in which an intermediary must remove after notice, it just says it has to happen "expeditiously."[37] Guidelines can be more restrictive than what the ECD requires since it's a directive and not a regulation but not if it substantially changes the incentive structure or regulatory outcomes for ISPs (Claussen 2018). Article 15, again, prohibits the requirement of ISPs to constantly monitor; this 24-hour guideline to many scholars effectively enforces a *de facto* monitoring requirement to try and preempt the many requests it gets throughout the day. Furthermore, this whack-a-mole on content had long term implications on journalists seeking to post legitimate news that functions to dilute objectionable content.

In a 2020 interview[38] with the Committee to Protect Journalists, Matthias Kettemann of the Leibniz Institute for Media Research in Hamburg elaborated on this bill's deleterious effect on benevolent users. He noted that reporters whose stories report on hate speech or other crimes that are listed as keywords for automated moderation algorithms could end up on databases that further prohibit them from publishing legitimate articles. Further, he explains that COVID-19 played an exacerbating role in this. With a large proportion of human moderators being laid off or sent home, content moderation was more heavily executed by algorithms and there was virtually no ability to appeal either fraudulent flags of legitimate stories by the algorithm or

[37] *Directive*, Recital 46, Article 13-14
[38]Earp, Madeline. 2020. "Germany Revisits Influential Internet Law as Amendment Raises Privacy Implications." Committee to Protect Journalists. https://cpj.org/2020/10/germany-revisits-influential-internet-law-as-amendment-raises-privacy-implications/ (March 16, 2022).

heckler's vetoes by malicious users looking to dilute legitimate stories. In an August 2020 Facebook (now Meta) statement[39], they admit that they'll be temporarily unable to offer appeals for the foreseeable future to users who were flagged and that the majority of such flagging is coming from technological and not human moderation.

Conclusion

This chapter examined the *preemptive framework* endorsed by the European Union and all the various ways it errs on the side of removal and affords ISPs limited liability protections for the speech that they support. I argue that, with the E-Commerce Directive as the foundation, German regulation has repeatedly created opportunities for objectionable content to dilute the redemptive power of legitimate speech by leaving open the door for, if not amplifying the effects of, heckler's vetoes. Beginning with the E-Commerce Directive, such haphazard regulation began as a byproduct of (1) textually taking inspiration from the US DMCA without incorporating all of the underlying principles and (2) its implementation as a directive rather than a regulation. Both of these missteps can be traced back to the way that the European Union and its lawmakers conceptualized the internet and its potential, as limited, intranational and best regulated by constant monitoring provisions on the sites themselves. From this rocky beginning, this chapter traced the legislative line through 2021's amendments to the NetzDG and argued that its status as a directive coupled with Germany's lackluster history of transcribing directives accurately and timely created a volatile legal environment at every iteration. Starting with the Tele Media Act in 2007, the lack of oversight by the European Union on ensuring an accurate translation led to contradictory statutes regarding the obligation of ISPs to constantly monitor.

[39]Rosen, Guy. 2021. "Community Standards Enforcement Report, August 2020." Meta. https://about.fb.com/news/2020/08/community-standards-enforcement-report-aug-2020/ (March 22, 2022).

Such a “have your cake and eat it too” incongruity created a perverse incentive for ISPs to err on the side of removing any reported content. The next update to the Germany intermediary liability governance, the NetzDG, reified Germany’s ongoingly contradictory understanding of ISP monitoring provisions. Originating with the Recital 48 v. Article 15 debate in the ECD, the TMA did through its Section 7(2) contradiction what the NetzDG did through its 24-hour guideline. Although in different forms, each iteration of German intermediary liability regulation perpetuated the incentive created in the ECD to err on the side of removal (even if wrongfully removing legitimate content). Contrary to the problems of the U.S. regulatory regime which center around the free speech rights afforded to malicious actors, objectionable content is able to flourish in Europe as a result of malicious actors being able to instead silence legitimate content providers.

Chapter 3: United States

What's more emblematic of the American dream than leaving a job you hated and setting off on your own to plant your own stake in the world? At 34, Bradley Rukstales of Illinois did just that. With nothing but a refurbished, early 2000s laptop and a passion, he launched his own company, Consumer Asset Consulting Group. Outside his successful career in marketing analytics, Brad's a self-proclaimed lover of music, smoker of meats, and "dog dad" to his Maltese Yorkshire Terrier, Daisy. His community knew him well as a patron of "Homes of Hope" which helps orphanages in India, "Smile Train" which assists children with cleft palate, the Special Olympics of Illinois, microlending in South Africa, and building homes for the troops. Suffice to say, Brad and his wife, Jill, along with their two daughters are the epitome of a stereotypical, wholesome, American, suburban family. Brad's personal blog[40] remarks about his family voting their conscience and the importance of unbiased news. A now-CEO at his own "Big Tech" marketing analytics firm, Brad was no stranger to railing against the manipulation that sites employ to drum up business. Like many Americans, he deleted traditional social media applications in a self-announced "social media cleanse" as a result of the reductive and polarizing information he was being presented. In 2020, however, Brad redownloaded some of these platforms and was exposed to an internet he previously did not know. Months later, this meat-smoking, dog dad from Illinois would be caught on CCTV footage in the Capitol Building throwing metal chairs at police officers and later being detained by no less than three officers in the Capitol Visitors Center.

Rukstales was no stranger to candidates and arguments affiliated with "The Big Lie" of the 2020 U.S. Presidential Election. FEC filings show that he gave $500 to the National

[40] https://www.bradrukstales.com/

Republican Congressional Committee in 2003 but then took a fifteen-year hiatus before giving his next donation. From 2018 to 2020, Rukstales gave $26,000 to President Trump and Congressional candidates that affiliated with his ideologies within the Republican Party. But none of his conduct before this chair-slinging episode is particularly out of the ordinary. Donating to Trump is not in and of itself any way a red flag for predicting the kind of behavior Rukstales exhibited on January 6th, 2022. He had no prior criminal record, strong ties to the community, a clear self-awareness of the dangers of social media, strong financial security, and a rather politically open family (as evidenced by his eldest daughter who is known to be a very outspoken Democrat in Chicago-area politics). Even now in court appearances, he goes so far as to say he has "no excuse for [his] actions," it was the "single worst personal decision of [his] life," and that he "condemn[s] the violence and destruction that took place."[41] So what happened? How did this ordinary family man with 54-years of unimpeachable character end up on the ground of the Capitol Visitor Center?

History of Domestic Radicalization

Although Rukstales' radicalization is not solely attributable to digital factors, the role of radical, inflammatory content on social media is undeniably a part of this story. It should come as no surprise that social media sites would be ideal mediums to spread messages. Since its inception, this has been the case. As mentioned in Chapter 1, groups like the "Aryan Nation Liberty Net" in the 80s and "Stormfront" in the 90s were predominantly-online, domestic terror groups (Walther and McCoy 2021) (Conway et al. 2019). In 2009, even the research arm of the Department of Homeland Security acknowledged the incidence of terror groups on mainstream,

[41]Goudie, Chuck. 2021. "Inverness Man, Ex-Schaumburg Tech CEO Sentenced to 30 Days in Prison in US Capitol Riot Case." ABC7 Chicago. https://abc7chicago.com/us-capitol-bradley-rukstales-inverness-il-cogensia/11227769/ (March 2, 2022).

commercially accessible platforms and wrote an entire report detailing the ways it has reached a demographic of people not even looking for such content, children (Homeland Security Institute 2009). Its frequency is well established, but its efficacy is the recent revelation.

As mentioned in Chapter 1, there has been a dramatic rise (+320%) in the past (5) five years of right-wing, domestic extremism in the world and Americans are being exposed to that content almost 30% faster than they were in 2005 (Institute for Economics and Peace 2019) (Jensen et al. 2018). What January 6th has illuminated is that this social media radicalization is affecting "unaffiliated citizens" at a much higher rate than ever before. Social media radicalization has been made more accessible. No longer do you need to be affiliated with an organized hate group or other collective to be radicalized by this content. In 2021, many of those who stormed the capitol were influenced to do so by family members, friends, or other role models and did not travel with an organized group like the Oath Keepers, The Proud Boys, or the Three Percenters. From 2015-2020, "unaffiliated citizens" accounted for 52% of extremist terror, but at the January 6th insurrection that number rose to 89%.[42] This rise in unaffiliated, domestic terror comes as research finds that social media is increasingly to blame for each one. A 2018 study by the National Consortium for the Study of Terrorism and Responses to Terrorism explains that from 2005-2016, social media played a role in radicalizing and mobilizing unaffiliated terrorists in 68.12% of all domestic terror attacks reported to the PIRUS (Profiles of Individual Radicalization in the United States) database. However, in 2016 alone, that number rose to 90% (Jensen et al. 2018). Although just a sliver of a much larger body of work, this research supports the argument that the attack at the Capitol was an aberration from our historical

[42] Pape, Robert, and Keven Ruby. 2021. "The Capitol Rioters Aren't Like Other Extremists." The Atlantic. https://www.theatlantic.com/ideas/archive/2021/02/the-capitol-rioters-arent-like-other-extremists/617895/ (March 29, 2022).

conceptualization of online extremism manifesting in reality, that such an aberration is the growing trend, and that social media plays a growing role in this trend.

Year	Date	Incident	Casualties	Platforms Used
2021	January 6	**U.S. Capitol Insurrection** Armed insurrectionists organized on Gab and Parler behind the extremist idea that the 2020 Presidential Election was being stolen and that votes were not being counted. On the day when Congress was going to certify the votes, 2000-2500 Trump-aligned rioters stormed the Capitol and called for violence against Congressmen and women certifying the "fraudulent" vote.	Seven (7) dead, 150 officers injured, countless more officers reporting trauma and PTSD-like symptoms	Gab, Parler
2019		**Ned Pepper's Bar Shooting** Shooter heavily subscribed to alt-left, Antifa content on Twitter and was inspired by their violent ideologies. He walked into a bar in Dayton, OH and opened fire. Complete motive is unsure, but reports say it was likely linked to a history of mental illness and constructions of reality presented to him online.	Nine (9) dead, seventeen (17) injured	Twitter
	August 3	**El Paso Walmart Shooting** A far-right hate monger, Crusius, posted on 8chan a manifesto supporting the extremist "Great Replacement Theory" and the 2019 Christchurch mosque shootings. The shooter was reading this material extensively before posting his own manifesto online detailing the shooting and his inspiration. He would then go into the El Paso Walmart and	Twenty-three (23) dead, Twenty-three (23) wounded	8chan

		open fire allegedly targeting Latinx customers.		
2018	October 27	**Tree of Life Synagogue Shooting** Shooter opened fire on a synagogue in session in Pittsburg, PA. He was allegedly radicalized by anti-semetic content posted on Gab and created content of his own.	Eleven (11) dead, Seven (7) injured	Gab
2016	June 12	**Orlando Night Club Shooting** Anti-LGBTQ and anti-Latinx shooting that took place over three hours inside Pulse, a historically gay night club in Orlando. The suspect was allegedly active in Facebook groups and read Facebook content from ISIS that radicalized him to comit the second most deadly deadliest terrorist attack since 9/11.	Fifty (50) dead, Fifty-eight (58) injured	Facebook
2015	December 2	**San Bernardino Shooting** Inspired by international jihadi organizers on various chat rooms, a husband and wife opened fire at a holiday party and training event for municipal employees before attempting to plant a bomb and flee on foot.	Sixteen (16) dead, Twenty-four (24) injured	Various chat rooms
	June 17	**Charleston Church Shooting** A racially motivated shooting at a church in Charleston historically known as a beacon for civil rights organizing. The shooter was a known white supremicist radicalized by online forums and a manifesto.	Nine (9) dead, One (1) injured	Various chat rooms
	May 3	**Curtis Culwell Center Attack** A foiled attack on a 'Draw Muhammad' art contest in Garland, TX. The shooters were killed before they could enter the exhibit but were	Two (2) dead, One (1) injured	Twitter

		allegedly radicalized by the content of Australian jihadist Joshua Ryne Goldberg online who posted instructions about the event and gave a detailed map of how they should do it.		

Figure 3: Selected domestic terrorism attacks perpetrated with the help of radicalizing content online

American Internet Exceptionalism

A critical reader would argue that this is no different than the way the radio or television could be used to radicalize. After all, Orson Welles' 1938 radio transmission *The War of the Worlds* created mass hysteria and corrupted public information exchanges in an analogous way to the misinformation currently being spread. But the internet is inherently different. It's transnational, translatable, simultaneous, exponentially reproducing, and allows for the clandestine storage of information within its depths in a way that other mediums like newspapers, bookstores, radio broadcasts, television shows, or movies cannot. We would expect a result of this increased functionality would be a proportionately increased regulation, but such is not the case. As this chapter will explain, the internet is remarkably unregulated. This is a concept explored in Chapter 1, "internet exceptionalism," or the idea that the internet has unique growth potential and therefore should be subject to the lowest level of transaction costs and regulatory oversight. Legal theory in the U.S. makes clear this distinction between the internet and other information conduits by contrasting "publisher liability" with "distributor liability" with full immunity.

"Publisher liability" is the most stringent level of liability to which information conduits can be held. Traditionally, newspapers, movies, television shows, and radio broadcasts are held to this standard. It generally holds a medium liable for any illegal, defamatory, or objectionable content that it publishes to the same level of liability as the person who said it first. This is

because publishers have the "knowledge, opportunity, and ability to exercise editorial control over the content of its publications."[43]. Take for example the case of *Ross v. Santa Barbara News-Press* (2004). This 2004 appeal to the Supreme Court was denied certiorari but concerned a 1989 case where, essentially, Ross sued the Santa Barbara News-Press newspaper for misrepresenting the facts of an investigation about allegedly defrauding investors. The paper published a story about Ross' business partner who was sentenced to jail time and incorrectly linked Ross to the events. The newspaper lost the case and was held liable for the damages they caused to his reputation. The court held the paper has editorial content over the information it receives on a given story and therefore has an obligation to ensure the veracity (to a certain extent) of the content to which they give a platform.

"Distributor liability" is a less stringent level of liability to which newsstands, bookstores, and libraries are held. Such distributors are only held liable if they knew of something illegal in their offerings but are otherwise under no obligation to search for illegal or objectionable material otherwise. Take the case of *Smith v. California* (1959) for example. Put simply, Smith owned a bookstore that sold a book that contained indecent material and California sought to punish her for it. She sued and alleged there was no way for her to read every book and prevent violating the state statute against indecent books. The Supreme Court agreed and vacated the state law but said that she was aware the book was indecent and could be punished for selling the book again now that she's been made aware of its illegality. Distributors are only liable if they've received notification and continue to offer access to that material. This is the level of liability to

[43] Digital Media Law Project. "Immunity for Online Publishers Under the Communications Decency Act." Digital Media Law Project. https://www.dmlp.org/legal-guide/immunity-online-publishers-under-communications-decency-act (March 29, 2022).

which the letter of the law in the European Union and Germany intended to hold sites accountable.

In contrast to the treatment of American publishers or ISPs in the EU (distributor liability), the U.S. ISP industry is essentially fully immunized. The previous chapter argued that the European Union and Germany's *preemptive framework* incentivized systematic and proactive removals of content and that such actions created pathways for objectionable content to flourish in a way that would run counter to the legislative intent of the E-Commerce Directive. The U.S., on the other hand, represents the opposite approach and one that I will refer to as the *deregulatory framework*. The U.S. regulatory regime is governed by the Communications Decency Act (CDA) within the Telecommunications Act of 1996 and, specifically, a section of it that added a 230th Section to the Communications Act of 1934 (hereinafter "Section 230").[44] This chapter will explain the important legislative intent behind Section 230, explain the result of the tension that lies within it, and lay out how that tension has manifested itself in objectionable, radicalizing content for U.S. viewers. As a result of the way the Telecommunications Act of 1996 was enacted, objectionable content flourishes with neither a legal incentive to remove it nor an organized body to proactively police it.

