THE FUTURE OF FREE SPEECH

POLICY PAPER SERIES

I BUILT THIS ALGO BRICK BY BRICK

WHY YOU SHOULD CARE ABOUT ALGORITHMS, FREE SPEECH, AND SECTION 230

SEPTEMBER 10, 2025

Ashkhen Kazaryan
Senior Legal Fellow

INTRODUCTION

"I built this algo brick by brick."

You might have seen that sentence in the comments section while scrolling TikTok's "For You Page" or Instagram Reels. It usually appears as a humorous reference to the fact that the commenter is seeing something in their feeds that they didn't search for but were "fed" by the algorithm based on their previous engagements with content that they find interesting or relevant. While it may appear to be just a meta-joke among users, this comment reflects our collective and intuitive understanding of how our modern Internet functions: users and algorithms work together to shape our overall experience. Through every click, scroll, and pause, users train the system that determines what they see next. The result is a feed that feels curated not by chance, but by design.

At the core of that design is algorithmic curation, which refers to the use of computational systems to filter, rank, recommend, deprioritize, and remove content from the millions of posts uploaded each day.

A growing chorus of voices is arguing that the legal foundations that built the Internet as we know it need to change, including the algorithms that we interact with on a daily basis.[1] They contend that platforms are either moderating too much or too little content, and that the algorithms used to perform this function should be subject to government scrutiny. These arguments often stem from a misunderstanding about how our current legal framework has shaped the modern Internet and what would happen if we were to abandon it.

Without these algorithms — along with the legal discretion for platforms to make content moderation decisions — platforms would be unable to manage the scale and complexity of user activity online. And if they can't do that, users will be left with a fragmented version of the online ecosystem. With algorithms, platforms make a series of editorial decisions, shaped by their policies, incentives, and legal obligations, about what content to elevate, downrank, or exclude altogether.

This paper argues that algorithmic curation is not just a technical necessity. It is a form of editorial discretion protected by the First Amendment and essential to the functioning of the digital ecosystem. Section 230, in turn, provides the statutory safety net that makes such moderation feasible without constant legal peril.

1 Jaron Lanier, Allison Stanger & Audrey Tang, *Sunset and Renew: Section 230 Should Protect Human Speech, Not Algorithmic Virality,* Ash Center for Democratic Governance & Innovation, Harvard Kennedy School (Oct. 30, 2024), *available at* https://ash.harvard.edu/articles/sunset-and-renew-section-230-should-protect-human-speech-not-algorithmic-virality/.

First, the First Amendment safeguards a platform's right to exercise editorial judgment, including decisions made through algorithmic systems about what to highlight, downrank, or remove. Second, Section 230 of the Communications Decency Act shields platforms from liability for third-party content, allowing them to moderate and curate that content.

Efforts to restrict these protections through must-carry legislation, government action[2], or litigation could risk undermining the Internet's foundational legal framework. If platforms lose the ability to moderate and curate without liability, they face an untenable choice: allow all content, including material that platforms deem harmful to users, or engage in excessive censorship to avoid liability.[3] Either approach diminishes free expression and leads to governmental control over editorial decisions. For us, the users, and our society at large, it means fewer spaces for dialogue, less access to information relevant to us, and greater exposure to content we don't like. Algorithmic curation helps users navigate billions of pieces of content to find community, interact, and engage with the digital world in a meaningful way.

1. UNDERSTANDING ALGORITHMIC CURATION

Before deciding whether or how to regulate the tools at the center of this debate, we must first understand what they are, how they function, why they matter, and what the consequences of regulating them will be.

From social media to e-commerce and streaming services, algorithmic curation determines what content users see and in what order. These systems are essential to how most of the Internet organizes the overwhelming volume of information posted online every second.[4] Without them, navigating today's digital world would be like trying to find a single book in a library with no catalog, no sections, and no staff... just millions of pages scattered at random.

