

April 29, 2025

The Implications of Section 230 for Black Communities

*Opportunities and Challenges for Black Communities
Due to Content Moderation*



Danielle A. Davis, Esq.
Director, Tech Policy



Introduction

This series of issue briefs provides a summary of “*The Implications of Section 230 for Black Communities*” law review article that examines Section 230 of the Communications Decency Act (Section 230), which grants online platforms immunity from liability for user-generated content. The first brief explored the dual-edged nature of the statute, analyzing both the opportunities and challenges it presents for Black communities by protecting tech companies from liability for third-party content. The first brief highlighted how this immunity from liability for third-party content empowers Black communities by enabling free expression and fostering innovation, while also examining how the same immunity allows harmful practices to persist. This second brief will examine the nuances and legal implications of the statute, specifically how it affects platform moderation practices, including potential biases in content moderation that often disproportionately impact Black users. The third brief will explore the implications of various Section 230 reform proposals with a focus on their unique impact on Black communities ■

Building upon the examination of Section 230’s immunity from liability for third-party content in our first brief, it’s important to also examine Section 230’s content moderation immunity, which provides protection for online platforms, such as Facebook, when they engage in content moderation.¹

Courts have interpreted Section 230(c)(1), in combination with the explicit protections in Section 230(c)(2), to grant platforms broad content moderation immunity for their decisions. As a result, private platforms can proactively remove or downrank harmful content that affects Black communities, while also amplifying content that counters anti-Black racism and disinformation.² However, Section 230’s content moderation immunity presents not only these benefits but also significant challenges for Black users.

The reason: Section 230’s content moderation immunity also gives

platforms the ability to engage in discriminatory content moderation such as removing content from Black creators, even if it follows the guidelines.³ Content moderation immunity also provides platforms with immunity for choosing not to remove harmful anti-Black content for reasons such as boosting user engagement, increasing ad revenue, or responding to political pressure.

This second issue brief will first explore the positive ways in which Section 230’s content moderation immunity enables platforms to reduce disinformation, prevent discrimination through updated practices and algorithms, and combat hate speech. Then, this issue brief will examine the negative impacts of content moderation on Black communities, including the challenges of discriminatory content moderation and the promotion of anti-Black activities for revenue or political gain.

The Positive Impact of Content Moderation on Black Communities

Section 230’s immunity for content moderation has arguably enabled tech platforms to create virtual spaces that are safer and more welcoming for Black communities. For example, this immunity provides platforms with the legal protection to downrank or remove harmful content, such as harassment, incitements to violence, and discrimination, creating a safer digital space for Black users. This protection extends beyond unlawful content to include harmful material that, while not illegal, can negatively affect Black communities through misinformation and hate speech.

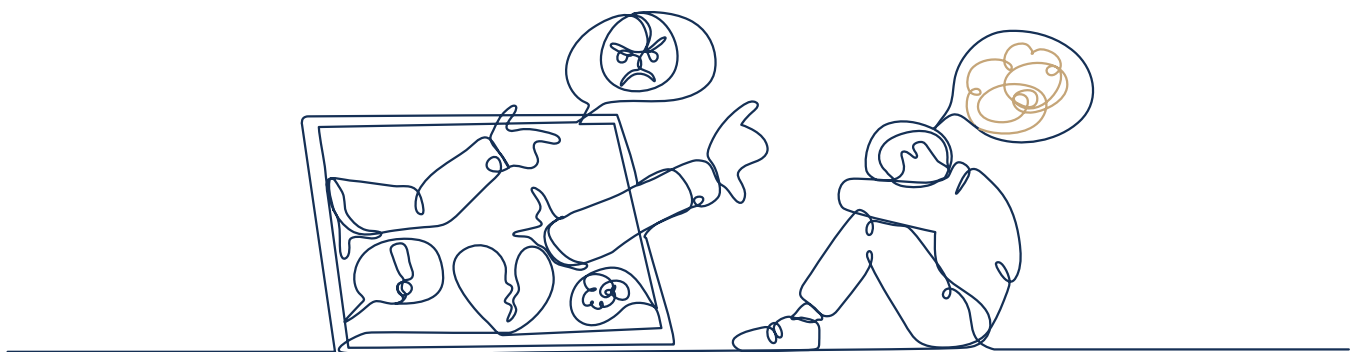
This section of the issue brief explores a specific type of content that content moderation immunity allows platforms to moderate with significant freedom: “lawful-but-awful” activities. These include the spread of racialized disinformation and harmful ideologies, such as non-Black users pretending to be Black, encouraging Black communities not to vote, or promoting white supremacist theories. While this content might be constitutionally protected from government regulation,⁴ content moderation immunity gives platforms the explicit protection to remove or downrank it.