[44] Interestingly, many commentators refer to it as "Section 230 of the Communications Decency Act." No such section exists. The statute they are referring to is 230th Section within Title 47 the U.S. Code. The "Communications Decency Act" is an instruction within the Telecommunications Act of 1996 that created a 230th Section to the Communications Act of 1934 (also known as Title 47 the U.S. Code). This is a niche technicality, but goes to show how much of a household name this controversial section has become.

Conceptualization of the Internet (Framework)	Institution for Redress	Outcome	Primary Benefit	Primary Drawback
Deregulatory	Courts	**(U.S. Status Quo)** Content is assumed to remain online unless challenged in court and ISPs are <u>not</u> incentivized to undertake their own filtering	Maximizes total information available on the internet	Maximizes amount of <u>total</u> objectionable content
Deregulatory	Governmental Body/Agency	Content is assumed to remain online but the agency can incentivize ISPs to undertake their own filtering pursuant to Section 230(c)(2)(A)	Maximizes total information available on the internet	Costly and time consuming for the government
Deregulatory	Internet Service Providers	(Not a compatible combination)	N/A	N/A
Preemptory	Courts	(Not a compatible combination)	N/A	N/A
Preemptory	Governmental Body/Agency	Agency would have to constantly monitor all content online	Minimizes amount of total objectionable content	Costly and time consuming for the government
Preemptory	Internet Service Providers	**(E.U. Status Quo)** ISP would have to constantly monitor all content online	Minimizes amount of total objectionable content	Incentivizes systematic removals of content

Figure 4: Chart of potential avenues of internet governance with the corresponding effect such would have on ISP regulation.

Legislative History of Section 230

Section 230 was originally created to vest content moderation in the hands of parents and generally to make the internet a safer place for children. With this in mind, it's ironic that the legislation would only come to be as a result of Jordan Belfort. You read that correctly; beyond

the self-described drug use, habitual infidelity, and criminal record, this subject of the 2013 film *The Wolf of Wall Street* would indeed have a direct hand in the creation of sweeping internet decency law originally intended to protect children. Jordan will have to momentarily take a backseat, however. The important legislative history of Section 230 begins with a 1991 New York District Court decision, *Cubby v. CompuServe*.

Don Fitzpatrick ran a daily newsletter, "Rumorville," that covered gossip about reporters and the general field of journalism. The newsletter was accessible on CompuServe, a hosting platform that hosted many different forums and newsletters to which CompuServe members could subscribe.[45] A competing newsletter, "Skuttlebut," was created by Robert Blanchard and his company, Cubby. Blanchard and Cubby claimed that Fitzpatrick and "Rumorville" published allegedly false and defamatory remarks about "Skuttlebut." Cubby sued Fitzpatrick, which is a separate case, and also sought to hold CompuServe liable for the speech to which it was giving a platform. The court dismissed the case in a win for CompuServe by alleging CompuServe was a distributor rather than a publisher of the information.[46] CompuServe exerted no more editorial control than a public library or bookstore and exercised, nor was expected to exercise, any editorial control over the content. Five years later, Belfort comes into the picture and he loses where Cubby won.

Prodigy was a similar site to CompuServe hosting newsletters and forums on different topics and for different affinity groups. In October 1994, someone posted on "Money Talks," a financial gossip newsletter hosted by Prodigy, that Stratton Oakmont and its president had committed financial crimes in connection with a recent business deal. In the same Southern

[45] This is a business model very popular in the late-1990s and early-2000s but faded out of popularity as forums began specializing in one type of content, vertically integrating, and having their own hosting site to save costs. But in the early days of the internet, this was a leading business model.
[46] *Cubby, Inc. v. CompuServe Inc.*, 776 F. Supp. 135 (S.D.N.Y. 1991)

District Court of New York, Stratton Oakmont, Belfort's trading firm, sought to hold Prodigy liable for the content of the anonymous user because they alleged it was defamatory and hurting their ability to do business. Stratton Oakmont noted that Prodigy was a very content-conscious forum that boasted its cleanliness and family-friendliness. Therefore he argued that they exercise editorial control that makes them not simply a "distributor" like Cubby but a "publisher" with a higher standard for content that they host. The judge agreed that Prodigy not only boasted its content moderators and the guidelines they enforced but went so far as to include it in promotional materials like in newspaper advertisements that they took out.[47] Because of this, the judge ruled in favor of Stratton Oakmont and held Prodigy liable for the speech of its anonymous user.

Lawmakers in Congress were outraged by such a result.[48] Prodigy was punished for exerting ethical content moderation while Cubby was rewarded for doing nothing. In an unlikely bipartisan effort, Representative Chris Cox (R-CA) and then-Representative (now-Senator) Ron Wyden (D-OR) got together over ice cream to draft legislation that they would add to the Senate's recently passed companion bill overhauling the internet (S652 - Telecommunications Act of 1996).[49] The legislative intent of the Cox-Wyden text was clear: the internet had the potential to blossom as an economic powerhouse, it should be allowed to grow unencumbered by Federal regulation, and that content is best left up to parents and not the government to police. Not only would they belabor this point in their floor speeches, but they went so far as to include such in the text of their bill. Section 230(b)(2) notes that, "[i]t is the policy of the United States to preserve the vibrant and competitive free market that presently exists for the Internet and other

[47] *Stratton Oakmont, Inc. v. Prodigy Services Co*., 23 Media L. Rep. 1794 (N.Y. Sup. Ct. 1995)
[48] 141 CONG. REC. H8470 (daily ed. Aug. 4, 1995) (statement of Rep. Christopher Cox).
[49] This is an unfortunately short summary of an extremely interesting story which was done for the sake of making my larger argument more cogent. The full legislative story is the topic of Chapter 3 of Kosseff (2019) and is well worth a read beyond my necessarily-reductive summary.

interactive computer services, unfettered by Federal or State regulation." Of particular interest is Rep. Wyden's remarks on the floor of the House:

> "Now what the gentleman from California [Mr. Cox] and I have proposed does stand in sharp contrast *to the work of the other body*. They seek there to try to put in place the Government rather than the private sector… In my view that approach, the approach of the other body, will essentially involve the Federal Government spending vast sums of money trying to define elusive terms that are going to lead to a flood of legal challenges." [50]

In his use of "work of the other body," Rep. Wyden is referring to the Exon amendment.

The Exon amendment was a piece of legislation that was attached to the recently-passed and aforementioned Senate companion bill. Senator James Exon (D-NE) was a grandfather concerned about his grandchildren stumbling across internet pornography online. As Kosseff notes, Exon carried around what was referred to as "The Blue Book" which contained pornographic images he alleged were freely accessible on mainstream sites of the then-nascent internet (62, 2019). While maybe his showing it to congressional colleagues in the Senate cloakroom to prove his point was a shade unnecessary, his point was certainly valid. Online pornography was legal for adults to view (see *Jacobellis v. Ohio* (1964) and *Miller v. California* (1973)), there was a growing amount of pornography online that was unsolicited and accessible to children. His original amendment would vest the power to the FCC in the same way the agency already had authority over content broadcast on television and radio. This was done pursuant to their existing authority to fine companies, stations, and actors for violating decency guidelines or doing indecent things outside of established "safe harbor hours" (10 P.M. to 6 A.M.).

[50] 104th Cong., 1st Sess., 141 Cong. Rec. Part 16, 22045-46 (August 4, 1995) (statement of Rep. Ron Wyden). (emphasis added)

There was opposition to vesting content moderation authority in the FCC, however. This came particularly from Senator Leahy (D-VT). First, he questioned the constitutionality of such an undertaking citing the Supreme Court case *Sable Communications of California, Inc. v. FCC* (1989) which had recently said that sexual expression is protected under the First Amendment and that restrictions passed in the name of protecting children cannot impede on the access of consenting adults. But more importantly and in siding with Cox and Wyden's eventual rationale, he felt that such FCC was restricted in its role as an independent agency and that it would only harm the future growth of the internet. He cited the fact that their attempt at regulating "dial-a-porn," pornographic content transmitted via traditional cable telephone, took almost 10 years of litigation to do constitutionally. Instead, he proposed handing off authority to the Department of Justice so that speech would be less infringed upon and so that actions would only be undertaken in the event of rooting out criminal activity.[51] However, neither the DOJ nor the FCC wanted jurisdiction at the time. In addition to recognizing the chilling effects that their regulation would have on the industry's growth, they both note the subjectivity of moderating internet speech and how this is a job left to the legislature.[5253]

Suffice it to say that the legislative history of Section 230 is riddled with a complicated debate about regulatory authority that would coalesce with the Cox-Wyden text eventually prevailing. The Cox-Wyden statute firmly applied the U.S. "court order standard," where everything is presumed to stay accessible online until a court says

[51] 141 Cong. Rec. S8342 (daily ed. June 14, 1995) (statement of Sen. Leahy).
[52] Letter from Kent Markus, Acting Assistant Attorney General, Department of Justice, to Sen. Patrick J. Leahy (undated), reprinted at 141 Cong. Rec. S8343-44 (daily ed. June 14, 1995)
[53] Benjamin Wittes, Interview: Reed Hundt Picks His Battles, Legal Times, Sept. 4, 1995, at 10, 12.

otherwise, to the internet.[54] This is in contrast to the *preemptive framework* which was chosen by the European Union and which would arguably have been implemented in the U.S. if the FCC was vested with initial authority.[55] Once the bill eventually got to the President, there was little media coverage or dispute about the Cox-Wyden addition. Telephone and cable companies were too concerned with other areas of the bill that more clearly affected them (Kosseff 2019). This niche debate over vesting authority over the then-nascent internet industry was of little concern to them at the time (which they would very quickly live to regret). Once signed by President Clinton on February 8, 1996, Section 230 would only be modified two times. First in 1998 to address a technical clarification about how the section interacted with a recently passed act about child safety online. The second time is in 2018 with SESTA/FOSTA, which was an act that amended Section 230 to no longer immunize ISPs that were facilitating sex trafficking. SESTA/FOSTA is a fascinating case study for another thesis argument and more so reflects the institutional failures of political participation in law drafting rather than lending credence to my argument about the contrasting frameworks. For this reason, it should be mentioned as evidence of the fact that Section 230 has otherwise gone untouched for 25-years but will not be explored in greater detail for the sake of clarity in my argument.

The problem that my argument seeks to address is how this U.S. framework has itself bolstered objectionable content. With the background of U.S. digitally-inspired

---

[54] Riley, Chris, and David Morar. 2021. "Legislative Efforts and Policy Frameworks within the Section 230 Debate." *Brookings*. https://www.brookings.edu/techstream/legislative-efforts-and-policy-frameworks-within-the-section-230-debate/ (March 29, 2022).

[55] See *Barrett v. Rosenthal*, 146 P.3d 510, 520 (Cal. 2006) (In comparing the notice-and-takedown regime under the DMCA to the liability shield of Section 230, the court held that "Congress did not intend to permit notice liability under the CDA.")

terrorism and the legislative history of its supposed solution in mind, I'll now break down what Section 230 specifically spells out (and doesn't spell out) from both a textual and administrative perspective. First, I'll argue below that the text of Section 230 allows sites to create content moderation safeguards but doesn't incentivize them to do so. Sites are under a perverse incentive to host objectionable speech in so far as they appear just marginally less objectionable than their competitors. Second, I'll argue that there exists a contradicting textual basis of FCC jurisdiction. The legislative history of the act makes it clear that Congress didn't want the FCC to have jurisdiction over the internet, and yet left open certain legal doors that allow the FCC to be granted a circular pathway to authority over the internet.

Text of Section 230

Section 230 is broken up into four major sections. The first of which (subsection (a)) is the "findings" of Congress that lay out the facts that they felt distinguished the internet from other mediums that are regulated. The second (subsection (b)) is the "policy" of Congress which spells out the legislative intent of where Congress wanted to place the onus of enforcing this text. The third (subsection (c)) gives the clear and concise immunity blanket for both content left up and good faith moderation efforts to take content down. This section directed future courts in interpreting the "court order standard" that was created by this immunity. The final section (subsection (e)) is exemptions to the blanked immunity offered to ISPs. Originally four exemptions, but after SESTA/FOSTA that became five exemptions to the protections of Section 230.

Section 230(a) offers a window into how Congress valued and conceptualized the internet in 1996. This section could be understood as evidence that Congress intended to distinguish the internet from other mediums and the liabilities they're held to (distributor or publisher liability). Of the five "findings," two are of particular relevance. Section 230(a)(2) notes that, "These services *offer users a great degree of control* over the information that they receive, as well as the potential for even greater control in the future as technology develops."[56] No other medium under the U.S. regulatory regime offers users or sites the ability to repackage the information. Newspapers are static and the words are the same for everyone, radio broadcasts are not personally recorded for each listener, and books are not reprinted depending on the font preferences of the reader. But the internet is interactive, which is what the statute codifies as a distinction here. Finally, Section 230(a)(4) confirms again the legislative intent of both the House (Cox and Wyden) and the Senate (Leahy) by writing, "The Internet and other interactive computer services have flourished, to the benefit of all Americans, with a minimum of government regulation."[57] It's clear that the Wyden-Cox-Leahy perspective on internet exceptionalism prevailed over the Exon perspective that the internet should be held to the same regulatory level as other mediums. Was that the right choice? Such will be explored later in this chapter, but Section 230(a) is important evidence that Congress made up its mind (in theory).

Section 230(b) from the perspective of "the forest" appears to represent evidence of Wyden-Cox-Leahy's perspective again, but at the level of "the trees" leaves open legal loopholes for supporters of Exon's perspective for agency regulation. Section 230(b)(2) is one of the most widely quoted sections of the text. It explains that "it is the policy of the United States to *preserve the vibrant and competitive free market* that presently exists for the Internet and other

[56] 47 U.S.C. § 230(a)(2) (1996) (emphasis added)
[57] 47 U.S.C. § 230(a)(4) (1996)

interactive computer services, *unfettered by Federal or State regulation.*"[58] With that being said, the rest of the subsection implies a need for a Federal regulator. Section 230(b)(1, 3-5) write:

> "It is the policy of the United States –
> (1) to promote the continued development of the Internet and other interactive computer services and other interactive media; […]
> (3) to encourage the development of technologies which maximize user control over what information is received by individuals, families, and schools who use the Internet and other interactive computer services;
> (4) to remove disincentives for the development and utilization of blocking and filtering technologies that empower parents to restrict their children's access to objectionable or inappropriate online material; and
> (5) to ensure vigorous enforcement of Federal criminal laws to deter and punish trafficking in obscenity, stalking, and harassment by means of computer."[59]

Such language further muddies the water between Exon's perspective and that codified elsewhere in the statute. As the FCC will later argue, "promoting," "encouraging," "removing disincentives," "and ensuring vigorous enforcement" all require an overseer. In the eyes of the FCC, Congress' intention to prevent the fettering of the internet needs to be carried out and prevented by a Federal overseer. This will be later explored in this chapter, but represent a crucial contradiction between deregulation and enforcing that deregulation.

Section 230(c) is the famous twenty-six words that Kosseff alleges created the internet. This is a line that the in-house attorneys at Facebook and civil rights groups alike know by heart. Section 230(c)(1) gives immunity to sites for things they choose to leave up by proclaiming, "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."[60] In response to pages upon pages of briefs and amicus filings, these twenty-six words are all a site needs to submit at times to get a case against them dismissed. Section 230(c)(2)(A) was written as a direct response to the holding in *Prodigy*. It immunizes ISPs for good faith efforts to restrict access to objectionable

[58] 47 U.S.C. § 230(b)(2) (1996) (emphasis added)
[59] 47 U.S.C. § 230(b)(1, 3-5) (1996)
[60] 47 U.S.C. § 230(c)(1) (1996) (emphasis added)

material. Sites do not open themselves up to civil liability from users whose content was removed as long as it is done in a good faith effort to preserve the internet ecosystem. For example, anti-competitive motives, like Twitter removing Zuckerberg's posts on Twitter, are not protected. Finally, Section 230(c)(2)(B) enables sites to offer users the opportunity to self-select what they individually would like to see. This clause provides liability protection to Facebook for offering a "hide" button or TikTok for offering a "don't show me more like this" feature. Section 230(c) has both a sword and a shield; subsection (c)(1) is a shield from liability for what they choose to keep up and subsection (c)(2) is a sword that allows sites to remove objectionable content.