---

2 Press Release, Off. of the Mo. Att'y Gen., Attorney General Bailey Files Groundbreaking Rule to End Big Tech's Censorship Monopoly and Protect Online Free Speech (May 6, 2025), https://ago.mo.gov/attorney-general-bailey-files-groundbreaking-rule-to-end-big-techs-censorship-monopoly-and-protect-online-free-speech/.

3 *See* Daphne Keller, *Amplification and Its Discontents*, 21-05 Knight First Amend. Inst. (June 8, 2021), https://knightcolumbia.org/content/amplification-and-its-discontents.

4 *See generally,* Tomas Apodaca & Natasha Uzcátegui-Liggett, *How Automated Content Moderation Works (Even When It Doesn't),* The Markup (Mar. 1, 2024), https://themarkup.org/automated-censorship/2024/03/01/how-automated-content-moderation-works-even-when-it-doesnt-work.

At its core, algorithmic curation is about arrangement, prioritization, deprioritization, and often personalization. When you type a phrase into a search engine, the platform decides which results to show you first based on relevance, credibility, popularity, and your past search behavior.[5] When you scroll through your feed on TikTok, Instagram, YouTube, or X, the content has been sorted and ranked through complex models that learn what you like, ignore, and what people like you engage with most. Algorithms don't just sort content; they select it based on values, patterns, and policy choices.[6]

Crucially, algorithms are not limited to entertainment or convenience. Platforms also use them to enforce community guidelines and safety policies at scale.[7] Content promoting suicide, eating disorders, child exploitation, or targeted harassment can appear in millions of forms, across countless languages and formats.[8] Human moderators alone cannot catch it all. And every platform has its own definitions and standards when it comes to what type of content they permit, with some adopting notably more expansive or restrictive standards than others.[9] Vulnerable users like adolescents can benefit from algorithms that assist by detecting patterns in real time and flagging or downranking content deemed harmful to these users.[10]

Terms like "algorithmic amplification" or "recommendation systems" are often used to describe aspects of algorithmic curation, but they refer to the same idea: the process of using algorithms to highlight certain content while downplaying other content. It is how platforms shape their environments and how they express what is important, trustworthy, or allowed under their rules. In practice, it is what makes digital platforms functional[11], and, in many cases, safer.[12]

---

5 Google, Ranking Results, How Search Works (last visited Aug. 28, 2025), *available at* https://www.google.com/intl/en_us/search/howsearchworks/how-search-works/ranking-results.

6 *See* Adam Mosseri, *Shedding More Light on How Instagram Works*, Instagram Blog (June 8, 2021), *available at* https://about.instagram.com/blog/announcements/shedding-more-light-on-how-instagram-works; Ben Smith, *How TikTok Reads Your Mind*, N.Y. TIMES (Dec. 5, 2021), https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html.

7 *See* Spandana Singh, *Everything in Moderation An Analysis of How Internet Platforms Are Using Artificial Intelligence to Moderate User-Generated Content*, New America, July 22, 2019, https://www.newamerica.org/oti/reports/everything-moderation-analysis-how-internet-platforms-are-using-artificial-intelligence-moderate-user-generated-content/.

8 *Id.*

9 Jacob Mchangama, Abby Fanlo and Natalie Alkiviadou, Scope Creep: An Assessment of 8 Social Media Platforms' Hate Speech Policies," The Future of Free Speech, (June 1, 2023), *available at* https://futurefreespeech.org/scope-creep/.

10 *See* Hilbert, Martin and Cingel, Drew P. and Zhang, Jingwen and Vigil, Samantha L. and Shawcroft, Jane and Xue, Haoning and Thakur, Arti and Shafiq, Zubair, #BigTech @Minors: social media algorithms have actionable knowledge about child users and at-risk teens. (December 19, 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4674573; Luke Balcombe & Diego De Leo, The Impact of YouTube on Loneliness and Mental Health, 10 Informatics No. 2, art. 39 (Apr. 20, 2023), https://www.mdpi.com/2227-9709/10/2/39.