The First Amendment allows social media platforms to manage content as they see fit, but content moderation immunity empowers them to do so swiftly and decisively.⁵ This Section 230 immunity empowers platforms to address key challenges faced by Black communities online, including reducing disinformation, preventing discrimination through updated practices and algorithms, and combating hate speech and white supremacist content.

Reducing Disinformation

Section 230 content moderation immunity provides social media platforms with the ability to combat disinformation, including content that is specifically aimed at Black users, by downranking or removing such content.

Disinformation refers to deliberately false or misleading information created to deceive and manipulate, often targeting sensitive issues including race, social justice, or the integrity of the electoral process. For example, between 2016 and 2019, non-Black extremists infiltrated the #Blaxit (Black exit) movement within the Black community, using fake accounts to pose as Black users. They spread misleading memes and crafted false narratives to simulate a growing African repatriation movement among Black Americans, an act described by one observer as “[] catfishing an entire race.”⁶ This tactic, known as digital blackface, was used to sow confusion and undermine authentic movements.



Digital blackface, while harmful, is a form of expression likely protected by the First Amendment, making government restrictions difficult.⁷ However, Section 230's content moderation immunity allows platforms to moderate content by removing users engaged in such deceptive practices.⁸ Many platforms also have policies against identity misrepresentation.⁹

Section 230's content moderation immunity empowers platforms to manage disinformation that aims to suppress Black voter turnout, such as deceptive messages from political candidates.¹⁰ Platforms are protected from liability when they choose to remove or downrank false allegations of voter fraud,¹¹ which were often aimed at the legitimacy of Black votes after the 2020 presidential election.¹² Similarly, this immunity allows platforms to remove content discouraging participation in crucial processes such as the U.S. Census,¹³ where Black individuals have been historically undercounted.¹⁴

Platforms also utilize Section 230's content moderation immunity to combat medical misinformation. Users spread dangerous falsehoods during the early stages of the COVID-19 pandemic,¹⁵ which disproportionately affected Black communities. By leveraging this immunity, platforms could swiftly remove misleading content that endangered public health, contributing to a more informed and safer online environment.

Updating Practices and Algorithms to Prevent Discrimination

Section 230's content moderation immunity empowers platforms to refine their policies, algorithms, and tools to combat discrimination more effectively. This includes identifying, downranking, or removing content linked to anti-Black users, hate speech, white supremacist activities, and racialized disinformation.

For example, after evidence revealed that Airbnb hosts were disproportionately rejecting Black guests,¹⁶ the platform required all users to commit to treating everyone in its community with respect and without bias, regardless of race, religion, ethnicity, disability, gender identity, sexual orientation, or age.¹⁷ As a result, Airbnb has denied access to its platform or removed more than 2.5 million users who refused to comply with this commitment.¹⁸

Airbnb also introduced measures to reduce discrimination, such as displaying guest photos only after a booking is accepted and allowing most bookings to be made without host approval. Airbnb also improved their reporting system for inappropriate content in messages and listings.¹⁹ Furthermore, Airbnb responded to findings that non-Black hosts were out-earning Black hosts²⁰ by implementing a machine learning tool to help hosts set competitive prices.²¹ This tool has significantly narrowed, though not entirely closed, the earnings gap between Black and white hosts.²²

Removing Hate Speech and White Supremacy

The First Amendment generally prevents the government from restricting hate speech unless it incites imminent illegal actions or poses specific threats of violence toward an individual or group.²³ However, the amendment also protects private entities' rights to disassociate from or refuse to promote such speech. Section 230's content moderation immunity further empowers platforms by allowing them to reduce the visibility of or remove hate speech and white supremacist content and to ban users who spread these harmful ideologies.²⁴

For years, white supremacist groups have used online platforms for various activities, including recruitment, radicalization, planning, and spreading the ideology of white cultural and genetic superiority—often for free.²⁵ From the early days of bulletin boards in the 1980s²⁶ to mainstream platforms including Facebook, Myspace, YouTube, and X in the 2000s,²⁷ these groups have used the internet to expand their reach.