Finally, Section 230(e) offers exemptions to these liability protections. Section 230(c) protection does not apply to: Federal criminal law, IP infringements, state laws consistent with Section 230, sites that host certain illegally obtained information[61], and any sites that knowingly engage in sex trafficking or intend to promote it. A critical reader of my argument would see Section 230(e)(1) about Federal criminal law and argue that sites hosting inciting speech do not fall within the protections. That would make sense; inciting violence against the U.S. government is a Federal crime that this subsection says is exempt from immunity protections. Therefore we would expect sites to more actively monitor for this speech because they're not immune.

To which, there are two responses. First, rarely are users on Facebook explicitly saying, "I would like you all to storm the Capitol and hurt someone on this day at this time." Generally, the speech on these sites is inflammatory but fails to reach a bar where it constitutes a call to explicit action. As it relates to the application of this specific crime to the specific incidence of January 6th, it seems more likely than not that the speech being put out on Facebook or Gab or

[61] See the Electronic Communications Privacy Act of 1986

Parler was general calls to actions and restatements of misinformation where people drew their own conclusions to storm the Capitol together. Once people have been amassed, their planning of the "imminent lawless action"[62] takes place on encrypted sites like Telegram or Whatsapp, some type of alt-tech site, or on the dark web. Even if a user explicitly posts on Facebook that they intend to do something criminal and incites others to join him or her in that specific crime, Facebook needs to be notified that something said on their site broke a Federal statute and have time to respond before being held liable. There are too many federal criminal statutes in existence that a requirement to monitor for all of them would result in the 'constant monitoring' dilemma that has played out in Europe for the past twenty-five years. All of that being said, some good comes of this criminal law exemption. But that good is not a panacea for objectionable content, especially those that fail to hit high bars of federal proof or hit those high bars of proof on sites that are harder to police (dark web, alt-tech sites).

Site	Founding	Alternative to...	Status
4chan	2003	Instagram	Active
Epik	2009	Amazon Web Services	Active
MeWe	2012	Facebook	Active
8chan	2013	Instagram	Active
Rumble	2013	YouTube	Active
Telegram	2013	AIM Instant Messenger	Active
Signal	2014	AIM Instant Messenger	Active
Minds	2015	Facebook	Active
Odysee (LBRY)	2015	YouTube	Active
Triller	2015	YouTube	Active
WeSearchr	2015	Crowdfunding	Active

[62] From the "clear and present danger" doctrine set forth by *Brandenburg v. Ohio* (1969) for defining when the government can curtail your freedom of speech.

DTube	2016	YouTube	Active
Gab	2016	Facebook/Twitter	Active
WASP Love	2016	Dating Site	Active
BitChute	2017	YouTube	Active
DLive	2017	YouTube	Active
GoyFundMe	2017	Crowdfunding	Defunct
Hatreon	2017	Crowdfunding	Defunct
PewTube	2017	YouTube	Defunct
Slug	2017	Facebook	Active
SubscribeStar	2017	Crowdfunding	Active
WrongThink	2017	Facebook	Active
Parler	2018	Facebook/Twitter	Active
Thinkspot	2019	Crowdfunding	Active
Gettr	2021	YouTube	Active

Figure 5: Non-exhaustive, sample list of alt-tech ISPs that accomplish traditionally-mainstream functions but boast stronger encryption or less access to the general population.

Judicial Manifestations of Section 230

As mentioned, this text creates a deregulatory framework and a "court order standard" where there is not only an incentive to keep the internet free from government intervention but also that any attempts for government intervention are decided by the courts. Section 230 gets an infamous reputation among myopic onlookers as a statute that allegedly condones terrible things being on the internet. There are many ways that the U.S. could have created their regulatory regime, but they chose to do so by pairing a deregulatory framework with a court order standard. While this combination does maximize the total amount of information available to consumers of the internet, it also allows for the largest amount of objectionable material to remain online. The cases below will illustrate that such a combination has manifested in objectionable material

remaining online. With neither a governmental body nor agency to police content not brought before a court nor an incentive for sites to maximize their content moderation efforts, a body of case law is created that highlights the objectionable content being created through this platform complacency.

The first major Section 230 case, *Zeran v. America Online, Inc.* (1997), arrived at the courts less than a year after its passage. Kenneth Zeran was an ordinary citizen living an ordinary life. The only thing that made him unique was that his phone rang non-stop starting in April 1995. Shortly after the bombing of the Alfred P. Murrah Federal Building in Oklahoma City, an anonymous user of AOL posted an advertisement selling shirts that glorified the bomber and his motives. The anonymous user listed Ken by name and listed his phone number at the bottom of the listing. Zeran reached out to AOL in the hopes that they would remove the fraudulent posting and an employee said that the platform would do so. However, the posting would since be reposted multiple times by other anonymous users of AOL despite the original one being removed. Zeran alleges AOL lost its protections as a distributor once it took down the advertisement but failed to screen for future iterations of the content. Ken and his lawyer, Leo Kayser, acknowledged that they couldn't hold AOL liable as a publisher because of Section 230, but they alleged that the site was given notice of the illegal, defamatory material and should therefore be liable for giving a platform to such content's reproduction.

The famous ruling by Judge Wilkinson sided with AOL in holding that a distributor is the same as the word *publisher* and that Section 230's use of publisher included those that didn't even curate the content to which they gave a platform. Wilkinson wrote that "subjecting a computer service provider to liability based on the provider's knowledge would 'reinforce service providers' incentives to restrict speech and abstain from self-regulation.'"[63] This *Zeran*

[63] *Zeran v. Am. Online, Inc.*, 958 F. Supp. 1124, 1126 (E.D. Va. 1997) at 333

ruling was a major case for proving Section 230's purpose because it (1) distinguished the US regulatory regime from the European model and its downsides and (2) it conferred immunity to both sites that moderate content (like Section 230(c)(2) (A) intended) and those that do not (like Section 230(c)(1) intended). Kenneth would continue to receive death threats, his phone would continue ringing as this content was republished, and the anonymous posting prevented him from receiving any recourse. Furthermore, cases like Ken's would be just the beginning of a case law riddled with victimized people not receiving recourse.

Case	Year	What went unpunished as a result of a Section 230 defense?
Blumenthal v. Drudge	1998	The Drudge Report posted an article alleging that Sydney Blumenthal, a White House staffer, assaulted his wife, also a White House staffer. Both husband and wife denied the assault happened and sued for defamation. The Drudge Report pays AOL to amplify its content. This article was amplified through Capitol Hill circles but **the Court ruled that, while the Drudge Report could be held liable, AOL was immune from liability for reproducing and amplifying the content.**
Doe v. America Online	2001	An 11-year old boy was lured into sexual acts with an adult defendant via AOL chat rooms. The defendant recorded the encounter and marketed this video online. While the defendant could be tried for these crimes, **Justice Wells ruled that AOL was immune from liability and had no legal obligation to prevent the future reproduction of this content by other users on the site.**
Batzel v. Smith	2003	Smith accused Batzel, a California attorney, of owning art stolen by the Nazis and being a descendant of Heinrich Himmler. **Batzel's career was upended and still lost her case.**
Carafano v. Metrosplash.com, Inc.	2003	Famous Star Trek actress, Christianne Carafano, was being stalked by a man in Berlin. He eventually created a fake dating profile with her name and image and posted her personal phone number and address (which violated the

		dating site's terms of service). Other stalkers would leave sexually explicit voicemails and threats against her and her son. She and Jeremiah had to flee their house and live hotel-to-hotel for months until the dating site eventually took down her post. She alleged the dating site should have had content filters to spot this infraction and rectify the situation faster upon receiving notice. **The Court disagreed and said matchmaker.com had no obligation to seek or disable such content.**
Doe v. MySpace	2008	A 13-year old girl was lured into meeting a 19-year old boy she met on MySpace and was sexually assaulted. **The Court ruled that MySpace was not liable for hosting such content or for being an accomplice to the actions of the sexual preditor.**
Barnes v. Yahoo!	2009	Barnes' ex-boyfriend posted revenge pornography of her online following their breakup and pretended to be her on multiple chat rooms to entice men into sexual relationships with Barnes. These postings listed her address and phone number. These men would repeatedly call and arrive at her home and office under the promise of solicited sex. This case had to deal with a legal principle known as *promissory estoppel* as the central issue. But the **Court ruled that Yahoo! was immunized under Section 230 from having to take down these posts even if they promised they would.**
Jones v. Dirty World Entertainment Recordings LLC	2014	Posts were made online (and endorsed by the site's moderator) about high school teacher and Bengals cheerleader, Sarah Jones, alleging she was sexually promiscuous, STI-positive, and having inappropriate relationships with her students. **The Court ruled that the site and its moderator were both immune under Section 230 and Jones had to pay Dirty World Entertainment $6,000+.**
Klayman v. Zuckerberg	2014	Klayman sued Facebook after they dragged their feet in removing the "Third Palestinian Intifada" Facebook page which called for Muslims around the world to rise up and kill Jewish people. **Facebook and Zuckerberg easily dismiss the case under Section 230.**
Doe v. Internet Brands	2014	Doe was using a website that connects aspiring models to shoots looking for models. The site had both an anecdotal and analytical history of being used to lure women into nonconsensual sexual encounters. Doe was lured into a fake

		audition, drugged, raped, and recorded. **The Court ruled that the site was immune from punishment despite this history of abuse on their site.**
Doe v. Backpage	2016	A group of parents filed suit in Massachusetts alleging that the sex trafficking of their children was a result of Backpage.com. They argue the business model of Backpage incentivized such illegal actions and facilitated advertisements for their children. **The Court denied the parents standing to sue and denied that the site's behavior constituted facilitation of such trafficking.** This case would later be the impetus for the SESTA/FOSTA Amendment to Section 230 and the 2017 film *I Am Jane Doe* which details the trial and the lives of the trafficked girls.
Fields v. Twitter	2016	Two American service members, Lloyd "Carl" Fields, Jr. and James Damon Creach, were killed overseas in Jordan by Anwar Abu Zaid. It was discovered he was a member of an ISIS radicalization ring on Twitter and ISIS claimed responsibility. **The Court ruled that Twitter was neither negligent nor liable for the deaths of Fields and Creach.**
Herrick v. Grindr	2019	Herrick's ex-boyfriend began impersonating Herrick on Grindr, a dating app, and instructing suitors that he ("Herrick") had a hardcore rape fantasy and preferred role play. Some men would come to Herrick's office and physically assault him and threaten him and his co-workers. **The Court ruled that Grindr was not liable for hosting the fraudulent profiles despite notification of illegal activity.**
Force v. Facebook	2019	Families of five Americans who were killed or hurt in Israel by attacks perpetrated by Hamas, a defined terror organization. Plaintiffs allege that Facebook knowingly hosted Hamas profiles and their algorithm funneled radicalizing content and terror plots to like-minded, but unaffiliated anti-Israel terrorists. **The Court ruled that Facebook's algorithm did not constitute a form of editorial discretion over the content and that Section 230 could immunize Facebook.**

Figure 6: A selection of a few, high-profile cases that sided with ISPs and ruled in favor of their Section 230(c)(1) immunity to allow such content to remain online. These are strictly summaries and do not include the many levels of nuance that these cases present. The objectionable content is highlighted and the result is bolded. (Source: Kosseff (2019) and Brannon and Holmes (2021))

Limitations of the "court order standard" employed by Section 230

The central argument of this chapter is not proven by the cases themselves, such is out of the scope of my question. Let's assume that all the cases above are justly decided given the facts of each case. From an institutional perspective, vesting the courts (rather than with the ISPs or with a government agency) with the power to provide recourse for victims came with drawbacks. First, judicial remedy requires that the injustice occurs before the issue at stake can be clarified. Second, these clarified issues are often decided on narrow grounds that depend on the facts of the case. This makes it difficult to use these cases as the building blocks of a clearer legal landscape for the internet ecosystem. While an agency rule or clarification can have wide applicability, the court ruling in *Zeran* or *Force* or *Barnes*, for example, only serves to prevent future infractions of the same circumstances and in the same jurisdictions. After all, Ninth Circuit rulings are not binding in Virginia in the same way that Second Circuit rulings are not binding to cases that arise in Silicon Valley. Such "circuit split" became evident recently when the Third Circuit ruled in 2021 that celebrities could sue Facebook for unauthorized use of their images in state court under Section 230(e)(2)[64] but such was a direct contradiction to the holding in another famous Section 230 case in 2007 in the Ninth Circuit.[65] All of this is to say that there exist issues with vesting recourse in the courts.

This is not to say that a judicially regulatory pathway is bad in and of itself. After all, the newspaper industry has used it for years. As will be later explained, the content of newspapers is not under the jurisdiction of the FCC or another government agency. Therefore victims of what they publish do not receive recourse through future rulemaking clarifications but only through the courts evaluating their claims. The difference between the judicial oversight of newspapers

[64] *Hepp v. Facebook, Inc.*, 465 F. Supp. 3d 491 (3d Cir. 2021)
[65] *Perfect 10, Inc. v. CCBill LLC*, 488 F.3d 1102 (9th Cir. 2007)

and of the internet is where the regulation places the onus. Section 230 specifically precludes ISPs from publisher liability. The burden of proof under "publisher liability" is the responsibility of newspapers in a way that the internet is not burdened as such. If allegations are made against a newspaper in court, the defendants (the newspaper) have to prove that they are entitled to immunity. If allegations are made against an ISP in court, the plaintiffs (victim of the speech) have to prove that the ISP is not entitled to protection. Immunity is the internet's to lose where it's the newspaper industries' to gain. For the internet, Section 230 takes away one of the crucial checks and balances, the onerous burden of proof, that allow such industries not under the regulatory jurisdiction of a government agency to responsibly be a part of a judicial regulatory scheme.

The question then becomes, why isn't the internet under the jurisdiction of the FCC? The FCC already regulates radio, television and phone content through things like safe harbor hours, explicative guidance, and guidelines on nudity and enforces such with fines, sanctions, censures, and licensing penalties. So why is the internet exceptional? The answer is… (very) complicated. But it essentially boils down to the way that early Congressional instruction and FCC interpretation hamstrung the future regulation of the internet. It should be noted that, while the FCC is an independent agency, its chairpeople who vote are politically appointed. This is evident in the countless 3-2 votes undertaken by the panel of five in charge of overseeing the FCC that almost always falls upon partisan lines. While the FCC is politically volatile, it's a nimble, proactive, and dynamic regulatory complement to the courts.

FCC Authority over the Regulation of the Internet

In contrast to the courts and Congress, the FCC has more access to industry knowledge. As the literature has illuminated, there exists a pervasive exchange between the FCC and industry leaders that some have dubbed a "revolving door" (Gilens and Page 2014; Popiel 2018). Their regulators go on to work at broadband, cellular, telecommunications, and radio companies and vice versa. As a result of the regularity of this knowledge and the proactive nature of agencies in comparison to courts, the FCC can more agilely make rulings that adapt to new technologies. We've seen examples of their hands-on regulatory abilities in the past: the modification of explicative standards, changing definitions of sexual innuendo, and inclusion or exclusion of other types of speech as they become more or less societally acceptable (Levi 2008). The FCC has power over most things related to communication, telecommunications, and broadcasting with hands in both the economics and production of all this material. From a speech perspective, the FCC has control over what can be said, when it can be said, who can say it, and the punishments for violating such rules. But the FCC chairs are politically appointed and the FCC is therefore subject to changing politics as administrations change.