11 *Id.*

12 *Id.*

2. THE FIRST AMENDMENT DOESN'T JUST PROTECT SPEECH. IT PROTECTS CHOICES ABOUT SPEECH

When a platform chooses what to show, recommend, downrank, or remove, whether it's highlighting reliable information, filtering out hate speech, or promoting a video you might enjoy, it is making a decision about speech.

Courts have long recognized that this kind of curating, editing, and prioritizing is part of the freedom to speak.[13] These editorial rights are protected by the First Amendment. Editorial rights, at their core, represent the autonomy and authority to make decisions about speech. They include editorial discretion — the selection of what will be published, what will be omitted, the framing and contextualization of narratives, what is highlighted, and what is downplayed. These decisions, made with autonomy and authority, play a pivotal role in shaping the informational ecosystem of our society.

One of the clearest examples comes from a 1974 Supreme Court case involving a newspaper.[14] Florida passed a law requiring newspapers to give political candidates space to reply to editorials that criticized them.[15] The *Tornillo* Court struck down the law, ruling that the newspaper's editorial judgment was protected speech under the First Amendment.[16] The Court explained that forcing a newspaper to carry someone else's message was just as much a violation of free speech as stopping it from speaking in the first place.[17]

That principle didn't stay limited to newspapers. In later cases, the Court extended it to other private organizations that curate or present messages:

- In ***Pacific Gas & Electric Co. v. Public Utilities Commission of California*** (1986), the Supreme Court struck down a state mandate requiring a private utility to include third-party messages in its billing envelopes.[18] The Court held that compelling the utility to distribute unwanted speech

13 *See* Miami Herald Publ'g Co. v. Tornillo, 418 U.S. 241 (1974), Hurley v. Irish-Am. Gay, Lesbian and Bisexual Group of Boston, 515 U.S. 557, 574 (1995), Moody v. NetChoice, LLC, 144 S. Ct. 2383 (2024).

14 Miami Herald Publ'g Co. v. Tornillo, 418 U.S. 241 (1974).

15 *Id.*

16 *Id.*

17 *Id.*

18 *Pac. Gas & Elec. Co. v. Pub. Utils. Comm'n of Cal.*, 475 U.S. 1, 16–17 (1986) (plurality opinion) ("[F]orced associations that burden protected speech are impermissible. The choice to speak includes within it the choice of what not to say.").

violated its First Amendment right to editorial discretion.[19] It explained that the government cannot force a private speaker to use its communication channels to carry messages it does not wish to promote.[20] This reinforced the principle that the right to exclude speech is just as protected as the right to express it.

- In ***Hurley v. Irish-American Gay, Lesbian and Bisexual Group of Boston*** (1995), the Court upheld a parade organizer's right to exclude a group whose message would have changed the meaning of the parade.[21] Justice Souter, writing for the unanimous Court, stressed that speech is also choosing not to speak: "one important manifestation of the principle of free speech is that one who chooses to speak may also decide 'what not to say.'"[22]

- In ***Manhattan Community Access Corp. v. Halleck*** (2019), the Court reaffirmed that private entities have the right to control speech on platforms they create.[23] Writing for the majority, Justice Kavanaugh explained that "when a private entity provides a forum for speech," it "may... exercise editorial discretion over the speech and speakers in the forum."[24] This principle reinforces that editorial judgment does not become public action simply because it occurs on a forum open to others.

- In ***303 Creative LLC v. Elenis*** (2023), the Court held that the First Amendment protects intermediaries who vet and shape the speech of others.[25] The case involved a website designer who declined to create custom wedding sites for same-sex couples, arguing it would compel her to convey a message contrary to her beliefs.[26] The Court agreed, emphasizing that even when speech originates with a third party, the act of selecting, shaping, or refusing to disseminate it is itself expressive and constitutionally protected.[27] Relying on *Hurley*, the Court concluded that combining or curating another's expressive content is itself a protected form of speech.[28]

These cases share a common theme: private actors have a right to control their message, which includes deciding what speech to include and/or exclude. The same principle applies to social media platforms. They decide what content to remove, what to promote, and how to rank or label posts.