In response to public pressure, major platforms have strengthened their content moderation policies and filtering systems.²⁸ Section 230 has played a crucial role in enabling these platforms to make these improvements. For example, Meta, the parent company of Facebook, defines hate speech as including “violent or dehumanizing speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing, and calls for exclusion or segregation.”²⁹ Many platforms also prohibit dehumanizing comparisons, such as likening individuals to animals or criminals, perpetuating harmful stereotypes including blackface,³⁰ and using symbols such as swastikas or inciting harassment or violence against protected groups.³¹

The Negative Impact of Content Moderation on Black Communities

While Section 230’s support for content moderation provides important benefits to Black communities, it also presents significant challenges. The broad immunity it grants allows platforms to disproportionately remove content from Black creators and apply potentially discriminatory moderation algorithms with minimal legal consequences.³² This subsection complicates efforts by Black communities to challenge moderation decisions that permit the spread of anti-Black content, even when such content clearly violates the platform’s own guidelines. This issue may arise from platforms’ indifference or their desire to cater to advertisers or influential political groups.

Section 230’s content moderation immunity also shields tech companies from having to disclose the reasoning behind their moderation decisions. Unlike newspapers or other private entities, tech companies enjoy broad deference under this immunity,³³ allowing them to make moderation decisions with little accountability. This lack of transparency—often described as a “black box”—makes it difficult for Black communities to understand or effectively contest these decisions, whether in court or public discourse.³⁴ As a result, Black users are often left in the dark about why certain content, such as a Black Lives Matter video, was removed and whether community guidelines are being applied consistently or fairly.³⁵

Even without Section 230’s content moderation immunity, platforms have significant autonomy under the First Amendment to moderate content as they see fit.³⁶ This autonomy could lead to the promotion of anti-Black content to increase user engagement or satisfy powerful regulatory bodies. Discrimination against Black users in content moderation is likely to persist, particularly in states without laws or court rulings that apply public accommodations laws to online platforms.³⁷

Understanding the interplay between the First Amendment and the limitations of legal protections against online discrimination is essential when considering potential legal reforms or incentives to improve content moderation practices for Black communities.

Discriminatory Content Moderation

Section 230’s content moderation immunity grants platforms significant authority, leading to the disproportionate removal and downranking of content from Black users and the implementation of discriminatory moderation algorithms. Studies reveal that Black individuals face higher rates of account removals compared to white users, despite following guidelines. Across multiple platforms (Facebook, Instagram, X, and YouTube, among others), Black users are 60 percent more likely to experience account removals.³⁸ Internal Facebook research also showed Black Instagram users in the United States were 50 percent more likely to have their accounts automatically disabled.³⁹

Content moderation systems, often based on flawed algorithms, fail to recognize cultural nuances, resulting in the disproportionate silencing of Black voices.⁴⁰ Platforms including TikTok and X have faced criticism for algorithms that unfairly target phrases associated with Black activism, such as flagging “Black Lives Matter” while ignoring “white supremacy.”⁴¹

In the future, platforms could use Section 230’s content moderation immunity protections to further downrank or remove content related to racism, systemic inequality, and diversity. This is particularly concerning as states enact legislation limiting discussions on racial justice and related topics.⁴² Section 230’s immunity has already been used to restrict information on sensitive issues such as abortion,⁴³ suggesting that similar restrictions could be applied to content on racial injustice and diversity, especially in states with restrictive policies.

Promoting Anti-Black Activity for Revenue or Political Gain

Section 230’s broad authority limits the options available to Black communities when platforms violate their own guidelines and allow the spread of anti-Black content. Although platforms are protected under the First Amendment, the promotion of harmful materials raises significant concerns about the costs to Black communities, which must be considered alongside any benefits.