As a result of this, the FCC's policymaking can best be described as they're writing a chain novel. The chairs have five-year terms and each successive administration of chairs gets to write a new chapter over the next four years but it has to be in line with the chapters made by the administrations before. They can chalk a few chapters up to being a "dream sequence" or a flashback but they can't ignore the precedents and backstories of characters in the novel. As a result, the FCC's regulatory scheme can best be described as a patchwork. As Thompson (2010) notes using a gardening analogy, they can plant new policies but some of those policies can't touch other policies in the garden and some need more or less sunlight than others. So trying to

adapt this garden to a whole new plant, or a whole new industry like the internet, can create issues of jurisdiction that are hindered by previous plantings. In addition to the issues that the FCC has with incorporating this new plant into their metaphorical garden, there are also issues of whether they're allowed to be the gardener in the first place.

As I will argue in the subsequent sections, the FCC has questionable authority over the regulation of the internet and as such has refrained from issuing heavy-handed regulation over speech on the internet to avoid their jurisdiction of the internet as a whole from being challenged in court. The argument begins with Congress' simultaneous definition of the FCC as both the administrator and not the administrator over the internet in their 1996 bill. From there, the FCC started a chain novel by pulling at textual straws in the bill to establish jurisdiction over the internet through a constantly modifying precedent. This further complicated the existing patchwork of a regulatory regime. By the 2010s, the FCC took a big step in the direction of genuine internet jurisdiction. But in doing so, it linked speech on the internet to "net neutrality," which would further complicate regulation as those two issues became part of politically opposing agendas by the early 2020s. With all of this in mind, my argument closes by addressing the way that legislative corrections to this patchwork regulatory regime are nearly impossible as a result of the uncontrollable rise of these ISPs and the radically different political environment that had developed by the time the U.S. could realize they had a problem.

Was the FCC vested with regulatory jurisdiction over the internet?

The FCC would like the answer to be "yes" but Congress constructed the 1996 bill in a way that left such ambiguous. Congress made very explicit within Section 230 that the FCC was to have no hand in regulating the internet, but legally Section 230 is placed within a larger bill

that amends Title II of the Communications Act of 1934 which is explicitly under the FCC's jurisdiction. Questions arise of policy statements and legislative intent and how the two reconcile when in conflict. This section will first recap the way Congress and the FCC fought over the jurisdiction of the internet. Second, it will explain how the FCC fought with its past self to establish jurisdiction over the internet.

The framers of Section 230 unambiguously proclaimed that "it is the policy of the United States to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, *unfettered by Federal or State regulation*."[66] This in addition to the failure of the Exon Amendment and the statements on the floor of the House mentioned above all support the textual proof that Congress intended Section 230 and ISPs be unencumbered by federal regulation, including the FCC. Congress felt that the internet was different from other services (TV, radio, telephone, etc.) under the FCC's umbrella because the "basic justifications for regulating traditional communications (scarcity and monopolization) [did] not apply" (Kennedy and Zallaps 1999, 18). There are endless computers, packets, and passageways to and within the internet that were exponentially growing without the need for physical infrastructure, unlike traditional FCC services that needed telephone wires or radio and television stations.

Despite this exceptionalism coded within Section 230, there exists a strong opposition that similarly cites legislative intent and textual grounding to find FCC jurisdiction. Prior to a committee conference that struck it, there existed a Section 104(d) in the 1996 bill which went even further than the "unfettered" clause within Section 230(b)(2); it specifically enumerated that the FCC was precluded from overseeing the regulation of the internet (Crawford 1997). Section 230(b)(2) is strictly a statement of legislative intent, not of policy direction. Proponents of FCC

[66] Id. at 58

jurisdiction further cite the dissonance within Section 230(b), the same section that contains the "unfettered" clause. As mentioned above, Section 230(b)(1, 3-5) represents a hidden escape valve of sorts that was added to the bill by someone from Exon's side of the issue. They imply that an actor will "promote… development," "encourage… development," "remove disincentives" to the development of the internet in addition to "[ensuring] vigorous enforcement of Federal criminal laws" but do not name who the actor is (or is not).[67] Such has been construed to imply that the FCC, whose Communications Act is being amended by the 1996 bill, is the unspoken actor. This dissonance and the contrasting signals of legislative intent would riddle the next two decades of attempts made by the FCC to establish authority. To understand better the regulatory patchwork they would create thanks to this ambiguity, we need to understand the tools that the FCC could use to create it.

Pathways of Acquiring Jurisdiction

The FCC has two major legal pathways to authority over an industry: statutory authority and ancillary authority. The FCC has statutory authority over "common carriers." Common carriers are services under Title II of the Communications Act of 1934. They are defined as a service engaged "in interstate or foreign communication *by wire or radio* or interstate or foreign radio transmission of energy."[68] This refers to telephones, radio transmissions, and the original medium of cable television over wires. Under this authority, the FCC has complete jurisdiction over these mediums. It can regulate the communication that goes over the wires, the installation of the wires, the merging of these wired companies, and anything else that is done pursuant to defending the principles of common carrier communications like competition, fairness, and

[67] Id. at 17
[68] 47 U.S.C. § 151(11) (1996) (emphasis added)

access. An important distinction is that while all common carriers are wired, not all wired communications are common carriers.

Although the early days of the internet used dial-up and other telephone-adjacent technologies, the internet was not cleanly classified as a common carrier. Although the "interactive computer services" (ISPs) are listed in a provision (Section 230) within Title II, they are defined as a type of "information service" under Title I whose two-way, interactive nature distinguishes it from the characteristics of common carriers. Title I services are "unregulated services" of the FCC not under the statutory authority. But, because Title I services are still under the larger Communications Act of 1934, they can be regulated but only if such is tangentially related to one of its Title II authorities. This is called a *Chevron* defense. In 1984, environmental groups challenged the EPA's ability to quantify provisions of the Clean Air Act that were left unquantified by Congress' original text. A unanimous Supreme Court found [69] that Congress did not enumerate a specific intention for the interpretation of the term and that the EPA was in the best position to create a reasonable policy choice. Therefore, the EPA was permitted to issue clarifying policy. This legal leeway was clarified again in 2001 when the Supreme Court ruled that any agency interpretations of statutes "qualify for *Chevron* deference when it appears that Congress delegated authority to the agency to make rules carrying the force of law, and that the agency interpretation claiming deference was promulgated in the exercise of that authority."[70]

Only once (2015-2017) did the FCC have statutory authority over "interactive computer services" and it only came as a result of decades of litigation and the precedent established in a major 2005 Supreme Court case. Otherwise, the FCC has had a mix of ancillary authority or no authority for the remainder of its interactions with ISPs. Even that statutory authority in 2015

[69] *Chevron U.S.A. Inc. v. Natural Resources Defense Council, Inc.*, 467 U.S. 837 (1984)
[70] *United States v. Mead Corp.*, 533 U.S. 218, 226–27 (2001)

would eventually be overturned in 2017 when former President Trump would take office. The complicated story of patchwork regulatory attempts to secure ancillary jurisdiction over the internet begins in the 1960s when the internet was at its absolute infancy. From here, I'll be illustrating the haphazard way the FCC has written its chain novel over the years as a way of demonstrating how the court order standard in the U.S. failed to prevent the rise of objectionable material. Section 230(c)(1) protects content left up and Section 230(c)(2) immunizes ISPs for good faith efforts to enact self-moderation. The court order standard privileges ISPs who can hide behind Section 230 for the content left up and the only check that could have existed on the system is if the FCC had less volatile jurisdiction over these ISPs and could have incentivized them to undertake more good faith efforts to self-moderate.

Volatile History of FCC Jurisdiction over ISPs

In 1966, the FCC initiated *Computer Inquiry I*, a policy that sought to differentiate computer communication from data processing. To the commission, communication was defined as a "basic" service of computers, while data processing was defined as "enhanced."[71] Computer communication at the time (called "message switching") was a very rudimentary form of sending code between terminals with little to no reformatting. The commission created rules that prevented competition over these wires between terminals but it deregulated, encouraged growth, and promoted competition when it came to creating additional safeguards for data processing. The FCC wanted technologies with the power to not just transmit code but reformat it and be interactive to shift their business models to one that rewarded the company with the most

[71] In re Regulatory & Policy Problems Presented by the Interdependence of Computer and Communication Services & Facilities, Notice of Inquiry, 7 FCC 2d 11 (1966)

safeguards for security and privacy (Kennedy and Zallaps 1999, Esbin 1999). This would set the early stages for modern ISPs who take on the functions of "enhanced services."

In the late-1970s and mid-1980s, the FCC refined this with their *Computer II*[72] and *III*[73] *Inquiries*. The internet was now leaving the jurisdiction of government offices and the military. With the internet rolling out to consumers, the FCC had to clarify its jurisdiction. They "determined that the enhanced services market was competitive, and that consumers were deriving benefits from this competition" so "the Commission declined to regulate enhanced services as common carriage under Title II of the Act" (Esbin 1999, 60). This would be the first ruling by the FCC: they explicitly reject jurisdiction over ISPs that reformat the text being sent but still retain jurisdiction over simple "message switching" ISPs who don't modify the information. Esbin (1999) simplifies this perfectly when she writes, "therefore, a service that allows subscribers to calculate the Dow Jones average using software located off the subscribers' premises would not be a cable service, even though a service that makes the Dow Jones average available to all subscribers would be a cable service" (85).

After this lack of jurisdiction was seemingly codified in a parallel way with Section 230, phone companies were upset that the internet was not being held to the same standard as they were. From an economic standpoint, they feared that internet telephony would drive traditional telecommunications out of business. As a result, their trade association, America's Carriers Telecommunication Association (ACTA), petitioned[74] the FCC for rulemaking to take control of the internet in the same way it regulated their traditional telephone clients. Likely recognizing

[72] Second Computer Inquiry, Tentative Decision and Further Notice of Inquiry and Rulemaking, 72 FCC 2d 358, 45 Rad. Reg.2d (P & F) 1485 (1979)
[73] Amendment of Sections 64.702 of the Commission's Rules and Regs, Report and Order, 104 FCC 2d 958, 45 Rad. Reg.2d (P & F) 603 (1986)
[74] In re Provision of Interstate and International Interexchange Telecommunications Service Via the "Internet" by Non-Tariffed, Uncertified Entities, Petition for Declaratory Ruling, Special Relief and Institution of Rulemaking of America's Carriers Telecommunication Association, RM 8775 (1996)

either the validity of their concern or the costs of displeasing the telecommunications industry, the FCC immediately began seeking ancillary authority over the internet ("enhanced services" as it was known at the time). In a 1996 order, the FCC cited no less than four textual grounds for ancillary authority.[75] Two years later, a new round of FCC chairs would declare, in their bi-annual address to Congress[76], that the use of cable doesn't make the internet subject to Title II jurisdiction. This, in conjunction with their 1998 proposed rulemaking order[77] which attempted to codify such a sentiment, represents a stark contrast from their 1996 reasoning. To boot, the 1998 and 1996 orders both cited the same sections (Section 271-272) to justify their contradictory reasonings. Even worse, this 1998 reasoning would be reversed again in a 1999 FCC order[78] that used a definition technicality as an attempt to "backdoor" their way into internet regulation. From this point, the chain novel had commenced. Subsequent FCC attempts to regulate and then disavow regulation over "enhanced services" would need to be ever more complex in order to not disrupt the legitimacy of previous orders.

[75] Implementation of the Non-Accounting Safeguards of Section 271 and 272 of the Communications Act of 1934, as amended, 11 FCC Rcd 21905 (1996)
[76] In re Federal-State Joint Board on Universal Service, Report to Congress, 13 FCC Rcd. 11501, para. 73 (1998)
[77] In re Deployment of Wireline Services Offering Advanced Telecommunications Capability, Memorandum Opinion and Order, and Notice of Proposed Rulemaking, 13 FCC Rcd. 24012-13, (1998)
[78] In re Inquiry Concerning the Deployment of Advanced Telecommunications Capability to All Americans in a Reasonable and Timely Fashion, and Possible Steps to Accelerate Such Deployment Pursuant to Section 706 of the Telecommunications Act of 1996, 14 FCC Rcd. 2398 (1999)

Text Cited	Argument Summary	Where?
Section 151	The internet uses cables and is, therefore, a common carrier under Title II regulation.	1996 ACTA Petition
Section 271/272	These sections give the FCC regulatory authority over "Bell Operating Companies," which are the local internet connection providers that resulted after a 1982 breakup of AT&T. If the FCC can control the internet connection, it, therefore, has the power to control ISP access for consumers. If it can control ISP access, it can control access to ISP content by extension.	1996 FCC Order
Section 601	Section 601 says that nothing within the Telecommunications Act of 1996 should be construed as limiting or circumventing Federal authority unless otherwise stated. Section 230(b)(2) is not a legally binding statement and therefore its preclusion of Federal regulation is not reason enough to prevent this section from giving the FCC jurisdiction.	
Section 254	Section 254 empowers the FCC to provide equitable "access to advanced telecommunications and information services should be provided in all regions of the nation."[79]	
Section 706	FCC argued that an "interactive computer service" was a type of "advanced telecommunications capacity" (ATC) provider which is under Title II.	1999 FCC Order

Figure 7: An illustration of the unique sections that had to be cited in each successive attempt by the FCC to get ancillary authority after being hamstrung by a previous FCC order. (Source: Crawford 1997, Fried 1999, Comstock and Butler 2000, Whiteley 2000, Chen 2001)

[79] 47 U.S.C. § 254(b)(2) (1996)

Although this is a simplified illustration of the nuance that underscores each iterative attempt by the FCC, it goes to show that the FCC hamstrung itself at an early point and soiled the metaphorical regulatory garden for years to come. As Lyons (2013) notes, "the Commission has invoked its ancillary authority roughly 65 times since 2007 – approximately the same number of times it has invoked its ancillary authority during the entirety of its history up to that point… the Commission is grasping at statutory language to justify imposing its own mandates" (5). As Esbin (2009) comically furthers, "Congress does not hide elephants in mouseholes… this exercise — searching for snippets and threads of regulatory authority over a communications medium as significant as the Internet in multiple, unrelated statutory provisions — itself should signal to the reviewing court that the exercise is unlikely to produce a credible source of authority" (2-3). This unclear regulatory designation would result in decades more of this haphazard chain novel.