19 *Id.*
20 *Id.*
21 Hurley v. Irish-Am. Gay, Lesbian and Bisexual Group of Boston, 515 U.S. 557, 574 (1995).
22 *Id.* at 573.
23 Manhattan Cmty. Access Corp. v. Halleck, 139 S. Ct. 1921, 1930 (2019).
24 *Id.* at 1930.
25 303 Creative LLC v. Elenis, 600 U.S. 570 (2023).
26 *Id.*
27 *Id.* at 588.
28 *See id.* at 585-86.

Sometimes this is done manually, but more often it's done through algorithms designed by people who set rules and priorities (and sometimes via a hybrid of both).

In ***Moody v. NetChoice and Paxton v NetChoice*** (2024), the Supreme Court affirmed that platforms like Facebook and YouTube are protected by the First Amendment when they make editorial decisions about user-generated content, including whether to publish, remove, promote, or demote that content through algorithmic systems.[29] Although the Court's formal holding focused on procedural rules for facial challenges, the majority issued a strong and explicit defense of editorial discretion in the context of social media platforms and their exercise of editorial functions.[30] Citing precedent from *Tornillo, Hurley,* and *PG&E*, the Court held that platforms enjoy the same constitutional right to curate and arrange third-party speech as traditional media.[31] It rejected the Fifth Circuit's view that social media companies were mere conduits lacking editorial rights and emphasized that when the government interferes with such editorial choices, it violates the First Amendment.[32]

Importantly, the Court made clear that these protections apply not only to the selection of content but also to how it is displayed, including when platforms use algorithms to personalize and organize feeds.[33] Decisions about how the display will be ordered and organized, and whether to present content the way the platform wants, are part of the expressive conduct that the First Amendment shields from government control.[34] Even a focused editorial choice, the Court said, is protected.[35] Platforms do not lose their rights simply because they carry a wide range of user speech while removing only some.[36]

Justice Barrett did briefly acknowledge in her concurrence an unresolved edge case.[37] She considered whether an algorithm that operates entirely without human guidance, responding only to user behavior with no input from editorial standards, might fall outside First Amendment protection.[38] But that door is narrow and mostly theoretical. As the Court noted, platforms typically design their algorithms based on policies, priorities, and expressive judgments set by human teams.[39] Platform content policies that inform the algorithmic design are often the result of extensive human deliberation, both internally within a company and often involving input from external stakeholders

29 Moody v. NetChoice, LLC, 144 S.Ct. 2383 (2024).
30 *Id.*
31 *Id.* at 729-731.
32 *Id.* at 727-728.
33 *Id.* at 710-711.
34 *Id.* at 729
35 *Id.* at 733.
36 *Id.* at 740.
37 *Id* "respond solely to how users act online … without any regard to independent content standards" established by humans. 2024 WL 3237685 at n.5 (Barrett, J., concurring).
38 *Id.*
39 *Id.* at 746.

like civil society, advertisers, researchers, regulators, and others.[40] So long as those human choices shape how content is curated and presented, the system remains an extension of the platform's editorial voice and remains protected.[41] In practice, there is no serious question that today's platforms' content moderation systems, even when automated, are designed and directed by human intent. That is the core of what the First Amendment safeguards.

The key idea that stems from these First Amendment precedents is this: the First Amendment does not only protect words. It protects choices about speech, including how it is arranged, what is emphasized, and what is excluded. This protection applies whether the speech appears in a newspaper, a parade, or an algorithmically ranked social media feed.

3. SECTION 230'S ROLE IN CONTENT MODERATION

Section 230 of the Communications Decency Act, codified at 47 U.S.C. § 230, is the backbone of the modern Internet's legal architecture.[42] The statute consists of two core provisions. Section 230(c)(1) states that "[n]o 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."[43] Courts have uniformly interpreted this language to bar most forms of liability based on content a third party created, even if the platform hosts, displays, or organizes that content.[44]

Meanwhile, Section 230(c)(2) provides that no provider shall be held liable for voluntary, good faith efforts to restrict access to material considered by the provider to be "obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable."[45] This language empowers platforms to moderate as they see fit without fear of incurring liability for choosing what to take down.

