No. 24-271

IN THE UNITED STATES COURT OF APPEALS
FOR THE NINTH CIRCUIT

X Corp., Successor in Interest to Twitter, Inc.,
Plaintiff-Appellant

v.

Robert Bonta, Attorney General of the State of California,
Defendant-Appellee

On Appeal from the
United States District Court, Eastern District of California
District Court Case No. 2:23-cv-01939-WBS-AC
The Honorable William B. Shubb, Judge

BRIEF OF THE LAWYERS' COMMITTEE FOR CIVIL RIGHTS UNDER LAW AS *AMICUS CURIAE* IN SUPPORT OF DEFENDANT-APPELLEE FOR AFFIRMANCE

Jon M. Greenbaum
Edward G. Caspar
Marc P. Epstein
LAWYERS' COMMITTEE FOR
CIVIL RIGHTS UNDER LAW
1500 K St. NW, Suite 900
Washington, DC 20005
(202) 662-8600
jgreenbaum@lawyerscommittee.org
ecaspar@lawyerscommittee.org
mepstein@lawyerscommittee.org

Michelle Quist
Lauren D. Wigginton
HOLLAND & HART LLP
222 South Main Street, Suite 2200
Salt Lake City, Utah 84101
(801) 799-5800
9555 Hillwood Drive, 2nd Floor
Las Vegas, Nevada 89134
(702) 669-4600
mlquist@hollandhart.com
ldwigginton@hollandhart.com

March 20, 2024

Counsel for Amicus Curiae

RULE 26.1 CORPORATE DISCLOSURE STATEMENT

Amicus is a nonprofit organization. It has no parent corporations, and no publicly held corporation owns any portion of it.

TABLE OF CONTENTS

Page

RULE 26.1 CORPORATE DISCLOSURE STATEMENT ii

IDENTITY AND INTEREST OF *AMICUS CURIAE* 1

INTRODUCTION 2

ARGUMENT 3

I. AB 587's Disclosure Requirements Would Improve the Online Experiences of Black People and Other People of Color with Minimal Burden on Social Media Companies. 4

II. AB 587 Empowers Black People and Other People of Color to Avoid Platforms that Fail to Curb Hate and Harassment and Election Disinformation. 10

A. Online hate and harassment chill the free speech of Black people and other people of color. 11

B. Online disinformation suppresses the votes of Black people and other people of color. 17

III. AB 587 Would Shed Light on the Effects of Social Media Companies' Content Moderation Practices on Content Creators of Color. 23

CONCLUSION 25

TABLE OF AUTHORITIES

Page(s)

CASES

Gonzalez v. Google LLC,
598 U.S. 617 (2023) 1

Ill. State Bd. of Elections v. Socialist Workers Party,
440 U.S. 173 (1979) 23

Moody v. NetChoice, LLC,
No. 22-277 (2023) 1

National Coalition on Black Civic Participation v. Wohl,
498 F. Supp. 3d 457 (S.D.N.Y. 2020) 17

NetChoice, LLC v. Bonta,
No. 23-2969 (9th Cir. 2023) 1

NetChoice, LLC v. Paxton,
No. 22-555 (2023) 1

Packingham v. North Carolina,
582 U.S. 98 (2017) 14

Vargas v. Facebook, Inc.,
No. 21-16499 (9th Cir. 2023) 1

Yick Wo v. Hopkins,
118 U.S. 356 (1886) 23

RULES AND STATUTES

Cal. Bus. & Prof. Code § 22677 ..... 6, 23, 25

Cal. Bus. & Prof. Code § 22680 6

California Assembly Bill 587 ..... passim

Fed. R. App. P. 29 1

OTHER AUTHORITIES

Ali Swenson & Will Weissert, *New Hampshire Investigating Fake Biden Robocall Meant to Discourage Voters Ahead of Primary*, AP News (Jan. 22, 2024) 22

