

ARTICLE

***Anderson, Algorithms, and Section 230 After
NetChoice: The Risk of a New Moderator’s Dilemma***

*Jennifer Huddleston**

Table of Contents

Introduction..... 728

I. Understanding Recent Jurisprudence Around the Intersection of Content Moderation, Algorithms, Section 230, and the First Amendment 729

 A. What Does Section 230 Actually Say? 730

 B. How the Courts Have Explored Algorithms and Other Technologies as Speech and the Relationship to Section 230 Liability Protections..... 731

 1. *Gonzalez v. Google* and a Potential Missed Opportunity....732

 2. *Force v. Facebook*: Second Circuit Finds Algorithmic Content Moderation Covered by Section 230.....733

 3. *Anderson v. TikTok*: The Third Circuit Denies Section 230 Protection for Algorithmic Content.....733

II. The Potential Consequences of Denying Section 230 Protection to Content Moderation by Algorithm 734

 A. The Impact of Denying Section 230 Protection to Algorithms on Expression and Content Moderation 734

 B. Impact on Competition.....737

 C. Impact on AI and Innovation More Generally..... 739

III. Solutions to a Potential Problematic Interpretation: How to Protect Algorithms Under Section 230..... 740

 A. Legislative Intervention May Create More Problems than It Solves..... 740

 B. The Supreme Court Provides Clarity 741

Conclusion 742

* Jennifer Huddleston is a senior fellow in technology policy at the Cato Institute and an adjunct professor at George Mason University’s Antonin Scalia Law School.

Introduction

The 2024 Supreme Court decision in *Moody v. NetChoice*¹ and *NetChoice v. Paxton*² (collectively referred to as the *NetChoice* cases) was widely lauded as a significant victory for free speech rights.³ All nine Justices of the Court agreed in remanding the decisions below back to the Fifth and Eleventh Circuits, respectively, finding that the lower courts had failed to evaluate a First Amendment facial challenge according to the proper elements of the law.⁴

However, the Justices did not agree on the underlying First Amendment issues at hand—including the question of whether algorithms are expressions of a platform. The majority opinion written by Justice Kagan includes significant language indicating that the decision to engage in content moderation is likely protected by the First Amendment rights of platforms.⁵ However, the Justices disagreed over the additional instructions to the lower courts and the interpretation of algorithms and content moderation decisions under the First Amendment. Even though they agreed with the decision to remand, Justice Alito's concurrence—joined by Justices Thomas and Gorsuch—indicated their belief that the analysis in Justice Kagan's opinion was merely dicta.⁶ Beyond what the Court itself thinks of the *NetChoice* cases decision, this apparent split on how to interpret the decision has yielded a potentially concerning line of emerging case law in lower courts. This line of cases could change a crucial understanding of the fundamental protection of Section 230, which has allowed online platforms to host various kinds of user speech and might also impact the development of emerging technologies like AI.

The Court has not clearly spoken on the question of whether algorithms are protected by an online platform's Section 230 immunity or the full extent of their interaction with First Amendment rights, having avoided these questions both in the 2024 *NetChoice* cases decision as well as the 2023 *Twitter, Inc. v. Taamneh*⁷ and *Gonzalez v. Google LLC*⁸ decisions.⁹ However, these questions continue to appear in litigation brought against social media companies and other online services.¹⁰ The

¹ 144 S. Ct. 2383 (2024).

² *Id.*

³ See Prithvi Iyer, *Reactions to the Supreme Court's NetChoice Cases*, TECH POLICY.PRESS (July 2, 2024), <https://perma.cc/N3ZF-V7J7>.

⁴ See *Moody*, 144 S. Ct. at 2408–09.

⁵ See *id.* at 2401.

⁶ *Id.* at 2422 (Alito, J., concurring).

⁷ 143 S. Ct. 1206 (2023).

⁸ 143 S. Ct. 1191 (2023).

⁹ See *Gonzales*, 143 S. Ct. at 1192; *Taamneh*, 143 S. Ct. at 1214–18.

¹⁰ See, e.g., *Anderson v. TikTok, Inc.*, 116 F.4th 180, 182–84 (3rd Cir. 2024).

recent *NetChoice* cases decision seem to further complicate the matter in lower courts, turning an initial victory for speech into a concerning risk of increased liability.

Most notably, a recent Third Circuit decision found that since content moderation algorithms are considered a platform's own speech under *NetChoice*, Section 230 may not serve as a shield against lawsuits based on the algorithm promoting or demoting certain content.¹¹ Such issues are likely to become even more complicated as artificial intelligence ("AI") programs face their first legal challenges for content they produce, because popular large language models ("LLMs") are, at their heart, the result of a large number of algorithms and their interactions.