---

40 Kate Klonick, The New Governors: The People, Rules, and Processes Governing Online Speech, 131 Harv. L. Rev. 1598 1649-1658 (2018).

41 *Id.*

42 47 U.S.C. § 230.

43 47 U.S.C. § 230(c)(1).

44 Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1098-99 (9th Cir. 2019) (analyzing that plaintiffs could not frame "website features as content" and that the site's recommendation and notification functions did not materially contribute to alleged unlawfulness of content); Force v. Facebook, Inc., 934 F.3d 53, 68 (2d Cir. 2019) (analyzing whether claims against Facebook for promoting particular content would make Facebook liable for information provided by another information content provider); Marshall's Locksmith Serv., 925 F.3d at 1271 (declining to treat search engines' conversion of fraudulent addresses from webpages into "map pinpoints" as developing content).

45 47 U.S.C. § 230(c)(2).

Congress passed Section 230 to overrule the holding in *Stratton Oakmont, Inc. v. Prodigy*[46], where a platform was treated like a publisher, and held just as liable for user content as the user because it had attempted to moderate content.[47] *Prodigy* created a perverse incentive: a platform could minimize its legal risk by doing nothing at all.[48] Section 230 changed that by allowing platforms to moderate while remaining shielded from the legal consequences of third-party speech.[49] The courts have largely adopted a broad reading of the (c)(1)protection.

- In ***Zeran v. AOL***, the Fourth Circuit held that AOL could not be held liable for messages posted by a user, even after being notified of their falsity.[50] The court emphasized that imposing liability on intermediaries would lead to a "chilling effect" on speech and would incentivize platforms to restrict access to lawful but controversial expression. This approach has continued to hold across circuits.[51]

- In ***Dyroff v. Ultimate Software Group***, the Ninth Circuit found that an online platform's algorithmic recommendations, based on user engagement data, did not strip it of Section 230 immunity.[52] The platform had not created the content at issue, and its recommendation system merely automated the delivery of user-generated posts.[53]

- Similarly, in ***Force v. Facebook***, the Second Circuit concluded that Facebook's algorithmic promotion of user content did not amount to its own speech or render it liable under the Anti-Terrorism Act.[54] The court reasoned that organizing and delivering third-party content, even algorithmically, falls squarely within the immunity granted by Section 230.[55]

- Recently, in ***Gonzalez v. Google,*** the Court declined to decide whether Section 230 shields algorithmic recommendations, remanding the case in light of its unanimous decision in ***Twitter v. Taamneh***.[56] There, the Court held that platforms do not incur liability under the Anti-Terrorism Act for merely hosting or recommending content when their relationship to the underlying conduct is remote and passive.[57]

---

46 Stratton Oakmont, Inc. v. Prodigy Services Co., 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995)
47 *See generally* Jeff Kosseff, The Twenty-Six Words That Created The Internet (2019).
48 *Id.*
49 *Id.*
50 Zeran v. AOL, 129 F.3d 327 (4th Cir. 1997).
51 *Id.*
52 Dyroff v. Ultimate Software Group, Inc., 934 F.3d 1093 (9th Cir. 2019).
53 *Id.*
54 Force v. Facebook, Inc., 934 F.3d 53 (2d Cir. 2019).
55 *Id.*
56 Twitter, Inc. v. Taamneh, 598 U.S. 471 (2023), Gonzalez v. Google LLC, 598 U.S. 617 (2023)
57 Daphne Keller, *The Long Reach of Taamneh: Carriage and Removal Requirements for Internet Platforms,* Brookings Inst. (Oct. 19, 2023), https://www.brookings.edu/articles/the-long-reach-of-taamneh-carriage-and-removal-requirements-for-internet-platforms/.