Date	Action	Effect on Internet Industry
2002	FCC determines that ISPs are not a Title II "telecommunications provider" and were solely a Title I "information service"	Deregulatory
2003	Ninth circuit ruled that ISPs had telecommunication functions and should be regulated under Title II	Regulatory
2004	FCC Chair Powell implements this ruling by asking for *voluntary* neutrality of ISPs in his "Preserving Network Freedom" order	Regulatory
2005	*National Cable & Telecommunications Ass'n v. Brand X Internet Services* (2005) - Supreme Court reverses the 9th circuit case (2003) and rules that FCC currently placed ISPs under Title I but that it has the power to change that	Both Regulatory and Deregulatory
2006	Comcast/AT&T contract establishes voluntary net	Regulatory

	neutrality in their contract (Esbin and Marcus 2008)	
2008	FCC censures Comcast/AT&T for breaking the voluntary "Preserving Network Freedom" order	Regulatory
2010	DC Circuit Court of Appeals overturns FCC's censure and vacates their "Preserving Network Freedom" order because it was not enacted in a way that passed the *Chevron* (1984) test	Deregulatory
2010	FCC alleges textual authority in a 2010 "Open Internet Order" under Section 706 (which SCOTUS implied in their 2005 decision) rather than ancillary authority	Regulatory
2011	Verizon sued the FCC alleging that it was applying Title II principles to Title I services and wins their case in 2014	Deregulatory
2015	An Obama FCC takes the leap and reclassifies ISPs under Title II	Regulatory
2017	A Trump FCC reverses Obama's Title II classification and vacates net neutrality	Deregulatory
2021	A Biden FCC announces it plans to reverse Trump's reversal and reinstate net neutrality	Regulatory

Figure 8: A synthesis of the back-and-forth between the FCC, the courts, and private ISPs over net neutrality and their attempts to patch together authority from 2001-2021 (Source: Esbin and Marcus 2008, Esbin 2009, Gioia 2008, Thompson 2010, Ruane 2011, Candeub 2020, Cheah 2021)

By the turn of the century, the courts finally had enough case law to hear a definitive case about the FCC's purported jurisdiction over the internet. In a landmark case, *National Cable & Telecommunications Association v. Brand X Internet Services* (2005), the FCC had to concede that it did not have statutory authority over the internet but the Supreme Court did agree that it had ancillary authority under *Chevron* (1984). The court said that the FCC's current classification of the internet left it under Title I but importantly affirmed that the FCC, not Congress, has the power to reclassify it. As a result, this issue received executive branch

attention and became an issue for presidential administrations rather than changing administrations within the FCC. From 2005 to 2015, the internet remained unregulated[80]. In 2015, President Obama was concerned with the growing inequality in the regulatory landscape. Networks could legally slow down access to certain sites on the internet. President Obama urged the FCC to reclassify it as a Title II common carrier in order to apply a Title II principle, equity of access," to networks on the internet in the same way it was required for telecommunications carriers. The FCC obliged and voted 3-2 along party lines to approve the change.[81] In 2017, President Trump urged the FCC, under Ajit Pai, to reverse that order and remove net neutrality principles. Pai obliged and voted 3-2 along party lines to reverse the reclassification and make the internet an unregulated service of the FCC again.[82] In 2021, President Biden passed an executive order directing the FCC to reverse that reversal and reinstate net neutrality[83]; that instruction is just waiting on the Senate to confirm Gigi Sohn to fill the vacancy left by Pai's retirement in 2021.

Impact of Volatile Jurisdiction

This patchwork regulatory framework is not just distressing to academics and policy wonks but has manifested itself in tangible harms for free speech. This created a volatile regulatory environment where especially nascent ISPs in the late 90s were forced to act as if they were in Europe. With uncertainty over whether they were subject to potential FCC fines at any given time, ISPs were incentivized to remove any content that could violate the FCC decency and obscenity guidelines established for common carriers. As Esbin (2009) notes, "unchecked

[80] A 2010 FCC, under Obama, attempted to install net neutrality prior to 2015. But that attempt was struck down by the Supreme Court in *Comcast v. FCC* (2010) because the way the FCC attempted to pass the change used authority the agency did not have.
[81] In re Protecting and Promoting the Open Internet, 30 FCC Rcd 5601 (2015)
[82] In re Restoring Internet Freedom, 33 FCC Rcd 311 (2018)
[83] Executive Order No. 14036, 86 Fed. Reg. 36987 (2021)

regulatory discretion under the amorphous doctrine of 'ancillary jurisdiction' is every bit as big a danger to a free and open Internet as any of the other dangers the FCC posits" (5). In addition to abstract incentives, this chain novel had real impacts on judicial equity. This volatility meant some ISPs were getting protection while others were not despite being sued for the same thing and it all depended on when the suit happened amidst changing, or otherwise unclear, regulations.

Take for example the *Cubby* (1991), *Stratton* (1996), and *Lunney* (1999) case study. In 1991, *Cubby* was protected because it was a "distributor" and exercised no editorial control. In 1996, Prodigy was punished for cleaning up the internet because it exerted some editorial control and was, therefore, a "publisher." In 1999, Prodigy was sued for the same charge (defamation) in the same Southern District of New York court but this time by teenager Alex Lunney. Following the 1999 FCC Order[84], Prodigy argued that it was akin to the telephone company. Prodigy argued that telephone companies are not responsible for tapping every phone, so they shouldn't be responsible for tapping every message. While this sounds logical, remember that this reasoning runs counter to the exceptionalist distinction that was made in the 1960s and 70s in the *Computer Inquiries* and codified in the 1996 bill. In this 1999 case, the judge bought Prodigy's rationale and gave them immunity under a 1974 case (*Anderson v. New York Tel. Co*) that is only supposed to apply to telephone companies. Although the 1999 Prodigy case ended the way Cox and Wyden would have wanted, it wasn't for the right reasons. ISPs shouldn't be able to classify as common carriers when it's convenient and unregulated services when it's not. If Prodigy really felt it was a common carrier, then the FCC would have had some oversight that could have prevented Lunney from being wronged in the first place. But Prodigy avoided FCC regulation

[84] Id. at 36

while Lunney was being defamed, but then called themselves an FCC-regulated common carrier before the judge when they needed to get out of their suit.

Although the FCC could be a great resource for incentivizing ISPs to utilize their Section 230(c)(2) self-moderation protections, I argue that the haphazard regulatory patchwork of the FCC has created a volatile legal environment and questions about the validity of FCC jurisdiction that preclude them from being an effective partner in the war on objectionable content. So the question remains: do ISPs need a Federal regulator to incentivize them to self-moderate? The answer is an unfortunate "yes." Will it happen? The answer is also an unfortunate "no."

Expired Pathways to ISP Self-Moderation

With a majority of the world's ISPs being based in the United States, the next logical question becomes: why don't they self-police themselves? After all, they have sweeping Section 230(c)(2) immunity for all good faith efforts at self-policing. Cleaner platforms attract more users and are more profitable than sites that are cesspools of objectionable content. So what's stopping Facebook, Instagram, Youtube, or Twitter from picking up the torch that the FCC dropped? The answer is threefold. First, "ask nicely" was an unsustainable policy. ISPs have a financial opportunity now that didn't exist in the early-2000s which incentivizes inflammatory content to be left up. They hide this financial opportunity behind a plausible and understandable (even if untrue) argument that there is just too much content. Under this sweeping immunity, content creation outpaced the growth of ISPs and their ability to keep up. Second, actions taken in 2015 linked content moderation to net neutrality. As a result of this merge, mandating stricter self-moderation could come to the detriment of another powerful internet principle, equity of

access. Finally, even if there was political consensus at a philosophical level *and* ISPs felt there was requisite public outcry to necessitate coming to the table, these technology companies have a stronghold on actual bill passage that gives them essentially a get-out-of-jail-free card behind the scenes.

Changing ISP Business Models

When the internet was smaller in scale, ISPs were hungry to attract customers and competed with each other for customers by offering them and their families the most content safeguards. Think back to the *Stratton* case in 1996 where Prodigy lauded themselves as a pioneer in content moderation and used newspaper advertisements to flaunt their safeguards. Although they were punished for it in a pre-Section 230 world, their business model was nonetheless the norm at the time. As these tech firms grew, their user bases became, for lack of a better word, addicted to them.

The algorithms kept users engaged and people were not choosing between Facebook or Instagram – the in-vogue thing was to have a Facebook *and* an Instagram account. In addition to the user base landscape changing, the landscape of ISPs themselves can also be said to have changed. This industry developed further as a result of the sweepingly immune legal landscape, companies merged, and those that didn't merge would eventually form a symbiotic relationship. They recognized that there was power in pseudo-collusion. In economic terms, the relationship that these platforms share can be likened to that of a cartel. They are so powerful as a result of this deregulation that they set the rules and govern the conduct of each other. Instead of having to fight to be the cleanest site, as long as they all exert the same level of content moderation the level itself can be rather low.

They make more money, after all, from inflammatory content which drives interactions and therefore drives up page views and ad revenue. Although mitigating objectionable content drives up user bases, these sites have hit an inflection point. Each inflammatory, objectionable post they allow is more financially valuable than the profit lost by that reciprocal number of users being turned off by it (Cohen-Almagor 2017, Gillespie 2020). Tsesis (2017) argues that there are many cases in which this manifests itself. Facebook, in particular, has "reviewed" content that is clearly against its written terms of service and then allowed it to stay online like "Death to Zionist baby killer Israeli Jews" and a page called "Stab Israelis" (609).

Many ISPs can hide this financial incentive for objectionable content behind a veil of technological incapability. In this, they have to ride a fine line between being the leader in technological knowledge and yet simultaneously alleging that they lack the technological ability to moderate such a large quantity of content. As Cohen-Almagor (2015) highlights in *Confronting the Internet's Dark Side: Moral and Social Responsibility on the Free Highway*, there is overwhelming evidence in the technologically-savvy community to show an ability to monitor this quantity of content, but reciprocally overwhelming evidence to show that there is an unwillingness to do so. The sheer scale of ISPs is a plausible excuse to the layman public for not catching enough objectionable content. Only when the financial harms of user outcry exceed the financial benefits of maximizing usership and viewpoints will sites begin to self-police. As Cohen-Almagor (2017) writes, "Facebook invests in profitable activities and does not invest in unprofitable activities." Of all the things on Facebook's plate, content moderation "isn't that important" (433). He furthers that Facebook takes objectively less pressing issues equally if not more seriously than policing for radicalizing content: for example, it has a team of specialists

that only investigates fake identities while that team could be better allocated to the content moderation division which yields much more severe consequences if understaffed (434).

To prove the existence of such ambivalence in the absence of public outcry, look no further than the case study of Anwar al-Awlaki. Anwar al-Awlaki was a leading propagandist for Al Qaeda that the U.S. killed in a 2011 drone strike. He was a well-known figure with a large following among terror cells but not by any means a household name like Osama Bin Laden or Abu Bakr al-Baghdadi. Despite his being deceased, his content remained online. Four years after his death, 68,400 videos of his still remained online. Six years after his death, that number only dropped by 0.6% to 68,000. His videos would become pervasive recruitment material for U.S. domestic jihadi cells. His videos would later go on to inspire 56 individual U.S. extremists whose executed terror plots resulted in the death of 93 civilians, 5 service members, and 417 Americans injured from 2001 to 2019. An additional 17 domestic terror plots within that period were fortunately foiled before coming to completion.

Most notably his videos were cited as inspirations by Nidal Hasan who murdered 13 and wounded 32 in the 2009 Fort Hood shooting rampage, the Boston Marathon bombers, the San Bernadino terrorists, the terrorists to attempted to kill at the "Draw Muhammad" cartoon contest in Garland, TX, Mohammad Youssef Abdulazeez who murdered four Marines in 2015, and the Pulse night club murderer in 2016.[85] After Hassan's rampage at Fort Hood, al-Awlaki glorified him with a post titled "Nidal Hassan Did the Right Thing" (Cohen-Almagor 2017, 443). There is little doubt that al-Awlaki's content was objectionable and radicalized domestic terrorists that harmed Americans. It's extreme to say that Facebook is complicit in this, but it is reasonable to say that Facebook is at the very least more ambivalent than they should be to such content. Hany

[85] Counter Extremism Project. "Anwar Al-Awlaki." Counter Extremism Project. https://www.counterextremism.com/extremists/anwar-al-awlaki (March 29, 2022).

Farid[86] puts it well, "It's not a technical problem, . . . [i]t's a policy issue. I think the speech and privacy issues are tricky. But to say there's nothing we can do about it *is cowardice*." In response to calls from the Counter Extremism Project to delete al-Awlaki's content, YouTube responded with a statement saying it permits "videos posted with a clear news or documentary purpose."[87] No matter where you watch your documentaries or get your news, I'm doubtful that anyone would consider videos with titles like "A Call to Jihad" to be a legitimate documentary or unbiased news broadcast.

Eventually, YouTube would take down almost 60,000 of al-Awlaki's videos in response to public pressure.[88] But such a "watershed moment" as the New York Times called it is in stark juxtaposition to the treatment that other more financially-impactful content received from these sites. As Shane explains, "copyrighted material, child pornography and beheading videos, for example, [pose] an obvious threat to their business" and such are removed at a far more responsive rate than al-Awlaki's content.[89] In another instance of seeming duality, Facebook's Director of Policy, Simon Milner, alleges that flagging potentially terrorist activity is a priority but deflects responsibility for preventing future content by noting that such content creators are best addressed by law enforcement and not repeated censorship by the ISP (Cohen-Almagor 2017).

---

[86] Farid is a professor of Computer Science at Dartmouth College. He is a co-creator of PhotoDNA, the software employed by Facebook and others to detect and delete child pornography.
[87] Shane, Scott. 2015. "Internet Firms Urged to Limit Work of Anwar Al-Awlaki." *New York Times*. https://www.nytimes.com/2015/12/19/us/politics/internet-firms-urged-to-limit-work-of-anwar-al-awlaki.html (March 2, 2022).
[88] Hern, Alex. 2017. "'Youtube Islamist' Anwar Al-Awlaki Videos Removed in Extremism Clampdown." *The Guardian*. https://www.theguardian.com/technology/2017/nov/13/youtube-islamist-anwar-al-awlaki-videos-removed-google-extremism-clampdown (March 03, 2022).
[89] Shane, Scott. 2017. "In 'Watershed Moment,' YouTube Blocks Extremist Cleric's Message." *The New York Times*. https://www.nytimes.com/2017/11/12/us/politics/youtube-terrorism-anwar-al-awlaki.html (March 03, 2022).

Even if we do take the above argument by ISPs at face value and assume that there is indeed too much content on the internet, Milner's statements reflect the fact that Section 230's sweeping immunity led us to that point. Milner's statements don't argue it would be hard to have enacted a more expansive and stringent self-policing policy, he argues it would be hard to "change to" such a system (Cohen-Almagor 2017, 431). This blanket immunity allowed for a rapid proliferation of content which far outpaced the ability of ISPs to grow in tandem (Gillespie 2020). Overall, sites have a financial incentive to keep inflammatory content up and couch that incentive behind a veil of technological inability that is plausible at the layman's level (even if disproven by technology experts).

Veto Power of Big Technology

Even if there is *both* political consensus on a solution and there is enough public outcry to outweigh the financial incentives of the status quo for ISPs, there is a low likelihood that meaningful legislation can be passed. ISPs are more than happy to come to the table but they know that such appearances are largely public relations endeavors. The CEOs of many of the largest social media platforms gave countless hours of testimony before Congress in 2021. Their soundbites were well-rehearsed, direct, and painted a picture that they were not opposed to regulation. This, coupled with the complimenting advertisements[90] of ISPs and increases in the number of Americans in favor of regulating content moderation[91], should signal a shift in the

[90] Facebook, Inc. 2021. iSpot.tv Internet Regulations: Remember the Internet in '96? iSpot.tv. https://www.ispot.tv/ad/OHcg/facebook-internet-regulations-remember-the-internet-in-96 (March 21, 2022). (A Facebook advertisement that ran on cable TV and YouTube shows pictures of technology in the 1990s and then a timeline with how that technology has evolved and concludes with the title "The internet has changed a lot since 1996… Facebook supports updated regulations.")

[91] Kemp, David, and Emily Ekins. 2021. "Poll: 75% Don't Trust Social Media to Make Fair Content Moderation Decisions, 60% Want More Control over Posts They See." *Cato.org*. https://www.cato.org/survey-reports/poll-75-dont-trust-social-media-make-fair-content-moderation-decisions-60-want-more#american-support-tech-industry (March 21, 2022).

Overton window away from the deregulatory history of U.S. technology policy. And yet no reciprocal legislation was passed in response to their seeming cooperation?

In 2021, Apple, Amazon, Google and Facebook alone spent $55 million in registered lobbying expenditures (a 62% increase from 2020). A large part of this increase is tied to these sites fighting allegations that their platforms bore responsibility for radicalizing terror attacks like the January 6th insurrection, the Orlando nightclub shooting, and others. For example, Google increased its lobbying efforts by 27 percent (from $7.5 million in 2020 to $9.5 million in 2021) in response to YouTube's alleged contribution to right-wing extremism.[92] This money goes into killing bills that seek to challenge the technological hegemony of the larger social media platforms with the money to lobby. Popiel notes that this is often successful. In addition to killing bills before they reach the floor this money is also valuable in killing support for bills that do make it to the floor of Congress. He cites Facebook and Google's defeat of the Commercial Privacy Bill of Rights Act of 2011 and Amazon, Facebook, Google, and Twitter's successful defeat of the Do-Not-Track Online Act of 2011 as examples of such power despite Congressional support (Popiel 2018).