Many of these legal challenges brought against online platforms over content moderation and algorithms have heartbreaking facts involving young people or adults who harmed themselves or others after interacting with certain types of content.¹² As the saying goes, bad facts make bad law, and it is not surprising that policymakers, parents, and judges seek someone to blame. Yet the consequences of bad legal interpretations that remove Section 230 protection for content moderation decisions made by algorithms would extend further than concerns about "big tech" or social media. It could greatly impact any number of technologies, such as AI, shift competition to make it more difficult for small players to offer users opportunities for expression, and create a new form of the "moderator's dilemma." While platforms may be later vindicated by the First Amendment in such challenges, there might still be significant costs to both innovation and expression due to litigation costs, particularly on smaller platforms.

Part I of this Article explores the emerging case law around the interactions between algorithms in content moderation, the First Amendment, and Section 230. Part II then considers the potential consequences for innovation, expression, and competition should courts not find that Section 230's liability protection covers algorithms, as was seen in the Third Circuit *Anderson v. TikTok, Inc.* decision. This Article concludes, in Part III, with potential solutions to correct the emerging trend through either the courts or legislation.

I. Understanding Recent Jurisprudence Around the Intersection of Content Moderation, Algorithms, Section 230, and the First Amendment

While it may seem like a small, specific niche of law, the courts have already partially encountered the question of whether algorithms used in

¹¹ *Id.* at 184.

¹² See Rachel Hall, *Parents Sue TikTok Over Child Deaths Allegedly Caused by 'Blackout Challenge'*, THE GUARDIAN, (Feb. 7, 2025, at 09:02 ET), <https://perma.cc/8UMS-E3VB>.

content moderation decisions are covered by Section 230. The Supreme Court was directly faced with this question in the *Twitter* and *Google* cases mentioned in the introduction, and lower courts have, over the years, considered the questions of algorithms and speech both in Section 230 context and beyond.

A. *What Does Section 230 Actually Say?*

Section 230 has become a flashpoint in the battle over online speech, with critics on both the political left and right decrying the law as the reason for either too much content moderation (i.e., in the form of censorship), or too little moderation of content considered “misinformation” or “hate speech.”¹³ But Section 230 emerged from a bipartisan effort to prevent a potentially concerning trajectory, which arose from a split among courts, wherein online platforms were held liable for their users’ content if they engaged in content moderation.¹⁴ This trend risked discouraging platforms from removing “lawful but awful” content for fear of greater liability.¹⁵ This result would also have limited online speech more generally and likely resulted in far fewer opportunities for connection and communication via the internet.¹⁶

The so-called twenty-six words that created the internet read as follows:

No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

(2) Civil Liability[.] No provider or user of an interactive computer service shall be held liable on account of—

(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or

(B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1).¹⁷

The law’s text does not dictate that platforms must engage in moderation, nor does it specify the specific tools platforms can use to

¹³ See Ash Johnson & Daniel Castro, *Fact-Checking the Critiques of Section 230: What Are the Real Problems?*, INFO. TECH. & INNOVATION FOUND., (Feb. 22, 2021), <https://perma.cc/WRM4-GLXF>.

¹⁴ *Section 230: Legislative History*, ELEC. FRONTIER FOUND., <https://perma.cc/JAH5-AUJY>.

¹⁵ See Conor Clarke, *How the Wolf of Wall Street Created the Internet*, SLATE (Jan. 7, 2014, at 16:29 ET), <https://perma.cc/J84L-SEQ5>.

¹⁶ See Billy Easley, *Revising the Law That Lets Platforms Moderate Content Will Silence Marginalized Voices*, SLATE (Oct. 29, 2020, at 17:43 ET), <https://perma.cc/ZS2G-VL44>.

¹⁷ 47 U.S.C. § 230(c).

engage in that moderation.¹⁸ Both of the law's authors, as well as a wide array of scholars across the political spectrum, have repeatedly pointed out that the law does not require platform neutrality or specific actions or inactions, thus allowing platforms to serve either a specific type of audience and beliefs or a general audience.¹⁹

When it comes to the question of whether Section 230 was designed to cover the use of algorithms in making these content moderation decisions, the law's authors have filed amicus curiae briefs for relevant cases before the Supreme Court.²⁰ In these briefs, they indicate that the use of algorithms in content moderation, including recommendation algorithms, was known at the time of Section 230's drafting and that they believed the use of algorithms to engage in content moderation fell under the law's liability protections.²¹

However, the authors have not clearly stated whether they believe Section 230 immunity will fully cover the content created by generative AI services such as ChatGPT.²² As will be discussed further in Part III, in the absence of legislative changes, the full extent of interactions between Section 230 and AI will likely be impacted by the precedents set around algorithmic decision-making and Section 230 protection.

B. *How the Courts Have Explored Algorithms and Other Technologies as Speech and the Relationship to Section 230 Liability Protections*

Even prior to the development of Section 230, the courts were faced with complicated questions around whether new technologies may be considered a form of expression. For example, in the 1996 case *Bernstein v. United States Department of State*,²³ the Northern District of California found a "right to code" as a form of expression protected by the First Amendment. This decision concerned potential controls that were intended to restrict the export of certain cryptographic software as munitions.²⁴ The Supreme Court has largely avoided the precise questions of whether algorithms are considered a platform's speech and if they are covered by Section 230 liability protection. Because the Supreme Court

¹⁸ See Emily Stewart, *Ron Wyden Wrote the Law That Built the Internet. He Still Stands by It—and Everything It's Brought with It*, VOX (May 16, 2019, at 06:50 ET), <https://perma.cc/LGW5-QQU2>.