Despite public pressure to combat disinformation and hate speech, mainstream platforms may hesitate to curb white supremacist content due to fears of being perceived as biased against conservative voices.⁴⁴ In some cases, executives have overridden their own content moderation teams to weaken policies on misinformation and hate speech, particularly to avoid alienating conservative users.⁴⁵ Financial incentives also play a role, as divisive and hateful content often drives higher user engagement and ad revenue, leading some platforms to promote anti-Black content.⁴⁶

Additionally, Section 230’s content moderation immunity provides white supremacists with a form of subsidy, enabling them to establish and profit from extremist platforms. As mainstream platforms tighten their moderation policies, many white supremacists have migrated to fringe platforms such as 8chan, Gab, MeWe, Rumble, Signal, and Voat, where overtly racist and violent content is allowed.⁴⁷ While the First Amendment protects the creation of these platforms, Section 230 grants them further procedural and substantive protections, such as immunity from liability for defamation, deceptive practices, unfair advertising, and coordinated harassment campaigns.

Conclusion

Section 230’s content moderation immunity provides online platforms with powerful tools to shape digital environments—tools that have both positive and negative implications for Black communities. On the one hand, this immunity empowers platforms to reduce disinformation, prevent discrimination through updated practices and algorithms, and combat hate speech and white supremacist content. On the other hand, it also enables discriminatory content moderation, lack of transparency, and the amplification of anti-Black content for revenue or political gain. As policymakers consider the future of Section 230, any reforms must carefully weigh its benefits against the structural harms it perpetuates, ensuring that content moderation practices promote equity and accountability for Black users.

In the next issue brief, we will explore the potential impact of Section 230 reforms on Black communities, examining how changes to this law could reshape the digital landscape and affect the online experiences of Black users.