Courts have also affirmed that platforms may remove or decline to remove content without incurring liability:

- In ***Barnes v. Yahoo!***, the Ninth Circuit held that Yahoo! could not be sued for failing to remove harmful content posted by a third party, even after allegedly promising to do so.[58] The decision reinforced that editorial choices of whether to take down or leave up content fall within Section 230's protections.

- Similarly, in ***Domen v. Vimeo***, the Second Circuit upheld Vimeo's removal of videos promoting conversion therapy, rejecting claims of discrimination and emphasizing that platforms may enforce content policies without becoming liable for viewpoint exclusion.[59]

These cases reaffirm that platforms are not responsible for user content simply because they arrange, amplify, or recommend it. That includes the use of algorithms that personalize feeds or suggest related posts. Courts have consistently viewed such systems as tools that help deliver user speech, not as creators of speech themselves.

58 Barnes v. Yahoo!, Inc., 570 F.3d 1096, 1103 (9th Cir. 2009).

59 Eric Goldman, *For the Third Time, a Second Circuit Panel Dismisses an Online Account Termination Case–Domen v. Vimeo,* Tech. & Mktg. L. Blog (Sept. 27, 2021), *available at* https://blog.ericgoldman.org/archives/2021/09/for-the-third-time-a-second-circuit-panel-dismisses-an-online-account-termination-case-domen-v-vimeo.htm.

4. WHAT IF AN ALGORITHM PROMOTES HARMFUL CONTENT?

While courts have long held that Section 230 protects platforms from liability for third-party content, they have also recognized a limit: if a platform materially contributes to the development of unlawful content, it may be treated as an information content provider.[60] "Material contribution" refers to actions that transform or meaningfully shape the content's illegality.[61]

But legal challenges to this current framework are growing. The theory is that, when a platform's recommendation system actively promotes harmful content in a way that allegedly causes injury, the platform is no longer a passive intermediary and thus should not be entitled to immunity.[62] In ***TikTok v. Anderson***, the Third Circuit considered whether TikTok could claim Section 230 immunity in a lawsuit brought by parents whose child tragically died after engaging in a dangerous viral challenge allegedly promoted by TikTok's algorithm.[63]

The Third Circuit held that Section 230 does not shield platforms from liability in these cases, and interpreted ***NetChoice v Moody*** to mean that algorithmically served content is a platform's own speech, and so when the claim targets the platform's own speech, specifically, the algorithmic promotion of harmful content, the platform is liable.[64] The Third Circuit distinguished its ruling from decisions like ***Dyroff v. Ultimate Software Group and Force v. Facebook***, concluding that TikTok's algorithm functioned as more than a neutral tool. It said that TikTok affirmatively directed the content to the child, raising questions of product liability and design negligence that fall outside Section 230's protections. This ruling creates a notable circuit split and signals a judicial willingness to reconsider the limits of platform liability.

But when the *Anderson* Court says TikTok's algorithm is "more than a neutral tool," it suggests that mere recommendation somehow crosses the line into content creation. However, it identifies no action by the platform that materially altered or added to the content itself. That stretches the concept of "material contribution" far beyond its established legal meaning. As a result, that risks collapsing

60 *See, e.g.*, Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1166, 1173–74 (9th Cir. 2008) (holding that website was an information content provider with respect to user preferences the website helped "develop" through mandatory questionnaires, but was not an information content provider with respect to information provided in a freeform text box); *See* FTC v. Accusearch, Inc., 570 F.3d 1187, 1200 (10th Cir. 2008) (holding that website materially contributed to alleged illegality of conduct when it collected and published confidential telephone records).

61 *Id.*

62 TikTok v. Anderson, 82 F.4th 459 (3d Cir. 2024).

63 *Id.*

64 *Id.*

the distinction between facilitation and authorship that Section 230 is designed to preserve.