[92] Birnbaum, Emily. 2022. "Tech Spent Big on Lobbying Last Year." *POLITICO*. https://www.politico.com/newsletters/morning-tech/2022/01/24/tech-spent-big-on-lobbying-last-year-00001144 (March 21, 2022).

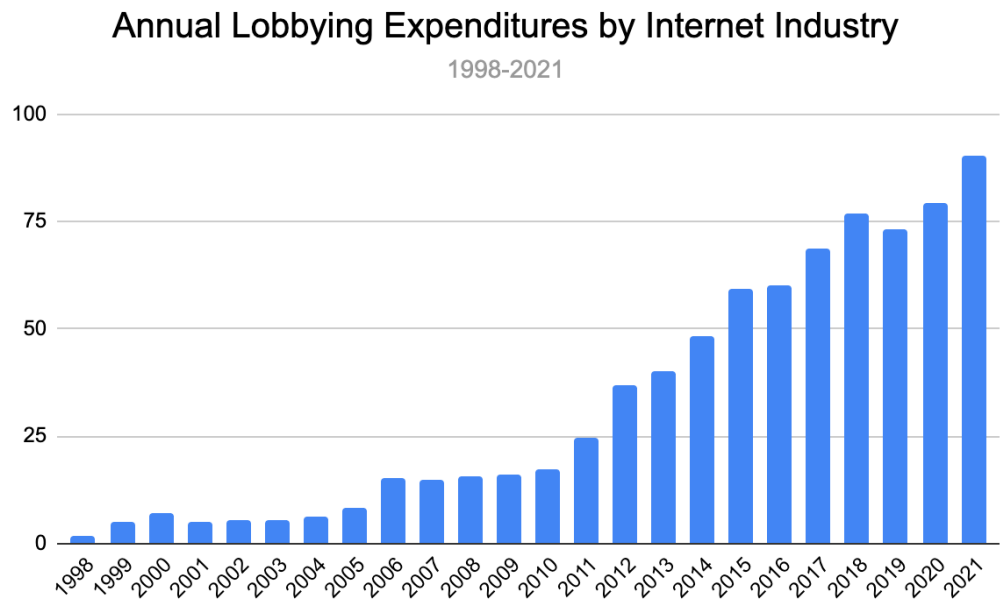

Figure 9: Cumulative lobbying expenditures by internet industry firms *(Source: OpenSecrets)*

In addition to the sheer power of the "Big Tech" lobby, their status as a now quasi-public utility can be used as a bargaining chip. Intelligence communities around the world have increasingly deepened their reliance on the services of these ISPs over the years for crucial signals intelligence (or SIGINT). This new, social media intelligence (dubbed SOCMINT) is used as an argument for why a deregulated internet is actually beneficial to the war against objectionable content and objectionable actions. SOCMINT "provides a new opportunity for law enforcement to generate operational intelligence that could help identify criminal activity, indicate early warning of outbreaks of disorder, provide information and intelligence about groups and individuals, and help understand and respond to public concerns" (Omand et al. 2012, 805). The deregulatory nature of the internet incentivizes speech. Even if objectionable, such speech still points intelligence communities in the direction of those hiring hitmen, those looking to plan attacks, those with psychologically violent tendencies, and pedophiles seeking to groom victims among other examples. As Google CEO, Sundar Pichai, ominously implied, "I

worry that if you regulate for the sake of regulating, it has a lot of unintended consequences … [including] implications for our national security."[93]

If we regulate the internet, this objectionable speech may go away from the public eye. But it will just migrate to the dark web where it is harder for law enforcement to gather SOCMINT. Surveillance-related national interests are often used as a reason to overlook the sector's monopolistic tendencies, economically inefficient copyright protections, and undemocratic commercialization of essential digital services (Popiel 2018). McChesney (2013) details the way that this is both a bargaining chip and a debt owed. In exchange for this seeming deregulatory deference, ISPs routinely comply with thousands of government requests for information. Specifically, Google received at least 10,000 requests for user information from national security organizations in 2011 and voluntarily complied with more than 93% of them (McChesney 2013). Although this relationship has become marginally more regulated as a result of updated data privacy regulations, there still exists an off-book relationship between the U.S. intelligence community and ISPs that affords ISPs a great deal of security from overly regulatory legislative action. Even if there was political will to pass something at an ideological level *and* public outcry was great enough to bring ISPs to the table, these platforms have two crucial insurance policies that prevent unwanted legislative action: their financially significant lobbying efforts and their role as a quasi-public utility within the intelligence community.

[93] Fiegerman, Seth. 2019. "Google CEO Reacts to Looming US Antitrust Probes for First Time." *CNN*. https://www.cnn.com/2019/06/14/tech/sundar-pichai-google-antitrust/index.html (March 21, 2022).

Contradictory Political Landscape

In 2014, President Obama urged the FCC to reclassify broadband under Title II of the Communications Act of 1934.[94] He was concerned about internet fairness and recognized that the only way to avoid the textual patchwork that was created from 2002-2013 (see *Figure 6*) was to fully reclassify ISPs under Title II. A 3-2 FCC voted to do just that and officially created a codified net neutrality jurisdiction. Although net neutrality is about equality of access, its reclassification of ISPs under Title II gives them many more regulatory avenues to control the quality of content on the internet in the same way it did for television and radio. From this point, net neutrality and intermediary liability were trailers tethered to the same truck, Title II. If a future administration wanted to modify content moderation jurisdiction, they would be affecting net neutrality and vice versa. The problem: these two issues cannot coexist under modern political priorities adopted by the two major parties.

Section 230 has been likened to both a sword and a shield in intermediary liability case law. Section 230(c)(1) shields from liability for the content they choose to leave up, so long as it's provided by a user. Section 230(c)(2) is a sword that allows sites to take down content provided by users. Democrats would like to weaken the shield and strengthen the sword. They feel sites can do more under Section 230(c)(2) to police for objectionable content and benefit too much from the shield of Section 230(c)(1). Republicans would like to strengthen the shield and weaken the sword. Republicans believe that sites have too much power under Section 230(c)(2) and allege they use it to censor right-leaning voices. One place we've seen manifestations of this political disagreement is in the lack of coordination in bringing bills to the floor. In the 116th Congress (2019-2020), twenty-six different bills were brought to the floor of the House or Senate

[94] The Obama White House. 2014. President Obama's Statement on Keeping the Internet Open and Free https://youtu.be/uKcjQPVwfDk (March 4, 2022).

that would have amended Section 230 (Brannon and Holmes 2021). This issue has become so ideologically complex that a team of researchers, called Future Tense[95] have started a constantly updating list of such bills. They divide all the bills into four categories: full repeal, limiting the scope, imposing new obligations, and modifying Good Samaritan protections.[96] Their research has tried to make sense of this legislative landscape but to ISPs and government regulators alike, the landscape for legislative reform looks more like a Jackson Pollock painting than a van Gogh or a Monet.

Concerned with the economic impact that regulation of networks was having on innovation and competition, the Trump administration implored the FCC, under Ajit Pai, to repeal net neutrality.[97] They would eventually do so and they did it in part by linking it to Section 230 and, more specifically, the "unfettered" clause in Section 230(b)(2). In fact, the FCC's 2015 Restoring Internet Freedom Order cited Section 230 eighty-four (84) times as statutory evidence that the FCC had no control over broadband internet equity and ISPs (referred to at times as "edge providers"). However, by forfeiting jurisdiction, Trump lost the regulatory arm that could have controlled content moderation practices by ISPs. It's in this way that the politics of the issue are divided. If Republicans want to prevent alleged internet censorship, they need to establish jurisdiction by conceding ISPs are under Title II, but doing so would mean they have to guarantee Title II protections to ISPs (like equality of access) and therefore reinstate net neutrality. As a result of this regulatory patchwork, it's become hard for Republicans to have their and eat it too. Conversely, Democrats are fearful that overly regulating ISPs could have deleterious effects on counterspeech (speech that seeks to dilute extremist points of view and

[95] A collaboration of researchers at Slate, New America, Arizona State University, American University, Stanford, and UNC Chapel Hill.
[96] Anand, Meghan et al. 2021. "All the Ways Congress Wants to Change the Internet's Most Fundamental Law." *Slate Magazine*. https://slate.com/technology/2021/03/section-230-reform-legislative-tracker.html (March 5, 2022).
[97] Collins, Keith. 2017. "Why Net Neutrality Was Repealed and How It Affects You." The New York Times. https://www.nytimes.com/2017/12/14/technology/net-neutrality-rules.html (March 10, 2022).

bring the debate back to a more moderate perspective), and yet simultaneously want a highly regulated internet landscape under net neutrality. FCC Commissioner Jessica Rosenworcel put it well when she said, “In the end, it’s not just the hypocrisy that disappoints, or the intellectual contortions required to make sense of this. It’s the dishonesty. It can’t be that the FCC points to Section 230 to disavow authority over broadband but then uses the same law to insist it can turn around and serve as the President’s speech police."[98]

Trump attempted to use a backdoor route around this legislative dissonance in his 2018 “Executive Order on Preventing Online Censorship." It directed the Secretary of Commerce *to direct* the National Telecommunications and Information Administration (NTIA) *to direct* the FCC to *consider* rulemaking.[99] As anyone can tell, it’s not particularly the most ironclad of legal orders and would later be held up by two pending court challenges over its constitutionality before being ultimately repealed by Biden in 2021.[100, 101]

Aside from the divided political consensus over content moderation itself, the patchwork has forced net neutrality and content moderation to ride the same statutory bus, Title II. Even if there was political consensus to clarify regulatory authority over ISPs, doing so would likely require throwing net neutrality to the wayside. As a result of the chain novel regulatory structure, legislators have been backed into some corners: is it the policy of the U.S. to strongly regulate content or promote network growth through deregulation? The tradeoffs come as a result of the

---

[98] Eggerton, John. 2020. “FCC Dems Can't Reconcile ISP Dereg, Section 230 Initiative.” *Multichannel News*. https://www.nexttv.com/news/fcc-dems-cant-reconcile-isp-dereg-section-230-initiative (March 10, 2022).
[99] Executive Order No. 13925, 85 Fed. Reg. 34079 (2020)
[100] Kelly, Makena. 2020. “President Trump's Social Media Order Will Endanger Voting Rights, New Lawsuit Claims.” *The Verge*. https://www.theverge.com/2020/8/27/21404554/trump-social-media-order-section-230-rock-the-vote-free-press (March 10, 2022).
[101] Spangler, Todd. 2020. “Lawsuit Says Donald Trump's Executive Order Targeting Twitter, Facebook Violates First Amendment.” *Variety*. https://variety.com/2020/digital/news/trump-sued-over-social-media-executive-order-first-amendment-1234623504/ (March 10, 2022).

patchwork regulatory regime and are only further exacerbated by the rise of platform power amidst that patchwork.

Conclusion

Of all the thousands of leads that the U.S. intelligence community tracks down every day, Brad Rukstales, "dog dad" and master griller from suburban Chicago, was nowhere near the top of their list. His exemplary behavior as a charitable father of two and self-aware critic of social media and alternative news are the antithesis of what we would expect when we're presented with pictures of January 6th insurrectionists. We're more disposed to believe someone that looks like the QAnon Shaman with his face painted and head adorned with a menacing fur pelt. But if the internet has taught us anything over the past three decades, looks aren't everything. Social media platforms and the algorithms that they employ have had a big role in curating content on the internet at a more individual and interactive level than previously available in other mediums like radio, television, or print journalism.

Contrary to these other mediums, the internet was uniquely classified as neither a distributor, nor a speaker, nor a publisher. A new class of intermediary liability was created in the name of bolstering the development of this nascent industry. The deregulatory framework employed by the U.S., in contrast to the approach taken by the European Union, incentivizes a higher quantity of content to stay up online. The European approach, on the other hand, would rather preemptively remove objectionable content than let the free marketplace of ideas and "counterspeech" dilute the objectionable material. This U.S. approach is codified in Section 230 which immunizes both content left up and good faith efforts to remove objectionable content. The problem is that this broad immunity is not coupled with a clear regulatory counterpart that

can incentivize ISPs to use their Section 230(c)(2) privilege to take bad content down without liability.

The FCC would be an ordinarily ideal regulatory agency to do just that. They already have fine structures they use on "common carriers" like radio, television, and telecommunications to compel self-moderation and punish attempts to circumvent existing decency guidelines. However, the conflicting legislative history and text of Section 230 create uncertainty of whether the FCC has jurisdiction over the internet. On one hand, Section 230 is within a larger piece of legislation (the Communications Act of 1934) that explicitly names the FCC as its regulatory agent – hence the jurisdiction over other mediums mentioned above. On the other hand, the text of the section itself, the floor debates about its passage, and the omission of the pro-FCC Exon amendment all suggest that Congress did not want the FCC involved in any way in internet oversight.

This contradiction and the patchwork regulatory nature of a politically appointed agency like the FCC creates a sort of "chain novel dilemma." Each successive administration of the FCC has to work with the chapters previously written and try to craft the chapter they've been empowered to write in a way that accomplishes their goals without trampling too much on the characters and storylines written by administrations before them. This unclear regulatory jurisdiction largely precluded the FCC from being involved in enforcing decency standards or incentivizing ISPs to use their Section 230(c)(2) privileges for the past twenty-five years. Only in 2015, was FCC jurisdiction finally clarified when the internet was moved under Title II of the 1934 Act rather than Title I (a privilege that was only recognized by the Supreme Court almost ten years after the passage of the CDA). From this point, however, the regulation of the internet

would leave the hands of FCC attorneys and become a more political issue that sat at the doorstep of Presidential administrations.

Trump's reversal in 2017 of Obama's net neutrality doctrine would cite Section 230's "unfettered" clause extensively in justifying the fact that the Federal government had no jurisdiction over broadband internet and the ISPs that used it. From this point, net neutrality and content moderation became trailers attached to the same truck, Title II. As a result of this complicated political landscape, it becomes difficult to reform either policy without affecting the other. Democrats want to see a strong net neutrality but not see the chilling effects on free speech that would come from regulating the platforms in the way Trump suggested. Republicans would like to dismantle net neutrality and see the broadband industry flourish, but in doing so would have to relinquish their jurisdiction over ISPs and their ability to prevent alleged censorship of conservative voices.

The question then becomes: "why don't ISPs just pick up the slack that the FCC has been unable to do and Congress is unable to change?" The answer: it's too cost-prohibitive to do so. Sites don't operate under the same business model now as when they did in the 90s. There is still demand for a clean platform, but technology companies can form a sort of oligopoly which allows them to set their content moderation standards relatively low so long as they keep parity with their competitors. The social media sector has grown tremendously in power both financially and politically since the 1990s and especially since the late 2010s. They now operate one of the largest lobbying industries in US politics and have cemented their value in the US intelligence community in a way that they are indispensable and can evade unwanted regulatory attempts despite the theater of being "in favor of new regulations." This is largely the reason why no meaningful, well-drafted content moderation legislation has been passed in over two decades.

SESTA/FOSTA was a farce in the eyes of the legal community and represents the only major edit to the foundational law of internet technology in the United States.