Endnotes

- ¹ 47 U.S.C. § 230(c)(2).
- ² Section 230(c)(2) explicitly states that platforms shall not be held liable for voluntary actions 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.” 47 U.S.C. § 230(c)(2). Following the Fourth Circuit’s decision in *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 330 (4th Cir. 1997), courts have widely interpreted Section 230(c)(1) to empower platforms to both “leave up” and “take down” (or uprank or downrank) third-party content. *Zeran*, 129 F.3d at 330.
- ³ Arguably there is a point at which such activity may cross a line and violate public accommodations laws or section 1981 if it amounts to purposeful discrimination against Black patrons of that business. See 42 U.S.C. § 2000a(a) (“All persons shall be entitled to the full and equal enjoyment of the goods, services, facilities, privileges, advantages, and accommodations of any place of public accommodation, as defined in this section, without discrimination on the ground of race, color, religion, or national origin.”); *id.* § 1981(a) (“All persons . . . shall have the same right in every State and Territory to make and enforce contracts . . . as is enjoyed by white citizens.”).
- ⁴ See *Platform Accountability Hearing* (Testimony of Mary Anne Franks, Professor of Law).
- ⁵ Eric Goldman, *Why Section 230 Is Better Than the First Amendment*, 95 Notre Dame L. Rev. Reflection 33, 36–39 (2019).
- ⁶ Brandi Collins-Dexter, *Butterfly Attack: Operation Blaxit*, The Media Manipulation Casebook (Oct. 16, 2020).
- ⁷ Speech does not lose First Amendment protection simply because it is false outside of special contexts (e.g., impersonating an officer, committing financial fraud, perjury). See *United States v. Alvarez*, 567 U.S. 709 (2012).
- ⁸ See Whitney Tesi, *When Disinformation Becomes “Racialized,”* ABC News (Feb. 5, 2022).
- ⁹ See, e.g., *Account Integrity and Authentic Identity*, Meta (last visited Feb. 10, 2024); *Impersonation policy*, YouTube Help (last visited Feb. 10, 2024); *Community Guidelines—Spam and Deceptive Account Behaviors*, TikTok (last visited July 17, 2023).
- ¹⁰ See *Care Comm. v. Arneson*, 766 F.3d 774, 785 (2014); *Susan B. Anthony List v. Driehaus*, 573 U.S. 149 (2014), remanded to 814 F.3d 466 (6th Cir. 2014); *Alvarez*, 567 U.S. at 719.
- ¹¹ Many platforms have guidelines prohibiting the posting of content aiming to mislead voters about the time, place, means, or eligibility requirements for voting, false claims to discourage voting, false claims of voter fraud, and incitement to interfere with voting procedures or democratic processes. See, e.g., *Elections Misinformation Policies*, YouTube Help (last visited Feb. 10, 2024).
- ¹² Ashley Nguyen, et al., *Anger Builds in Black Community Over Trump’s Claims of Voter Fraud in Big Cities*, *The Washington Post* (Nov. 20, 2020); Brandon Tensley, *The Racist Rhetoric Behind Accusing Largely Black Cities of Voter Fraud*, CNN (Nov. 20, 2020).
- ¹³ Many platforms have guidelines prohibiting suppression of census participation. *Misinformation Policies – Suppression of Census Participation*, YouTube Help (last visited Feb. 10, 2024); *Misinformation: Voter or Census Interference*, Meta (last visited July 17, 2023); *Community Guidelines—Civic and Election Integrity*, TikTok (last visited Feb. 10, 2024).
- ¹⁴ Eric Jensen & Timothy Kennel, *Detailed Coverage Estimates for the 2020 Census Released Today*, U.S. Census Bureau (Mar. 10, 2022).
- ¹⁵ Nambi Ndugga, et al., *COVID-19 Cases and Deaths, Vaccinations, and Treatments by Race/Ethnicity as of Fall 2022*, Kaiser Family Foundation. (Nov. 17, 2022).
- ¹⁶ See, e.g., Benjamin Edelman, et al., *Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment*, 9 AM. ECON. J.: Applied Econ 1, 2 (2017) (finding that guests with distinctively Black American names were 16% less likely to be accepted relative to identical guests with distinctively white names).
- ¹⁷ Airbnb, *Community Policy: General questions about the Airbnb Community Commitment* (2016) (last visited Feb. 10, 2024).
- ¹⁸ Airbnb, *A Six-Year Update on Airbnb’s Work to Fight Discrimination and Build Inclusion* (Dec. 13, 2022) at 7.
- ¹⁹ Laura Murphy, *The Rationale for and Key Elements of a Business Civil Rights Audit (2021) at 29*; Katie Benner, *Airbnb Adopts Rules to Fight Discrimination by Its Hosts*, *The New York Times* (Sept. 8, 2016); Airbnb, *General Questions About the Airbnb Community Commitment* (last visited Feb. 10, 2024) (requiring all within the Airbnb community to treat everyone with respect and without bias regardless of protected class status); see also *Airbnb Will Change Process to Fight Discrimination in Oregon*, U.S. News (Jan. 6, 2022) (responding to a lawsuit brought by three African American women by indicating that hosts in Oregon will start seeing an Oregon guest’s initials in place of the guest’s first name until a booking request is confirmed).
- ²⁰ Benjamin Edelman and Michael Luca, *Digital Discrimination: The Case of Airbnb.com* § 1 (Harv. Bus. Sch., Working Paper No. 14-054, 2014) (finding that non-Black hosts in New York City could charge 12% more on average than Black hosts).
- ²¹ Lane Lambert, *White Airbnb Hosts Earn More. Can AI Shrink the Racial Gap?*, *Forbes* (Aug. 11, 2021).
- ²² *Id.*
- ²³ *Brandenburg v. Ohio*, 395 U.S. 444, 447 (1969); *NAACP v. Claiborne Hardware*, 458 U.S. 886, 928 (1982); *United States v. White*, 670 F.3d 498, 513 (4th Cir. 2012).
- ²⁴ See *Twitter, Inc. v. Super. Ct. ex rel. Taylor*, Case No. A154973 (Cal. Ct. App. Aug. 17, 2018) (finding that Section 230 barred an unfair trade practices claim by a white supremacist Twitter user whose account was permanently suspended for violating “the Twitter Rules against being affiliated with a violent extremist group”); *White nationalist Jared Taylor Sues Twitter Over Ban*, BBC (Feb. 22, 2018).