¹⁹ See *id.*

²⁰ Brief for Senator Ron Wyden & Former Representative Christopher Cox as Amici Curiae in Support of Respondent at 1, *Gonzalez v. Google LLC*, 598 U.S. 617 (2023) (No. 21-1333).

²¹ See *id.* at 2–3.

²² See *Section 230 Co-Author Says the Law Doesn't Protect AI Chatbots*, MARKETPLACE (May 19, 2023), <https://perma.cc/2SV2-A3GP>.

²³ 922 F. Supp. 1426 (N.D. Cal. 1996).

²⁴ See *id.* at 1429.

has declined to answer, a circuit split has begun to emerge that could significantly challenge the scope of Section 230's protections.²⁵

I. *Gonzalez v. Google* and a Potential Missed Opportunity

In 2023, *Gonzalez v. Google* presented the Supreme Court with the question: Does Section 230(c)(1) of the Communications Decency Act immunize interactive computer services like Google from civil lawsuits when they make targeted recommendations of harmful content provided by their users?²⁶ This case was combined with another case, *Twitter v. Taamneh*, which explored similar questions about the potential for victims of a terrorist attack to recover against a social media platform where the terrorist had shared content under existing U.S. law.²⁷

In deciding these cases, the Court focused on the inability to recover under the relevant U.S. terrorism-related statutes rather than reaching the question about the applicability of Section 230 to the use of algorithms and algorithmic recommendations.²⁸ This reasoning was generally well received by free speech advocates such as ACLU and deemed supportive of online speech;²⁹ however, the decision did not reach the relevant question about Section 230 and algorithms. This was not entirely surprising following oral arguments where the Court appeared to tread carefully, indicating a likely narrow ruling restricted to the facts of the case.³⁰

The decision on these other grounds meant the Court did not potentially disrupt the status quo regarding the use of algorithms as part of content moderation tools or Section 230 more generally; however, the circuit courts have continued to deal with such questions and have combined various elements of the Supreme Court's recent online speech jurisprudence to reach different conclusions.

²⁵ See Petition for a Writ of Certiorari at 19–30, *Force v. Facebook, Inc.*, 140 S. Ct. 2761 (2020) (No. 19-859).

²⁶ Petition for a Writ of Certiorari at i, *Gonzalez v. Google LLC*, 598 U.S. 617 (2023) (No. 21-1333).

²⁷ See *Gonzales v. Google LLC*, 2 F.4th 871, 880 (9th Cir. 2021) (discussing the claims of the *Taamneh* plaintiffs in a combined case).

²⁸ See *Twitter, Inc. v. Taamneh*, 143 S. Ct. 1206, 1217 (2023); *Gonzalez v. Google LLC*, 143 S. Ct. 1191, 1192 (2023).

²⁹ ACLU of N. Cal., *ACLU Commends Supreme Court Decisions Allowing Free Speech Online to Flourish*, ACLU (May 18, 2023, at 11:00 ET), <https://perma.cc/V24E-M4C7>.

³⁰ See Will Duffield, *Supreme Court Treads Carefully in Gonzalez*, CATO INST.: CATO AT LIBERTY (Mar. 22, 2023, at 08:18 ET), <https://perma.cc/53FZ-SZMN>.

2. *Force v. Facebook*: Second Circuit Finds Algorithmic Content Moderation Covered by Section 230

*Force v. Facebook, Inc.*³¹ had a similar set of facts to the cases heard by the Supreme Court.³² Like the *Taamneh* and *Gonzalez* cases, the *Force* case concerned the ability to recover under terrorism-related laws against a social media platform for hosting terrorist content and the interactions of its recommendation algorithm with such content.³³

The *Force* case was decided by the Second Circuit before both *Gonzalez* and the *NetChoice* cases.³⁴ The Second Circuit found that a recommendation algorithm, such as that of Meta's Facebook, was a content moderation tool involved in the distribution of content and therefore subject to the liability protections of Section 230.³⁵ In a dissent, Judge Robert Katzmann questioned the idea that recommendation algorithms were merely neutral tools serving as a necessary part of the content moderation process and expressed concerns about the potential for such algorithms to lead individuals "further down dark paths."³⁶

This decision would indicate that even if algorithms are a form of First Amendment expression, their use in content moderation processes is still protected by Section 230.

3. *Anderson v. TikTok*: The Third Circuit Denies Section 230 Protection for Algorithmic Content

In 2024, the Third Circuit found that litigation against the social media app TikTok was not barred by Section 230.³⁷ This case concerns a young girl who passed away after participating in the "blackout challenge" that her family alleged she learned about and participated in due to the recommendations of TikTok's FYP ("For You Page") algorithm.³⁸ The plaintiffs argued that TikTok was liable for failing to warn users about the dangers of such actions due to their knowledge of such content and continued recommendation of this content to young users.³⁹

The case was initially dismissed by a Pennsylvania district court under Section 230; however, the Third Circuit on appeal found that the case was

³¹ 934 F.3d 53 (2d Cir. 2019).