While it's tough to prove causally that Rukstales' radicalization along with the radicalization of the many other domestic terrorists was a result of platform ambivalence, I have argued that Section 230 and its deregulatory framework have created incentive structures that increase the likelihood of such radicalizing content remaining online for impressionable eyes to see. The U.S. set out in 1996 to prevent the internet from becoming a haven of objectionable content visible for kids and clearly knew the risks associated with a deregulated internet. But they chose to counter this objectionable content not with preemptive removal, like in Germany, but by flooding the marketplace of ideas with counterspeech to dilute the prevalence of the bad content. This is not a bad plan in and of itself. But it failed to account for the fact that ISPs would eventually become more powerful (financially, politically, socially) than they were in the 1990s, no longer have an incentive to take the proper content down when it had the chance, and create an algorithm that curates this content for people circumventing the benefits of counterspeech. Without a nimble and proactive regulatory agency, like the FCC, to update guidelines and keep ISPs in line, the inmates begin running the asylum; content on the internet is regulated not by elected officials but by corporations and others who stand to benefit from inflammatory content and the ad revenue it brings.

Chapter 4: Alternatives and Interlopers

As the previous chapters have made clear, the U.S. and Germany exist on two linearly different ends of a spectrum. While the U.S. regime set out to prioritize the growth of the internet sector through a deregulatory framework and the German regime set out to prioritize the wellbeing of citizens through a preemptive framework, both the U.S. and Germany share one thing in common: their predilection for the freedom of expression and internet privacy. Although I place them at opposite ends of a spectrum, that spectrum in and of itself still operates in complete juxtaposition to the regulatory regimes undertaken by authoritarian nation-states like Russia and China. The heavy-handed internet censorship of Eastern autocracies could certainly cure many of the woes bemoaned in previous chapters related to radicalization and its translation into violence.

Although my larger argument critiques the conceptualization of the internet within Western democratic regulatory regimes, that does not mean Western democracy and its principles are incompatible with a more comprehensive and responsible regulation of internet speech. An alternative to democratic regulation would be dictatorial control over the pathways of information. While I will highlight in this chapter that such is highly effective in addressing the dependent variable (violence stemming from radicalization) in my argument, this chapter serves to juxtapose Western democratic regulation against their much worse alternatives: the internet censorship of authoritarian countries like Russia and China.

Before doing so, it's important to conceptualize why these alternatives are important to the argument. Although the frameworks of these two countries are not alternatives that democracies like the U.S. and Germany would be likely to adopt, they represent alternatives that nonetheless affect the U.S. and Germany. My previous chapters have not compared apples to

oranges; my argument does not evaluate whether a deregulatory framework would have worked in Germany or vice versa. This is not a comparative analysis between the U.S. and Germany. This thesis has offered an analysis of each regulatory framework and how it functions amidst the larger regulatory regime at play in the country. But it becomes important to realize that the internet is a transnational industry. With few exceptions, almost all content is available in some way everywhere at the same time. The success of the U.S. and Germany in curbing radicalizing content therefore cannot only come from reconciling their incongruous frameworks within the larger regimes but also must take into account the risk posed by authoritarian regimes trespassing within the openness of democratic internet regulations.

As a result of this transnational nature, the shared support for freedom of expression in the two democracies comes at a cost. The 'fallacy of liberalism' posits that by opening up the doors for free expression, you inherently leave the doors open for bad actors looking to exploit your openness. By having comparatively less restrictive liability regimes than Russia and China, the U.S. and Germany allow foreign interlopers to join their domestic marketplaces of ideas. As we saw in the 2016 U.S. election, incentivizing a free internet allows foreign meddling where fake identities can post propaganda that could influence the behavior of millions.[102]

Russia: "The Psychological Firewall"

Russia's crackdown on internet speech is a fairly recent concept compared to China's. Before 2007, Russian social media was outside the purview of Putin's master plan. He targeted mass media like television and newspapers and put them under state control (Nisbet et al 2017).

[102] Kim, Young Mie. 2020. "New Evidence Shows How Russia's Election Interference Has Gotten More Brazen." *Brennan Center for Justice*. https://www.brennancenter.org/our-work/analysis-opinion/new-evidence-shows-how-russias-election-interference-has-gotten-more (March 20, 2022).

The oversight of social media at the time was left up to a court-order standard like the U.S. but a less stringent one. Blocked sites could get around weak restrictions with technological fixes and were often only blocked in one region of Russia and not the whole country (Soldatov 2017). The functionality and popularity of modern social media platforms did not explode until around 2008. Before 2009, the "internet" was limited to primarily "message switching" where messages could be shared between terminals by sending messages to designated usernames, IP addresses, or onto message boards but there was no automatic curation. In 2009, however, came the advent of algorithms. Facebook created the first algorithm (called EdgeRank) which curated content using three factors (affinity, rank, decay). Two years later, that would be replaced with an algorithm called "News Feed" that used 100k factors, and the algorithm would only grow from there.[103] Suffice it to say, Putin was more concerned with the power of television and radio than the limited functionality of social media.

This is until the 2011-2012 riots. These riots were to protest the amateur videos posted on social media sites like Facebook and Twitter showing ballot stuffing and other voter fraud in the 2011 Parliamentary victory for Putin's party.[104] As a result of Facebook and other social medias' heavy hands in rallying support for the riots against him, Putin began to crack down on platforms in the same way he had previously done by nationalizing television and radio stations (Nisbet et al 2017). Because Russia had not always had a clamped down censorship system, Putin had to get creative on how to achieve his policy priorities. Russia uniquely put the onus on networks rather than ISPs or the government to moderate content. Putin characterized the internet as

---

[103] McGee, Matt. 2013. "EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close to 100k Weight Factors." *MarTech*. https://martech.org/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors/ (March 20, 2022).

[104] de Carbonnel, Alissa. 2011. "Insight: Social Media Makes Anti-Putin Protests 'Snowball.'" *Reuters*. https://www.reuters.com/article/us-russia-protests-socialmedia/insight-social-media-makes-anti-putin-protests-snowball-idUSTRE7B60R720111207 (March 20, 2022).

something being used by both domestic extremists and alleged "that foreign powers use the Internet to pursue political and military objectives against Russia" thereby necessitating his creation of a "bulletproof vest" of protection for Russians (Nisbet et al 2017, 960). Putin alleged that he was building a firewall for the psychological protection of all Russians and framed his actions as something akin to Germany's defensive democratic model.

The government under the Roskomnadzor (Federal Service for Supervision of Mass Communication) contracted NGOs like Safe Internet League and Mediaguard to blacklist sites and oversee the enforcement of the bans (Soldatov 2017). These NGOs and government agencies then used networks as the enforcement mechanism. ISPs (ie: Twitter) use networks (ie: AT&T or Verizon) to transmit their content after their users request it. Since no ISPs were domiciled in Russia at the time, Putin had to resort to the networks. Networks could lose their license for negligently allowing a blacklisted site to be transmitted into Russia. In addition, Russia imposed the cost of such enforcement on the networks themselves. All the tracking and filtering equipment was paid for and maintained by the networks (Soldatov 2017). This proved rather effective: by using the networks as a choke point, Putin could more easily control the spread of information. Networks can filter for content that is posted on all sites that go over their connection, whereas ISPs can only monitor for blacklisted content on their own sites. As a result, networks were able to catch hundreds of mirrors of the same content on hundreds of sites simultaneously.

Another problem: the Russian intelligence community helped draft the 2012 laws and inextricably linked surveillance and content moderation. There was little to no oversight on tapping the content of these platforms. They could access network data without network permission. While a warrant was necessary, FSB (Federal Security Service) employees only had

to ‘show’ the warrant to their supervisor in order to tap networks and access the personal data of any Russian citizen (Soldatov 2017). While this prevents radicalizing content from being amplified and reproduced amongst fringe actors, it also hinders the ability of citizens and journalists to receive and transmit information detrimental to the hegemony of the Russian oligarchy. Nisbet et al (2017) define this phenomenon as “window opening” whereby Russians are prevented from seeing the outside world and realizing the oppression of Vladimir Putin’s regime (959). Just recently, Russia shut down access to Twitter and Facebook, essentially cutting off Russians from information about Ukraine and allowing Putin to solely control the messaging of his invasion.[105] While a byproduct of this censorship is a logically lower incidence of domestic terror, it seems hardly worth vacating countless other hallmark principles of democratic governance to achieve it.

China: “The Information Curtain”

In a more extreme case than Russia, China has banned Facebook and Twitter since 2009 and never had the pre-2007 internet freedom that Russians enjoyed. It has always been a bastion of internet censorship but goes about it differently. China also logically has fewer instances of domestic terrorism but that’s because it doesn’t let any outside ideas violate Chinese digital sovereignty or as Hillary Clinton has called it “The Information Curtain” (Kyriakopoulou 2011, 20). While Russia (barring this month’s events) has historically supported Western platforms, China instead offers state-sponsored platforms as alternatives. Sites like Weibo and WeChat are

[105] Yuan, Li. 2022. “China's Information Dark Age Could Be Russia's Future.” *The New York Times*. https://www.nytimes.com/2022/03/18/business/chinas-russia-information.html (March 30, 2022).

platforms completely overseen by the Chinese government regardless of whether the user is in China or another jurisdiction.[106]

China similarly justifies its censorship as a necessary precaution for the national security of its citizens. China argues that state secrets are being shared on these platforms and China has a vested interest in monitoring them.[107] They use their army of government censors to monitor all this content and preemptively restrict access to any material that may jeopardize the hegemony of the Chinese state (Nisbit et al 2017). China clamped down on internet speech around the same time as Russia by rewriting its statutes about state secrets to be more vague and malleable in 2010. They can now claim jurisdiction over nearly all types of surveillance even on networks and communications that take place outside the state.

TikTok is a popular example of how an authoritarian nation can trespass within the openness and freedom of Western democratic regulatory frameworks. In 2020, Trump sought to force TikTok's parent company (which is Chinese) to merge with a U.S. parent company if it wanted to remain available in the United States. He and his then-Secretary of State alleged that TikTok was feeding information back to its parent company which was sharing the data directly with the Chinese Communist Party.[108] In addition to this surveillance, China has a history of flooding the U.S. with propaganda by paying influencers to peddle incorrect stories about their human rights, the origins of COVID, and other lies that are in the interest of Chinese hegemony.[109] By having this censorship, China too has the opportunity to pollute the U.S. social

---

[106] Feng, Emily. 2019. "China Intercepts WeChat Texts from U.S. and Abroad, Researchers Say." *NPR*. https://www.npr.org/2019/08/29/751116338/china-intercepts-wechat-texts-from-u-s-and-abroad-researcher-says (March 21, 2022).

[107] Xu, Beina, and Eleanor Albert. 2017. "Media Censorship in China." *Council on Foreign Relations*. https://www.cfr.org/backgrounder/media-censorship-china (March 30, 2022).

[108] Rodriguez, Salvador. 2021. "Tiktok Insiders Say Social Media Company Is Tightly Controlled by Chinese Parent ByteDance." CNBC. https://www.cnbc.com/2021/06/25/tiktok-insiders-say-chinese-parent-bytedance-in-control.html (March 30, 2022).

[109] Seitz, Amanda, Eric Tucker, and Mike Catalini. 2022. "How China's Tiktok, Facebook Influencers Push Propaganda." San Francisco Chronicle.

media landscape despite our best regulatory attempts. Such is the price that Western democracies pay for promoting a free internet.

Conclusion

Chapters 2 and 3 critiqued the current conception of internet speech regulation by Western democracies. It put the U.S. and Germany on opposing sides of a linear spectrum. This chapter sought to contextualize where that spectrum falls in relation to the other world superpowers and the regulatory regimes they employ. In addition to contextualizing the earlier chapters, this chapter highlighted the anti-democratic regime that necessarily has to accompany the harsher framework that would be required to appropriately squash the reproduction of objectionable content. This chapter illustrates the inherent trade-offs associated with the type of comprehensive regulatory reform necessary to eliminate objectionable content. If the U.S. and Germany want to maintain their democratic ideals, there are costs to be paid. They need to accept the risk that malicious actors will seek to influence their free internet and they need to understand that you'll never be able to live in a world without a little objectionable or radicalizing content unless you aggressively censor the free flow of information.

https://www.sfchronicle.com/news/article/How-China-s-TikTok-Facebook-influencers-push-17038060.php (March 30, 2022).

Chapter 5: Conclusion

This thesis identified reasons why two beacons of democracy, the U.S. and the European Union, not only chose mutually exclusive and linearly different regulatory frameworks but, more importantly, why neither framework proved to be better suited in responsively addressing the rise of radicalizing content on the internet. I argue that there are two narratives present. The micro-narrative has to do with the perverse incentive structures created as a byproduct of either the text of each bill or its implementation. This is bad in and of itself and has created demonstrable barriers to each country's ability to curb this objectionable content. The macro-narrative goes further and alleges that this cannot be blamed only on bad policy drafting or bad implementation. Rather, each of these case studies represents an incongruous relationship between the framework they intended and the regulatory regime that they either had in place or created to support that framework.

From the perspective of Germany, objectionable content drowns out well-intentioned counterspeech. This problem of the "heckler's veto" is a result of (1) Germany's defensive democratic underpinnings, (2) the fact that the E.U. took some provisions of the American D.M.C.A. but not all of it, (3) the unclarified statutory conflict between a sites' duty of care and its immunity from constant monitoring, and (4) the lack of oversight by the European Union on the accurate transposition of its directive into German national legislation. Each successive iteration of intermediary liability reform from 1997 to 2021 admittedly did clarify some aspects of the previous legislation but failed to address the incongruity that existed between the *preemptive* framework and the German regulatory regime that supported it.

The *preemptive* framework came as no surprise, Germany has been a leader in the defensive democratic model since the end of WWII. Their priority was to root out objectionable

content with force before it could take hold and damage national unity or the sanctity of the Constitution. Such defensive democratic ideals had existed before the internet via their Constitutional right to curtail speech if it threatens the unity of the state and such defensive democratic ideals continued with their notice-and-takedown policies that governed internet speech. This preemptive framework, however, was passed by the European Commission without harmonizing the underlying laws that the framework sought to enforce. The standard of proof for proving defamation/libel/slander, incitement, sexual harassment, threats to national sovereignty, and many other statutes were not harmonized from member state to member state. While the E-Commerce Directive made it clear that ISPs had to evaluate reported content in a certain period of time, it allowed for there to exist essentially twenty-seven (27) different sets of laws that could be applied. This created an incentive to just remove the content without a very critical eye. Although the ECD established the country of origin principle, it didn't create a clear pathway for how Germany can force Italy to remove (in a timely fashion or with any certainty) objectionable content posted on an Italian ISP that had made its way onto the smartphone of a citizen in Berlin.

To have properly implemented the preemptive framework, the E.U. would have needed to pass the ECD as a regulation and not a directive. They were experiencing a similar conflict with the 1995 Data Protection Directive before eventually reenacting it as the General Data Protection Regulation (GDPR) in 2016. Europe recognized the transnational nature of the internet in the policy statements of the ECD itself, and yet still chose to pass this as a directive. This directive-based regime made supporting the European Union and Germany's preemptive framework extremely difficult.

From the perspective of the United States, counterspeech can flourish more but so too can objectionable content. Objectionable content, therefore, exists as a smaller proportion of all content, but in larger quantities compared to the German case study. The text of Section 230 gives nearly blanket immunity to sites for both the content that they leave up and that which they takedown. As a result of shifting business models and a financial incentive to keep inflammatory posts that didn't exist in the early days of ISPs, objectionable content can flourish irrespective of how hard well-intentioned people try to dilute it with counterspeech.

The deregulatory framework existed as a way of allowing the nascent internet industry to blossom. This should come as no surprise either, the promise of the First Amendment and the United States' natural inclination to support capitalist, free-market solutions aligns well with this largely hands-off framework. The macro-narrative becomes problematic however when you examine what it would have taken to responsibly implement a deregulatory framework. Vesting redress powers with the court (rather than ISPs like in Germany) is great insofar as it only rights injustices after the plaintiffs have been wronged. There is no proactive regulatory agent, even if non-binding, to liaise with ISPs and other stakeholders about mutual ways to make the internet a cleaner place without having to retroactively litigate all of the effects of objectionable content. Many drafters of the bill, like Sen. Exon, believed the FCC could fill that vital role. But Cox and Wyden's insistence on having a court order standard without a complimentary and nimble voice of proactive guidance won the day. They believed ISPs were exceptional from other media methods despite ISPs assuming the same functional roles as televisions, radios, and newspapers. The U.S. wanted a deregulatory standard but believed that such could only exist without a proactive overseer. As a result, ISPs have exponentially increased in power in the U.S. and can essentially now set their own regulations.