- ²⁵ Daniel L. Byman, *When Hate Goes Viral*, Foreign Policy (Mar. 23, 2022); See *With Hate in their Hearts: The State of White Supremacy in the United States*, ADL (Mar. 3, 2017).
- ²⁶ See Heather J. Williams, et al., *The Online Extremist Ecosystem*, RAND, 1, 5 (Dec. 2021).
- ²⁷ See *id.*, at 6.
- ²⁸ See *Facebook Community Standards: Hate Speech*, Meta (last visited Feb. 10, 2024); *YouTube Hate Speech Policy*, YouTube Help (last visited Feb. 10, 2024); *Hateful Conduct*, X: Help Center (last visited Feb. 10, 2024); *Community Guidelines*, TikTok (last visited Feb. 10, 2024). Many platforms also have guidelines restricting dangerous organizations. See *Violent extremist or criminal organizations policy*, YouTube Help (last visited Feb. 10, 2024); *Dangerous Organizations and Individuals*, Meta (last visited Feb. 10, 2024); *Violent and Hateful Organizations and Individuals*, Meta (last updated Feb. 10, 2024).
- ²⁹ See, e.g., *Facebook Community Standards: Hate Speech*, Meta: Transparency Center (last visited Feb. 10, 2024).
- ³⁰ *Id.*
- ³¹ See *Hateful Conduct*, X: Help Center (last visited Feb. 10, 2024).
- ³² See Reducing Information Discussion.
- ³³ *Platform Accountability Hearing* (Testimony of Mary Anne Franks, Professor of Law) at 3–4.
- ³⁴ Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (2016); Susan Benesch, *Nobody Can See Into Facebook*, The Atlantic (Oct. 30, 2021).
- ³⁵ See, e.g., Jason Parham, *TikTok and the Evolution of Digital Blackface*, Wired (Aug. 4, 2020).
- ³⁶ See *NetChoice, v. Fla.*, 34 F.4th at 1220–21 (explaining why social media platforms are not common carriers, and holding that state restrictions on platform content moderation violate the First Amendment); *Platform Accountability Hearing* (Testimony of Mary Anne Franks, Professor of Law) at 3–4 (explaining that social media platforms are private entities with the First Amendment right to “fact-check, label, remove, ban, and make other interventions as they see fit about the content on their sites”). But see *NetChoice v. Paxton*, 49 F.4th at 455 (finding that social media content moderation was not protected speech under the First Amendment, and upholding Texas statute mandating viewpoint neutral content moderation by social media platforms). See also Frank Pasquale, *Platform Neutrality: Enhancing Freedom of Expression in Spheres of Private Power*, 17 Theoretical Inquiries L. 47 (2016) (observing the irony that while Section 230 prevents platforms from being treated as speakers, when government officials attempt to regulate how content appears, platforms have a “convenient identity crisis” and claim to be speakers that enjoy full First Amendment protections, and proposing platforms relinquish First Amendment defenses . . . in order to enjoy full Section 230 protections).
- ³⁷ David Brody and Sean Bickford, *Discriminatory Denial of Service: Applying State Public Accommodations Laws to Online Commerce*, Lawyers’ Comm. for Civil Rights Under L., 6 (Jan. 2020). As discussed above, there may be a point at which purposeful discrimination against Black patrons of an online business violates public accommodations laws or Section 1981. See 42 U.S.C. § 2000a(a); *id.* § 1981(a).
- ³⁸ See Oliver Haimson, et al., *Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas*, 5 Proc. of the ACM on Human-Computer Interaction 1, 13 (Oct. 2021) (finding that “16% of Black participants had accounts removed as compared to 10% of white participants”).
- ³⁹ Olivia Solon, *Facebook Ignored Racial Bias Research, Employees Say*, NBC News (July 23, 2020); see also Daphne Keller, *Toward a Clearer Conversation About Platform Liability*, Knight First Amendment Institute (Oct. 14, 2018).