³² *See id.* at 57–61.

³³ *See id.*

³⁴ *See id.* at 53.

³⁵ *See id.* at 65–66.

³⁶ *See id.* at 85–88 (Katzmann, C.J., concurring in part and dissenting in part).

³⁷ *Anderson v. TikTok, Inc.*, 116 F.4th 180, 184 (3d Cir. 2024).

³⁸ *Id.* at 181.

³⁹ *Id.* at 182.

not barred by Section 230 and partially reversed, vacated, and remanded the decision.⁴⁰ The circuit court rooted this decision in its interpretation of the Supreme Court's decision in *NetChoice*: If choosing how to compile and display the content its users create is a platform's own expression, then such actions, although protected by the First Amendment, are the platform's own expression and therefore not protected from the liability shield of Section 230.⁴¹ The Third Circuit's reasoning also focused on the fact that content was not found through a user input like a search, but rather through the platform's recommendation algorithm and FYP.⁴² Rehearing of the case en banc was denied.⁴³

The decision in this case sets up a split between the Second Circuit's decision in *Force* and the Third Circuit's post-*NetChoice* decision in *Anderson*, which may ultimately force the Supreme Court to consider the issue, as discussed further in Part III of this paper.

II. The Potential Consequences of Denying Section 230 Protection to Content Moderation by Algorithm

While there are strong emotions around TikTok related to both general debates over social media as well as specific debates to the app's potential ties to a Chinese parent company, the ruling in *Anderson* is more expansive than just one possible app. If other courts adopt a similar interpretation, it will result in significant consequences beyond just social media. As this Section will discuss, these consequences will not only impact speech and content moderation but might also impact the landscape of competition in social media and AI. In effect, denying Section 230 protection when algorithms are used could create a new form of "moderator's dilemma."

A. *The Impact of Denying Section 230 Protection to Algorithms on Expression and Content Moderation*

By denying Section 230 protection to content moderation decisions by algorithm or the use of recommendation algorithms, the Third Circuit may have unknowingly created a new moderator's dilemma. Effectively, this interpretation significantly narrows the existing interpretation of

⁴⁰ *Id.* at 184–85.

⁴¹ *Id.* at 184.

⁴² *Id.*

⁴³ See Aleeza Furman, *TikTok Opts Not to Take Section 230 Immunity Fight to U.S. Supreme Court*, LAW.COM: LEGAL INTELLIGENCER (Feb. 6, 2025, at 17:16 ET), <https://perma.cc/A5M2-N4TC>.

Section 230 and renders it inversely related to the First Amendment.⁴⁴ Not only will this interpretation significantly narrow the tools available to platforms to engage in moderation, but it might also result in a less productive and enjoyable speech environment for internet users.

The Moderator's Dilemma refers to the phenomenon that platforms displaying content are faced with the choice of engaging in heavy-handed moderation, greatly restricting any potentially offensive or questionable content, and silencing important conversations and voices in the process, or engaging in no moderation at all, resulting in a potential influx of "lawful but awful" speech such as spam, animal cruelty, and pornography.⁴⁵ Denying Section 230 protection to the use of algorithms—or, more specifically, the use of recommendation algorithms—does not fully result in the same dilemma, but could create a similar paradox for platforms, based on the scale of user-generated content, which often requires the use of tools beyond just simple human review, given the quantity of such content.⁴⁶

Removing Section 230 protection from algorithms discourages the use of these tools for content moderation and forces platforms to utilize more human moderators or cease to engage in moderation for certain types of content. In most cases, such a transition away from algorithms and toward human moderators is unlikely to be sustainable if services wish to continue to provide the same opportunities for user content.

Not only would relying solely on human moderators take longer for the same amount of content to be reviewed, but various articles and interviews with human moderators have documented the psychological and emotional trauma that repeated exposure to such extreme and problematic content can have on the individuals doing the moderation.⁴⁷ While companies have tried to provide additional resources and support in such cases, simply replacing algorithms with human moderators is unlikely to be sustainable for any platform of size. Recent statistics from

⁴⁴ Jack Fitzhenry, *Let the Algorithm Speak?: Third Circuit Indicates in Anderson v. TikTok That the First Amendment and Section 230 Are Inversely Related*, FEDERALIST SOC'Y (Sep. 30, 2024), <https://perma.cc/YKJ8-SHW9>.

⁴⁵ Matt Schruers, *Debate over Online Content Embodies "Moderator's Dilemma"*, DISRUPTIVE COMPETITION PROJECT (Sep. 4, 2019), <https://perma.cc/3EVR-F9RW>.

⁴⁶ See Andrew Hutchinson, *What Happens on the Internet Every Minute (2024 Version)*, SOCIALMEDIATODAY (Dec. 18, 2024), <https://perma.cc/TBQ2-HYAA>.