Finally, I juxtapose and contextualize the debate over Western democratic governance with a brief analysis of authoritarian interlopers. I point out the costs that come as a result of having a free and open internet and the unfortunate fact that achieving success (eliminating all objectively objectionable content) under my puzzle naturally requires countries to sacrifice democratic ideals. One of the fallacies of liberalism is that democracy can only flourish if a little bit of objectionable content does as well.

In making my argument critiquing democratic conceptualizations of internet speech regulation, I am straying from the existing literature which is oversaturated with legal solutions to reform the legislation. I am instead contributing a structurally comparative analysis that conceptualizes our failure to curb radicalizing content as one that results from the origins of the bills and not any exogenous factors. Western democracies believe that updating these pieces of legislation can be done with additional legislation. I offer instead the perspective that democracies will only be able to move the needle on policing objectionable content if they look at how their desired frameworks fit into the original regimes by which they were surrounded.

With this being said, additional research is nonetheless necessary to discover what else, in addition to internet regulation, is creating this culture of objectionable content. The internet is simply the canvas on which society is painting. I would suggest there are cultural issues at play with regards to the inclination of users to believe things that are distortions of reality. Feelings of alienation, isolation, poverty, oppression and general distrust in institutions among many other emotions likely drive people to seek out ways of distorting their suboptimal reality. For example, there exists a growing body of literature on *creepypasta*, or the sharing of clearly fictional horror stories online that they allege are based on real stories. Users choose to believe the implausible is plausible in these stories because it offers an escape from the mundane realities of life.

Conceptions of reality and its validity are further obfuscated by the government itself. When the U.S. government acknowledged in 2021 the fact that the armed forces surveilled and tracked UFOs, people's conceptions of reality were shattered. UFO deniers felt like the silly ones and it muddied the waters between what the government said, what "rational" people chose to believe, what "rational" people should be believing, and what is the actual reality.

The impacts of this question are unmistakable. At risk here is a fundamental aspect of what differentiates humans from other animals and what has allowed us to progress as a society for thousands of years, our ability to agree on reality. We may not always agree on what to build, but we've predominantly had the human decency to operate with the same tools in the toolbox. That's not to say that addressing this problem of internet speech regulation is going to be a panacea to returning us to a pre-truth era. That's also not to say that truth couldn't continue to be distorted even with a refresh on internet speech regulations. But what is definitive is the fact that this problem can only get worse if we do nothing. Under the status quo, life will continue to imitate art. Our conceptions of reality and truth will be about as malleable as these sentences in Google Docs. If we don't want our next generations growing up in a world whose metaphorical font, size, color, and spacing can be distorted with little effort or oversight then we must address this issue of internet distortion before we gradually continue to lose control of it.

Bibliography

Baistrocchi, Pablo A. 2002. “Liability of Intermediary Service Providers in the EU Directive on Electronic Commerce.” *Santa Clara Computer and High Technology Law Journal* 19(1): 111–30.

Bechtold, Eliza. 2020. “Terrorism, the Internet, and the Threat to Freedom of Expression: The Regulation of Digital Intermediaries in Europe and the United States.” *Journal of Media Law* 12(1): 13–46. doi: 10.1080/17577632.2020.1760474.

Bilazarian, Talene. 2019. “Countering Violent Extremist Narratives Online: Lessons from Offline Countering Violent Extremism.” *Policy & Internet* 12(1): 46–65. doi: 10.1002/poi3.204.

Brannon, Valerie C, and Eric N Holmes. 2021. “Section 230: An Overview.” *Congressional Research Service*: 1–54.

Candeub, Adam. 2020. “Bargaining for Free Speech: Common Carriage, Network Neutrality, and Section 230.” *Yale Law Review* 11: 391–433.

Cheah, Michael. 2020. “Section 230 and the Twitter Presidency.” *Northwestern University Law Review* 115: 199–222.

Chen, Jim. 2001. “The Authority to Regulate Broadband Internet Access Over Cable.” *Berkeley Technology Law Journal* 16(2): 677–728.

Claussen, Victor. 2018. “Fighting Hate Speech and Fake News: The Network Enforcement Act (NetzDG) in Germany in the Context of European Legislation.” *Journal of Media Law*: 110–36.

Cohen-Almagor, Raphael. 2015. *Confronting the Internet's Dark Side: Moral and Social Responsibility on the Free Highway.* Cambridge, MA: Cambridge University Press.

Cohen-Almagor, Raphael. 2017. “The Role of Internet Intermediaries in Tackling Terrorism Online.” *Fordham Law Review* 86: 425–53.

Collins, Matthew. 2001. *The Law of Defamation and the Internet.* 1st ed. Oxford: Oxford University Press.

Comstock, Earl W, and John W Butler. 2000. “Access Denied: The FCC's Failure to Implement

Open Access to Cable as Required by the Communications Act." *Journal of Communications Law and Policy* 8(1): 5–22.

Conway, Maura, Ryan Scrivens, and Logan Macnair. 2019. "Right-Wing Extremists' Persistent Online Presence: History and Contemporary Trends." *International Centre for Counter-Terrorism*: 1–24. doi: 10.19165/2019.3.12.

Conway, Maura. 2020. "Routing the Extreme Right: Challenges for Social Media Platforms." *The RUSI Journal* 165(1): 108–13. doi: 10.1080/03071847.2020.1727157.

Crawford, Henry E. 1997. "Internet Calling: FCC Jurisdiction over Internet Telephony." *Journal of Communications Law and Policy* 5(1): 43–54.

Cusumano, Michael A., Annabelle Gawer, and David Yoffie. 2021. "Can Self-Regulation Save Digital Platforms?" *SSRN Electronic Journal*. doi: 10.2139/ssrn.3900137.

Determann, Lothar. 1999. "Case Update: German CompuServe Director Acquitted on Appeal." *Hastings International and Comparative Law Review* 23(1): 109–24.

Esbin, Barbara, and Adam Marcus. 2009. "The Law Is Whatever the Nobles Do: Undue Process at the FCC." *Journal of Communications Law and Policy* 17(2): 535–656.

Esbin, Barbara. 1999. "Internet over Cable: Defining the Future in Terms of the Past." *Journal of Communications Law and Policy* 7(1): 37–118.

Esbin, Barbara. 2009. "Jurisdiction: The $64,000 Question." *The Progress and Freedom Foundation* 5(12): 1–7.

Fried, Neil. 1999. "Dodging the Communications Decency Act When Analyzing Libel Liability of On-Line Services." *Columbia Science and Technology Law Review* 1(1): 1–44.

Fried, Neil. 2021. "The Myth of Internet Exceptionalism: Bringing Section 230 into the Real World." *American Affairs Journal*. https://americanaffairsjournal.org/2021/05/the-myth-of-internet-exceptionalism-bringing-section-230-into-the-real-world/ (March 29, 2022).

Friedman, Jonathan A, and Francis M Buono. 2000. "Limiting Tort Liability for Online Third-Party Content under Section 230 of the Communications Act." *Federal Communications Law Journal* 52(3): 647–66.

Ganesh, Bharath, and Jonathan Bright. 2020. “Countering Extremists on Social Media: Challenges for Strategic Communication and Content Moderation.” *Policy & Internet* 12(1): 6–19. doi: 10.1002/poi3.236.

Ganesh, Bharath. 2018. “The Ungovernability of Digital Hate Culture.” *Journal of International Affairs* 71(2): 30–49. https://www.jstor.org/stable/26552328.

Gilens, Martin, and Benjamin I. Page. 2014. “Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens.” *Perspectives on Politics* 12(3): 564–81. doi: 10.1017/s1537592714001595.

Gillespie, Tarleton. 2020. “Content Moderation, AI, and the Question of Scale.” *Big Data & Society* 7(2): 1–5. doi: 10.1177/2053951720943234.

Gioia, Andrew. 2009. “FCC Jurisdiction over ISPs in Protocol-Specific Bandwidth Throttling.” *Michigan Telecommunications and Technology Law Review* 15(2): 517–42.

Goldman, Eric. 2017. “The Ten Most Important Section 230 Rulings.” *Tulane Journal of Technology and Intellectual Property* 20: 1–10.

Gross, Emanuel. 2003. “Defensive Democracy: Is It Possible To Revoke the Citizenship, Deport, or Negate the Civil Rights of a Person Instigating Terrorist Action Against His Own State?” *UMKC Law Review* 72(1): 51–122.

Guhl, Jakob, and Jacob Davey. 2020. “A Safe Space to Hate: White Supremacist Mobilization on Telegram.” *Institute for Strategic Dialogue*. https://www.isdglobal.org/wp-content/uploads/2020/06/A-Safe-Space-to-Hate2.pdf (March 29, 2022).

Hamm, Mark S, and Ramon F Spaaij. 2017. *The Age of Lone Wolf Terrorism.* New York, NY: Columbia University Press.

Hoeren, Thomas, and Silviya Yankova. 2012. “The Liability of Internet Intermediaries - The German Perspective.” *International Review of Intellectual Property and Competition Law* 43(5): 501–31.

Hoeren, Thomas. 2000. “Electronic Commerce and Law – Some Fragmentary Thoughts on the Future of Internet Regulation from a German Perspective.” *Computer Law and Security* 16(1): 113–17.

Homeland Security Institute. 2009. “The Internet as a Terrorist Tool for Recruitment and Radicalization of Youth.” *Homeland Security Institute.* https://perma.cc/2K65-62QQ (March 29, 2022).

Institute for Economics & Peace. 2019. "Global Terrorism Index 2019: Measuring the Impact of Terrorism," *Vision of Humanity.* http://visionofhumanity.org/reports (October 22, 2021)

Jacobides, Michael G. 2020. “Regulating Big Tech in Europe: Why, so What, and How Understanding Their Business Models and Ecosystems Can Make a Difference.” *SSRN Electronic Journal*: 1–42. doi: 10.2139/ssrn.3765324.

Julià-Barceló, Rosa, and Kamiel J Koelman. 2000. “Intermediary Liability in the E-Commerce Directive: So Far So Good, But It's Not Enough.” *Computer Law & Security Review* 16(4): 231–39. doi: 10.1016/s0267-3649(00)89129-3.

Keller, Daphne. 2018. “The Right Tools: Europe's Intermediary Liability Laws and the EU 2016 General Data Protection Regulation.” *Berkeley Technology Law Journal* 33(1): 287–364.

Kennedy, Leonard J, and Lori A Zallaps. 1999. “If It Ain't Broke... The FCC and Internet Regulation.” *Journal of Communications Law and Policy* 7(1): 17–36.

Kightlinger, Mark F. 2003. “A Solution to the Yahoo! Problem? The EC E-Commerce Directive as a Model for International Cooperation on Internet Choice of Law.” *Michigan Journal of International Law* 24(3): 719–66.

Koehler, Daniel. 2014. “The Radical Online: Individual Radicalization Processes and the Role of the Internet.” *Journal for Deradicalization* (1): 116–34.

Kosseff, Jeff. 2019. *The Twenty-Six Words That Created the Internet.* Ithaca, NY: Cornell University Press.

Kuczerawy, Aleksandra. 2015. “Intermediary Liability & Freedom of Expression: Recent Developments in the EU Notice & Action Initiative.” *Computer Law & Security Review* 31(1): 46–56. doi: 10.1016/j.clsr.2014.11.004.

Kyriakopoulou, Kalliopi. 2011. “Authoritarian States and Internet Social Media: Instruments of Democratization or Instruments of Control?” *Human Affairs* 21(1): 18–26. doi: 10.2478/s13374-011-0003-y.

Levi, Lili. 2008. “The FCC's Regulation of Indecency.” *Freedom Forum Institute* 7(1): 1–95.

https://www.freedomforuminstitute.org/wp-content/uploads/2016/10/FirstReport.Indecency.Levi_.final_.pdf (March 29, 2022).

Lyons, Daniel A. 2013. "Restoring Limits on the FCC's Ancillary Authority." *Free State Foundation Perspectives* 8(34): 1–8.

Martinet, Beatrice, and Reinhard J Oertli. 2015. "Liability of E-Commerce Platforms for Copyright and Trademark Infringement: A World Tour." *American Bar Association.* https://www.americanbar.org/groups/intellectual_property_law/publications/landslide/2014-15/may-june/liability-e-commerce-platforms-copyright-trademark-infringement-world-tour/ (March 29, 2022).

McChesney, Robert W. 2013. *Digital Disconnect: How Capitalism Is Turning the Internet Against Democracy.* New York, NY: The New Press.

Mehra, Salil, and Marketa Trimble. 2014. "Secondary Liability, ISP Immunity, and Incumbent Entrenchment." *American Journal of Comparative Law* 62(1): 685–705. doi: 10.5131/ajcl.2013.0041.

Michael, Jensen et al. 2018. "The Use of Social Media by United States Extremists." *National Consortium for the Study of Terrorism and Responses Terrorism.* https://www.start.umd.edu/pubs/START_PIRUS_UseOfSocialMediaByUSExtremists_ResearchBrief_July2018.pdf (March 29, 2022).

Middendorf, Lynn. 1997. "Internet Service Providers Immune from Liability for Third Party Defamation." *Loyola Consumer Law Review* 10(3): 212–15.

Nisbet, Erik C., Olga Kamenchuk, and Aysenur Dal. 2017. "A Psychological Firewall? Risk Perceptions and Public Support for Online Censorship in Russia." *Social Science Quarterly* 98(3): 958–75. doi: 10.1111/ssqu.12435.

Omand, David, Jamie Bartlett, and Carl Miller. 2012. "Introducing Social Media Intelligence (SOCMINT)." *Intelligence and National Security* 27(6): 801–23. doi: 10.1080/02684527.2012.716965.

Popiel, Pawel. 2018. "The Tech Lobby: Tracing the Contours of New Media Elite Lobbying Power." *Communication, Culture and Critique* 11(4): 566–85. doi: 10.1093/ccc/tcy027.

Ruane, Kathleen Ann. 2011. "The FCC's Authority to Regulate Net Neutrality After Comcast v. FCC." *Congressional Research Service*: 1–32.

Sheridan, David R. 1997. “Zeran v. AOL and the Effect of Section 230 of the Communications Decency Act upon Liability for Defamation on the Internet.” *Albany Law Review* 61(1): 147–80.

Soldatov, Andrei. 2017. “The Taming of the Internet.” *Russian Social Science Review* 58(1): 39–59. doi: 10.1080/10611428.2017.1275024333.

Tambini, Damian. 2021. “Reconceptualizing Media Freedom.” In *Regulating Big Tech: Policy Responses to Digital Dominance*, ed. Martin Moore. New York, NY: Oxford University Press. essay, 299–317.

Thompson, Marcelo. 2010. “The Sheriff of ‘Not-the-Internet:’ Reflections on Comcast Corp. v. FCC.” *Communications Law Review* 1(1): 1–18.

Tsesis, Alexander. 2017. “Social Media Accountability for Terrorist Propaganda.” *Fordham Law Review* 86: 605–31.

Walther, Samantha, and Andrew McCoy. 2021. “US Extremism on Telegram: Fueling Disinformation, Conspiracy Theories, and Accelerationism.” *Perspectives on Terrorism* 15(2): 100–124. https://www.jstor.org/stable/27007298.

Whiteley, Jason. 2000. “AT&T Corp v. City of Portland: Classifying Internet over Cable in the Open Access Fight.” *Brigham Young University Law Review* 1: 451–90.

Wiesner, Lucy. 2020. “Good Intentions and Unintended Consequences: SESTA/FOSTA's First Two Years.” *Temple Law Review* 93(1): 151–80.