While the First Amendment may still protect algorithmic curation as a form of editorial judgment, ***TikTok v. Anderson*** introduces legal uncertainty about where Section 230 ends and tort liability begins. The Section 230 application distinction made by the Third Circuit is deeply flawed. It conflates content presentation with content creation and misreads the nature of editorial discretion. Algorithms are simply scalable tools of editorial judgment. They implement platform policies to surface content that aligns with user preferences or community guidelines. Moreover, drawing a legal line between hosting and recommending content is unworkable. Platforms cannot realistically host third-party user speech without ordering and filtering what users see. An algorithm that highlights a trending topic or suggests a post based on user interest does not transform third-party speech into the platform's own. If that were the rule, no platform would be safe from liability because every act of curation could be reframed as a recommendation and therefore as a platform's own speech. Any ruling that penalizes recommendation tools would invite litigation against every form of algorithmic feed, whether a trending tab, a "For You Page," or a chronological list.

The emotional gravity of cases like this cannot and should not be ignored. It is precisely because the stakes are so high that the law and policymaking must respond with clarity, not a reactionary approach. The core question is not whether harm exists online; rather, it is whether the proposed solutions preserve the structure of free expression in the process of responding to that harm. By allowing platforms to host and organize third-party content without fear of liability, Section 230 ensures the internet remains an open and diverse forum.

Weakening Section 230 would not only expose platforms to crushing litigation risk but also deter them from engaging in content moderation altogether. That would produce one of two results: either rampant under-moderation, which enables abuse and legitimately harmful content, or over-moderation, which erases lawful but unpopular speech.

5. WHAT HAPPENS IF WE GET THIS WRONG?

Some argue that platforms use curation to silence disfavored viewpoints or manipulate political discourse.[65] But compelling platforms to carry speech against their editorial judgment would transform them from hosts into mouthpieces for those who hold the levers of power. Forcing platforms to carry certain viewpoints or structure their feeds in government-approved ways would violate settled First Amendment doctrine as laid out above.

The Third Circuit's decision in ***TikTok v. Anderson signals a shift in how some courts may treat algorithmic conduct and introduces uncertainty.*** It invites litigation that could severely narrow the scope of platform immunity, fundamentally alter how content moderation operates at scale, and drastically change the Internet.

If Section 230 protections or First Amendment editorial rights are rolled back, the consequences could be profound. Some platforms will over-moderate to avoid legal exposure, removing lawful but controversial content. Others will under-moderate, allowing harmful content to spread unchecked. Such a shift will not harm the powerful but the vulnerable, the dissenters, and the voices that depend on intermediaries to be heard.[66] Smaller platforms and start-ups may shut down or avoid hosting speech and change their business models altogether due to litigation risk.[67]

Child safety may also decline. Platforms use algorithms to demote age-inappropriate material and highlight trusted content. If recommendation tools are regulated out of existence, these safety functions disappear. And without legal protection, platforms may hesitate to remove dangerous content out of fear that their own moderation efforts could trigger liability. If platforms must rely primarily on automated systems to manage liability, children will not be shielded more effectively; instead, entire categories of educational or supportive content could be swept away while dangerous material slips through.

---

65 Jake Hawkes, Are Google and Facebook Really Suppressing Conservative Politics?, The Guardian (Dec. 4, 2018), https://www.theguardian.com/technology/2018/dec/04/google-facebook-anti-conservative-bias-claims; *See* Shannon Bond, Facebook and Twitter Limit Sharing 'New York Post' Story About Joe Biden, NPR (Oct. 14, 2020, 6:49 PM), https://www.npr.org/2020/10/14/923766097/facebook-and-twitter-limit-sharing-new-york-post-story-about-joe-biden.

66 Fight for the Future, Democrats Are Ignoring Marginalized Communities with Proposed Bill That Carves Up Section 230 (Oct. 14, 2021) (press release), *available at* https://www.fightforthefuture.org/news/2021-10-14-democrats-are-ignoring-marginalized-communities-with-proposed-bill-that-carves-up-section-230/; Fight for the Future, If the Supreme Court Undermines Section 230, Marginalized People Will Pay the Price (Oct. 3, 2022) (press release), *available at* https://www.fightforthefuture.org/news/2022-10-03-if-the-supreme-court-undermines-section-230-marginalized-people-will-pay-the-price.