⁴⁷ Ruth Spence, Antonia Bifulco, Paula Bradbury, Elena Martellozzo & Jeffrey DeMarco, *Content Moderator Mental Health, Secondary Trauma, and Well-Being: A Cross-Sectional Study*, 27 CYBERPSYCHOLOGY, BEHAV. AND SOC. NETWORKING 149, 153 (2024); see also Billy Stockwell, *Facebook Inflicted 'Lifelong Trauma' on Kenyan Content Moderators, Campaigners Say, as More than 140 Are Diagnosed with PTSD*, CNN (Dec. 22, 2024, at 06:23 ET), <https://perma.cc/UYF8-DRE6> (discussing trauma experienced by content moderators); Deepa Parent & Katie McQue, *'I Log into a Torture Chamber Each Day': The Strain of Moderating Social Media*, THE GUARDIAN (Sep. 11, 2023, at 01:00 ET), <https://perma.cc/6NND-QPG8> (discussing the impact of content moderation on a content moderator's mental health).

popular social media companies indicate that algorithms either conduct or are involved in over 90% of content moderation online.⁴⁸ Even when humans are frequently involved in the end decisions around content moderation or in a more decentralized system—as the amicus brief of sub-Reddit moderators in *Gonzalez v. Google* noted—algorithms still play a pivotal role in identifying and prioritizing the content most in need of review.⁴⁹ Furthermore, as this same amicus brief notes, the removal of liability for the use of algorithms also puts human moderators at risk for personal liability as well.⁵⁰

In reality, almost all content moderation is a hybrid of human and algorithmic interactions, though the exact balance may vary from company to company or community to community.⁵¹ Removing Section 230 protection from content moderation decisions that involve algorithms would increase the cost of content moderation and potentially create liability for content moderation strategies more generally. As a result, platforms might be forced to adopt a strategy they feel is subpar for the community's needs because of potential liability risk.

Algorithms also play an important role in allowing companies to respond to consumer demands and create more tailored experiences. For example, recommendation algorithms play a part in parental controls by preventing young people from being fed otherwise age-restricted content.⁵² Algorithms also play a role in helping identify potential online predators and identifying problematic interactions with young users.⁵³ Thus, in many cases, the removal of algorithms from content moderation is likely to make for a less enjoyable user experience. Even if Section 230 is found not to apply to recommendation algorithms, it could still greatly impact the ability of platforms to respond to growing demands to provide options and protections for young users in the face of bad actors and problematic content.

Section 230 protection has also been critically important for other types of platforms that carry user-generated content, such as review sites. These sites use algorithmic tools to engage in content moderation, including recommendation algorithms to assist and improve the user

⁴⁸ David Inserra, *A Guide to Content Moderation for Policymakers*, CATO INST. (May 21, 2024), <https://perma.cc/534C-Y7BR> (Policy analysis No. 974).

⁴⁹ Brief for Reddit, Inc. & Reddit Moderators as Amici Curiae in Support of Respondent at 21, *Gonzalez v. Google LLC*, 143 S. Ct. 1191 (2023) (No. 21-1333).

⁵⁰ *See id.*

⁵¹ *See* Valentine Crosset & Benoît Dupont, *Cognitive Assemblages: The Entangled Nature of Algorithmic Content Moderation*, SAGE J.: BIG DATA & SOC'Y (Dec. 12, 2022), <https://perma.cc/U95W-XNEA>.

⁵² *See The Rise of AI in Parental Control Tech*, TECHNOLOGY.ORG (June 3, 2025), <https://perma.cc/553V-V6PL>.

⁵³ Brian L. Huchel, *Algorithm Tool Works to Silence Online Chatroom Sex Predators*, PURDUE UNIV. (Apr. 17, 2018), <https://perma.cc/Q8PY-GSRS>.

experience.⁵⁴ For example, popular review sites like Yelp and TripAdvisor rely on user-generated content and use recommendations and other algorithms to guide users to relevant reviews based on high-quality content or other factors.⁵⁵ Removing protection for such practices could result in a business being overrun by fraudulent reviews or customers spending more time to find the desired information.

As in the case of the moderator's dilemma, some platforms might continue using algorithms, including recommendation algorithms, but with additional restrictions for content negatively impacting speech in the process. The alternative would be continuing use of algorithms but with a significant litigation risk, substantially limiting the types of content that platforms may recommend and magnifying negative impacts to marginalized voices. For example, it might be less likely that algorithms will recommend content associated with a certain religion, for fear that such content might be considered radicalization or religious extremism, or with healthy eating or exercise, for fears of promoting eating disorders, and so on.

Recommendation algorithms can positively help tailor a user's experience on the platform while also exposing them to content they might not have otherwise discovered, thus creating further opportunities for connection or sparking a new interest.⁵⁶ Their absence can mean that only the loudest and largest voices are heard, making it more difficult to discover content or voices that might be of interest but are not as widely popular to a general audience.⁵⁷ The result is that users would have less opportunities to connect or express themselves and platforms would have fewer options to tailor content moderation strategies to specific communities. In this way, removing Section 230 protection from algorithms or recommendation algorithms undermines the very intentions that led to Section 230's creation.