- ⁴⁰ Brief of Amici Curiae of Scholars of Civil Rights and Justice, in Support of Respondent, *Gonzalez v. Google LLC*, 143 S. Ct. 762 (Feb. 19, 2023), No. 21-1333 [hereinafter *Civil Rights Scholars Amicus Brief (Gonzalez)*] (“The content moderation systems of many platforms lack the ability to recognize cultural nuances not rooted in a white, male, straight context, resulting in disproportionate silencing of users of color and other underserved groups.”).
- ⁴¹ Charlotte Colombo, *TikTok Has Apologized for a ‘Significant Error’ After a Video That Suggested Racial Bias in Its Algorithm Went Viral*, Insider.com (July 8, 2021).
- ⁴² See *#TruthBeTold Campaign*, African American Policy Forum (last visited July 17, 2023). In 2022, Florida enacted the Individual Freedom Act (also known as the “Stop WOKE Act”), which prohibited workplace training or school instruction that teaches about unconscious bias or that people are privileged based on race, specifically prohibited teaching *The 1619 Project*, and states that such trainings or lessons constitute discrimination.
- ⁴³ Katherine Trendacosta, *Abortion Information Is Coming Down Across Social Media. What Is Happening and What Next*, Electronic Frontier Found. (July 28, 2022).
- ⁴⁴ Daniel L. Byman, *When Hate Goes Viral*, Foreign Policy (Mar. 23, 2022). This reluctance to intervene facilitates not only content that is explicitly anti-Black, but also anti-Black propaganda strategies that use tactics such as “keyword squatting.” See Brandi Collins-Dexter and Joan Donovan, *How a Racialized Disinformation Campaign Ties Itself to the 1619 Project*, Colum. Journalism Rev. (Mar 11, 2021) (finding that right-wing media co-opted the keyword “1619,” writing typically negative coverage three times more than left-wing sources, which skewed the top search results on Google and YouTube).
- ⁴⁵ See Ryan Mac and Craig Silverman, *“Mark Changed The Rules”: How Facebook Went Easy On Alex Jones And Other Right-Wing Figures*, BuzzFeed (Feb. 22, 2021); Keach Hagey and Jeff Horwitz, *Facebook’s Internal Chat Boards Show Politics Often at Center of Decision Making*, Wall St. J. (Oct. 24, 2021).
- ⁴⁶ *A Country in Crisis: How Disinformation Online is Dividing the Nation: Hearing Before the H. Comm. on Energy and Commerce, Subcommittee on Consumer Protection and Commerce*, 116th Cong. (June 24, 2020) (Testimony of Hany Farid, Ph.D) at 4 (“[T]he vast majority of delivered content is actively promoted by content providers based on their algorithms that are designed in large part to maximize engagement and revenue. . . . These algorithms have learned that divisive, hateful, and conspiratorial content engages users and so this type of content is prioritized, leading to . . . increased anger, hate, and intolerance, both online and offline.”).
- ⁴⁷ See, Heather J. Williams, et al., *The Online Extremist Ecosystem*, RAND, 1, 7–10 (Dec. 2021).; Steven Lee Myers and Stuart A. Thompson, *Racist and Violent Ideas Jump From Web’s Fringes to Mainstream Sites*, N.Y. Times (June 1, 2022).



Acknowledgments

The Joint Center is especially appreciative of the Knight Foundation for supporting this report. We also thank Spencer Overton and Catherine Powell, who co-wrote the law review article this report summarizes.

A special thanks to Dedrick Asante-Muhammad, Jessica Fulton, Victoria Johnson, Kennedy Fortner, and Ed Paisley for their support throughout the project and for closely reviewing the report to ensure the highest level of accuracy. This report was designed by Vlad Archin.

Media Contact

press@jointcenter.org | 202.789.3500 EXT 105

© Copyright 2025

All rights reserved.

Joint Center for Political and Economic Studies

633 Pennsylvania Ave., NW

Washington, DC 20004

info@jointcenter.org | jointcenter.org

 [@JointCenter](https://twitter.com/JointCenter)

 [linkedin.com/company/jointcenter](https://www.linkedin.com/company/jointcenter)

 [facebook.com/JointCenter1970](https://www.facebook.com/JointCenter1970)

 [@JointCenter](https://www.instagram.com/JointCenter)

 [youtube.com/@jointcenterYT](https://www.youtube.com/@jointcenterYT)