67 Ethan Wham, *The Economic Case for Section 230, Disruptive Competition Project*, Computer & Communications Industry Association (Sept. 6, 2019), *available at* https://www.project-disco.org/innovation/090619-an-economic-case-for-section-230/.

Critics often claim algorithms are at the core of mis- and disinformation and harmful content problems.[68] But personalized curation (for example, Pinterest conducts user surveys to personalize their feeds)[69], when paired with transparent standards and algorithmic design, can be part of the solution, helping to surface content users want to see, detect patterns of harm, and shield vulnerable users from abuse.

Solutions that don't lead us down the path of giving the government and those with more resources to sue tools to suppress speech? Rather than defaulting to top-down removal or centralized control, platforms could enable systems that allow users to choose from competing recommendation services that reflect different values and priorities. Platforms can expose the social provenance of content by labeling which communities engage with, endorse, or contest a given post.[70] Instead of removing controversial speech, platforms could label whether a piece of content reflects shared consensus across diverse audiences or primarily circulates within a narrow group.[71] This kind of context helps users assess credibility and relevance without suppressing expression.[72]

Another promising solution to content moderation frustrations is having middleware[73] as an additional layer of the Internet social media ecosystem.[74] Middleware refers to independent tools or services that users can choose to layer on top of existing platforms to filter, rank, or interpret content according to their own values and preferences.[75] But this kind of experimentation would be chilled or even foreclosed entirely if the *Anderson* decision were broadly adopted. By treating algorithmic curation as content creation, the decision threatens to saddle both platforms and middleware providers with potential liability for third-party speech. And without strong Section 230 protections, platforms would lack the legal breathing room to partner with or enable middleware in the first place, undermining one of the most promising paths forward in making content moderation better for society.

---

68 *See* Paulo Carvão, *The Supreme Court Has Spoken on Gonzalez v Google – Now It's Congress's Turn To Address Section 230,* Harv. ALI Soc. Impact Rev. (May 30, 2023), https://www.sir.advancedleadership.harvard.edu/articles/supreme-court-spoken-gonzalez-v-google-now-congresss-turn-section-230.

69 Leif Sigerson, Stephanie Chen & Wendy Matheny, Healthier Personalization with Surveys, Pinterest Engineering Blog (Medium) (May 30, 2025), *available at* https://medium.com/pinterest-engineering/healthier-personalization-with-surveys-65177cf9bea8.

70 Glen Weyl, Audrey Tang & Jacob Mchangama, *Building a Prosocial Media Ecosystem*, Noema (May 1, 2025), https://www.noemamag.com/building-a-prosocial-media-ecosystem/.

71 *Id.*

72 *Id.*

73 Mike Masnick, *Protocols, Not Platforms: A Technological Approach to Free Speech,* Knight First Amend. Inst. (Aug. 21, 2019), https://knightcolumbia.org/content/protocols-not-platforms-a-technological-approach-to-free-speech.

74 Daphne Keller, *The Future of Platform Power: Making Middleware Work,* 32 J. Democracy 168 (2021), https://www.journalofdemocracy.org/articles/the-future-of-platform-power-making-middleware-work/.

75 Red Hat, What Is Middleware?, Red Hat (Dec. 16, 2022), *available at* https://www.redhat.com/en/topics/middleware/what-is-middleware.

The bottom line is this: without Section 230, those who are most vulnerable will have fewer spaces online to speak, connect, and be heard. Preserving the combination of constitutional (First Amendment) and statutory (Section 230) protections is not a gift to social media companies; it is a commitment to free expression in the digital age and to the ability of platforms to responsibly protect vulnerable users.

THE FUTURE OF FREE SPEECH

www.futurefreespeech.org