B. *Impact on Competition*

While critics of Section 230 will disparage it as a handout or free ride for "big tech,"⁵⁸ removing Section 230 liability protection could make it more difficult for smaller user-generated, content-carrying platforms to compete. Even if an erosion of this protection is limited only to algorithms

⁵⁴ See, e.g., *Why Does Yelp Recommend Reviews?*, YELP, <https://perma.cc/WBS5-GXCS>.

⁵⁵ *Id.*

⁵⁶ Elizabeth Nolan Brown, *In Defense of Algorithms*, REASON (Jan. 2023), <https://perma.cc/D9JR-C8KT>.

⁵⁷ See *id.*

⁵⁸ See, e.g., Matt Stoller, *Judges Rule Big Tech's Free Ride on Section 230 Is Over*, BIG BY MATT STOLLER (Aug. 29, 2024), <https://perma.cc/M25D-UYR4>.

or recommendation algorithms, these changes are likely to be more acutely felt by smaller platforms.

Section 230 has pro-competitive effects by making it possible for platforms to allow users to share content without fear of company-ending liability if a user creates content they did not intend or if they allow or remove certain content in an effort to maintain standards or reach a certain audience.⁵⁹ Even if a company's actions are ultimately deemed lawful under First Amendment expression or association rights, the cost of litigation can be significant and potentially business-ending for a small business as even the discovery phase can cost a company over \$300,000.⁶⁰ Section 230 has also allowed platforms to create specific content moderation policies and utilize specific recommendation algorithms that may provide opportunities for serving a niche or smaller audience based on a shared interest or set of beliefs.⁶¹

Recommendations and other algorithms can play an important role in consumer enjoyment allowing newer or smaller platforms to establish their place in the market.⁶² Finding a way to recommend content that resonates with an audience of consumers, as well as decisions around hosting that content, are both parts of content moderation. Small platforms can also offer alternatives that focus on more limited or private factors, but these user controls still result in some sort of algorithmic interaction regarding what content is displayed.

Just as litigation has costs, so too does content moderation, whether done by human or algorithm. Algorithms can help reduce the cost of broadly agreed-upon content moderation, such as by removing spam or providing an initial flag for content that might violate the platform's rules, as algorithms can review a greater amount of content than a human in the same span of time and can also scan for trends more easily.⁶³ Algorithms, of course, are often imperfect with both false positives and false negatives; however, they still can reduce the burden on human moderators when dealing with the large amount of content most sites encounter.⁶⁴ Additionally, human moderators can and do make mistakes and bring with them their own biases. When user-generation content platforms face

⁵⁹ Jennifer Huddleston, *Competition and Content Moderation: How Section 230 Enables Increased Tech Marketplace Entry*, CATO INST. (Jan. 31, 2022), <https://perma.cc/4H9Y-DJ4Y> (Policy analysis No. 922).

⁶⁰ See Evan Engstrom, *Primer: Value of Section 230*, ENGINE (Jan. 31, 2019), <https://perma.cc/47PY-9D3N>.

⁶¹ See Huddleston, *supra* note 59.

⁶² See Alex Hern, *How TikTok's Algorithm Made It a Success: 'It Pushes the Boundaries'*, THE GUARDIAN (Oct. 24, 2022, at 01:00 ET), <https://perma.cc/Y6PR-M9BN>.

⁶³ See Juan Londoño, *Assessing the Impact of the Widespread Adoption of Algorithm-Backed Content Moderation in Social Media*, AM. ACTION F. (Jan. 25, 2022), <https://perma.cc/NFW3-8T29>.

⁶⁴ See Marc Faddoul, *COVID-19 Is Triggering a Massive Experiment in Algorithmic Content Moderation*, BROOKINGS INST. (Apr. 28, 2020), <https://perma.cc/8H44-QC4M>.

limited resources, either due to the size of the company or other reasons, algorithms are often necessary tools.

While platforms of all sizes might face a new form of moderator's dilemma if recommendation algorithms—or algorithms more generally—are not protected by Section 230, newer and smaller platforms are likely to have fewer resources to incur litigation costs or to continue to host user-generated content in a revised way. This change in costs and liability might impact the market dynamics and encourage concentration. Possible reactions would be to seek acquisition by a larger platform or to exit the market by no longer carrying user-generated content. Either of these solutions only increases the market share of the current largest players, who are most able to absorb the cost of a changed dynamic.⁶⁵

C. *Impact on AI and Innovation More Generally*

While AI is often portrayed as novel and disruptive, it is largely an expansion and combination of algorithms.⁶⁶ If courts continue on the path that denies Section 230 protection to content moderation decisions that involve algorithms, this may negatively impact or deter not only the use of AI tools in content moderation but also the development of AI, which is reliant on algorithms to help users create content.

The broader question of when a generative AI product is protected by Section 230 is a far larger topic and debate that deserves a full exploration beyond the scope of this Article. However, the removal of the liability shield of Section 230 for algorithms or recommendation features means that AI products are unlikely to receive this protection. This is because LLMs are effectively gathering large amounts of information and producing a form of recommendation in their response to such content requests.⁶⁷

A lack of clarity around the liability protection for AI could deter generative AI or the implementation of AI into existing user-generated content services. This includes not only AI products used for content moderation⁶⁸ but also AI products that might be integrated into the platform's services like Instagram's Imagine feature.⁶⁹

⁶⁵ See Huddleston, *supra* note 59.

⁶⁶ See Aaron Patzer, *Not All Algorithms Are AI (Part 1): How the Revolution Started*, FORBES (Aug. 30, 2023, at 06:30 ET), <https://perma.cc/3L5M-ZWM3>.

⁶⁷ See *LLMs vs. Traditional ML Algorithms – A Pragmatic Comparison*, MLOPSAUDITS.COM, <https://perma.cc/2GQW-6J83> (explaining LLMs versus traditional algorithms).

⁶⁸ See Rem Darbinyan, *The Growing Role of AI in Content Moderation*, FORBES (June 14, 2022, at 06:45 ET), <https://perma.cc/LK3D-MVJV>.

⁶⁹ *The Rise of AI Integration in Social Media Platforms*, GREEN CLOSET CREATIVE, <https://perma.cc/39Y8-CRSN>.

While AI, both as a term and a technology, has been explored since the 1950s,⁷⁰ the sudden popularity and accessibility of these products have created new opportunities for innovation and disruption. Much like the internet in 1996, removing Section 230 protection for algorithms and AI might prevent positive applications of these products in the speech space. While the full extent of Section 230's application to generative AI outputs is a question likely to see continued debate, the affirmative removal of the protection from AI or algorithms will likely deter innovation and investment in its potential applications for content moderation or user-generated content at a minimum.

III. Solutions to a Potential Problematic Interpretation: How to Protect Algorithms Under Section 230

Because removal of Section 230 protection from algorithms impacts not only content moderation on a range of internet platforms but also competition and innovation, a clarification that Section 230 does apply to algorithms used in content moderation is necessary. As with Section 230's origins in a circuit split over content moderation,⁷¹ the potential solutions are likely to come from either intervening Congressional legislation or the Supreme Court's resolution of the emerging circuit split.

A. *Legislative Intervention May Create More Problems than It Solves*

Given congressional consideration and political temperament around Section 230, Congress seems more likely to significantly narrow its scope than to protect it or expand its interpretation. Additionally, when it comes to tech policy in particular, Congress seems hesitant to act, providing both pacing problems and pacing benefits in the absence.⁷²

The most relevant congressional action from the 118th Congress, for example, failed to clarify that Section 230 provides protection for algorithms and rather sought to formally remove Section 230 protection from AI products.⁷³ Similarly, in the 119th Congress, the discussion of Section 230 has included a potential bipartisan effort to sunset the law in

⁷⁰ *The Birth of Artificial Intelligence (AI) Research*, LAWRENCE LIVERMORE NAT'L LAB'Y, <https://perma.cc/T7BT-XJWG>.

⁷¹ See Clarke, *supra* note 15.

⁷² See Adam Thierer, *The Pacing Problem, the Collingridge Dilemma & Technological Determinism*, TECH. LIBERATION FRONT (Aug. 16, 2018), <https://perma.cc/LBV9-KPJ5>.

⁷³ See Adam Thierer & Shoshana Weissmann, *Even if You Hate Both AI and Section 230, You Should Be Concerned About the Hawley/Blumenthal Bill to Remove 230 Protections from AI*, R ST. (Dec. 6, 2023), <https://perma.cc/R7MW-ZUTY>.

its entirety.⁷⁴ Additionally, the last legislative change to Section 230 weakened the law by creating an additional carveout for sex trafficking content, but with a far greater impact to a range of platforms including the loss of everything from groups for legal sex workers to Craigslist's personal ads.⁷⁵ Therefore, reopening Section 230 to legislative changes—even if for positive intentions—is likely to result in more problems from those on both the left and the right who have been awaiting such opportunities to potentially revoke or significantly limit the law.

B. *The Supreme Court Provides Clarity*

In the coming years, the Supreme Court will likely again have the opportunity to address the issue and resolve the emerging circuit split. While the Court avoided the question in *Gonzalez* as discussed in Part II, several cases provide a potential opportunity for the Court to clarify after *NetChoice*. However, lower courts ought to consider the interactions between Section 230 and the First Amendment, as well as the question of algorithms or AI and Section 230 more directly. The particular positioning of the *NetChoice* cases may allow the Court to further clarify the applicability of its decision and the ways the First Amendment interacts with Section 230. The case was remanded to the lower courts, and Justice Kagan's majority opinion provided instructions that content moderation decisions, including those by algorithms, merit First Amendment protection.⁷⁶ However, in its remand to the district court, the Fifth Circuit appears to seek additional information, particularly around the questions of algorithms and the state's algorithmic transparency requirements.⁷⁷ It is possible that these questions could end up back at the Supreme Court, even from the same set of cases, following further review.

In its recent jurisprudence, the Court has seemed hesitant to directly address questions related to online speech. In its decision in *TikTok Inc. v. Garland*,⁷⁸ for example, the Court largely relied on data security questions rather than answering the more difficult questions around speech.⁷⁹ In *Gonzalez*, it relied on other grounds to reach the decision that the plaintiffs were not able to recover.⁸⁰ The Court denied cert in other recent

⁷⁴ Lauren Feiner, *Lawmakers Are Trying to Repeal Section 230 Again*, THE VERGE (Mar. 21, 2025, at 17:13 ET), <https://perma.cc/2VBY-YBTJ>.

⁷⁵ See Laura Moraff, *How Online Censorship Harms Sex Workers and LGBTQ Communities*, ACLU (June 27, 2022), <https://perma.cc/96DX-AYE5>.

⁷⁶ *Moody v. NetChoice, LLC*, 144 S. Ct. 2383, 2409 (2024).

⁷⁷ *NetChoice, L.L.C. v. Paxton*, 121 F.4th 494, 499–500 (5th Cir. 2024).

⁷⁸ 145 S. Ct. 57 (2025) (per curiam).

⁷⁹ *Id.* at 66.

⁸⁰ *Gonzalez v. Google LLC*, 143 S. Ct. 1191, 1192 (2023).

cases where such questions might arise, such as *Malwarebytes v. Enigma Software Group*.⁸¹

TikTok has chosen not to appeal the *Anderson* decision to the Supreme Court;⁸² however, other evolving issues and cases in online speech might ultimately provide the Court an opportunity to grant cert and clarify critical intersections between Section 230, algorithms, and the First Amendment. Pending litigation involving generative AI in district courts, however, are likely to present further opportunities for consideration if they reach the Supreme Court.⁸³ Cases challenging the restrictions or requirements related to algorithms that serve content to minors, such as California's Age Appropriate Design Code, may also present opportunities for the Court to clarify how it views the relationship between algorithms, speech, and government restrictions.⁸⁴

Offline examples illustrate that even if algorithms are the platform's own speech and not protected from liability by Section 230, it is still likely that recommendation algorithms are insufficient to sustain claims of incitement to harmful actions and result in plaintiff recovery. The Supreme Court spoke clearly on the issue of the First Amendment in rejecting a California law that would have restricted the sale of "violent video games" based merely on concerns that such games could incite violence.⁸⁵ Additionally, lower courts dismissed claims that Ozzy Osbourne's song "Suicide Solution" was sufficient incitement in a teen's suicide.⁸⁶ Yet the fact that these claims are likely to fail does not mean that Section 230's protection for algorithms is unnecessary.⁸⁷

While *NetChoice* provided a roadmap for a potential strong statement of Free Speech, *Anderson* indicates that without further guidance from the Supreme Court, it might have opened a new set of concerns around when and how Section 230 and the First Amendment might intersect with regard to tools for and decisions related to content moderation.

Conclusion

The facts in *Anderson v. TikTok, Inc.* are without a doubt tragic; however, the consequences of such an interpretation extend far beyond

⁸¹ 141 S. Ct. 13 (2020); *see id.* at 13–14 (Thomas, J., concurring).

⁸² Furman, *supra* note 43.

⁸³ *See* Kevin Roose, *Can A.I. Be Blamed for a Teen's Suicide?*, N.Y. TIMES (Oct. 24, 2024), <https://perma.cc/2D4X-Z2KP>.

⁸⁴ *See* PETER J. BENSON, CONG. RSCH. SERV., LSB11071, *NETCHOICE V. BONTA AND FIRST AMENDMENT LIMITS ON PROTECTING CHILDREN ONLINE 3* (2023).

⁸⁵ *Brown v. Ent. Merchs. Ass'n*, 564 U.S. 786, 804–05 (2011).

⁸⁶ Greg Henderson, *Supreme Court Lets Stand Ruling for Ozzy Osbourne*, UNITED PRESS INT'L (Oct. 13, 1992), <https://perma.cc/8PPD-4BZ3>.

⁸⁷ *See* Huddleston, *supra* note 59.

one tragedy and one app and could be significant for speech, innovation, and competition. Section 230 notably does not preclude bringing claims against the users who post harmful content or use online platforms to engage in nefarious activities. Creating a two-tiered system of Section 230 where platforms that use algorithms for recommending content or otherwise engaging in content moderation must first litigate the case will undermine many of the intentions behind Section 230 by creating a new moderator's dilemma. Such problems are likely to become even more pronounced with the increasing popularity and integration of generative AI and AI tools. Despite avoiding it in the past, the Supreme Court is likely to confront these questions directly once again in the coming years and should take the opportunity to clarify the interaction between the First Amendment and Section 230, thereby ensuring that speech may continue to flourish.