Alyan Layug et. al, T*he Impacts of Social Media Use and Online Racial Discrimination on Asian American Mental Health: Cross-sectional Survey in the United States During COVID-19*, FRONT PUBLIC HEALTH, 20 (Jan. 12, 2021)..... 13

Amanda Lenhart, *Online Harassment, Digital Abuse, and Cyberstalking*, Data & Soc'y Rsch. Inst. (Nov. 21, 2016) 15

Amnesty Int'l, *Troll Patrol Findings* (Mar. 5, 2024) 12

Anti-Defamation League & Ctr. for Tech. & Soc'y, *Online Hate and Harassment: The American Experience 2022* (2022) 11

Anti-Defamation League & Ctr. for Tech. & Soc'y, *Online Hate and Harassment: The American Experience 2023* (2023) 11

Ashley Nerbovig, *Michigan's Immigrant Communities Hit With Misinformation on Closed Platforms*, Detroit Free Press (Nov. 27, 2020) 19

Ashley Nerbovig, *WhatsApp, KaKaoTalk Make it Easy to Spread Misinformation, Especially Among Immigrant Communities*, Detroit Free Press (Nov. 27, 2020) 20

Blake Peterson, *ICE, Dispelling Rumors, Says it Won't Patrol Polling Places*, ProPublica (Nov. 2, 2018)..... 19

Brendesha Tynes et al., *Trajectories of Online Racial Discrimination and Psychological Functioning among African American and Latino Adolescents*, 91(5) Child Dev. 1577, 1578 (2020) 12

Cmty. Rel. Serv., *2022 FBI Hate Crime Statistics*, Dep't. of Just. (Oct. 30, 2023) 12

Ctr. for Fam. Rsch., *Skin-deep Resilience*, Univ. of Ga. (Mar. 2021) 13

Daniel I. Weiner & Lawrence Norden, *Regulating AI Deepfakes and Synthetic Media in the Political Arena*, Brennan Ctr. for Just. (Dec. 5, 2023) 21

Dhanaraj Thakur & DeVan Hankerson Madrigal, *An Unrepresentative Democracy: How Disinformation and Online Abuse Hinder Women of Color Political Candidates in the United States*, Ctr. for Democracy & Tech. 8 (2023) 15-16

Elizabeth Dwoskin et al., *Facebook's Race-Blind Practices Around Hate Speech Came At The Expense Of Black Users, New Documents Show*, Wash. Post (Nov. 21, 2021) 24

Emily A. Vogels, *The State of Online Harassment*, PEW RSCH. CTR. (Jan. 13, 2021) 11, 12

George Veletsianos et al., *Women Scholars' Experiences with Online Harassment and Abuse: Self-Protection, Resistance, Acceptance, and Self-Blame* (2018) 14-15

Ian Vandewalker, *Digital Disinformation and Voter Suppression*, Brennan Ctr. for Just. (Sept. 2, 2020) 20

Jason Guerrasio, *'Star Wars' actress Kelly Marie Tran Left Social Media After Racist and Sexist Trolls Drove Her to Therapy*, Insider (Mar. 3, 2021)..... 15

Jessica Guynn, *Facebook While Black: Users Call It Getting 'Zucked,' Say Talking About Racism Is Censored As Hate Speech*, USA Today (Apr. 24, 2019) 24

Kat Tenbarge, *Fake News YouTube Creators Target Black Celebrities With AI-Generated Misinformation*, NBC News (Jan. 30, 2024) 22

Louis D. Brandeis, *What Publicity Can Do*, Harper's Weekly (Dec. 20, 1913) 5

Lucina Fisher & Brian McBride, *'Ghostbusters' Star Leslie Jones Quits Twitter After Online Harassment*, ABC News (July 20, 2016) 15

Maeve Duggan, *Online Harassment 2017*, Pew Rsch. Ctr. (July 11, 2017) 14

Marianna Spring, *Trump Supporters Target Black Voters with Faked AI Images*, BBC News (Mar. 4, 2024) 22

Megan McCluskey, *These TikTok Creators Say They're Still Being Suppressed for Posting Black Lives Matter Content*, Time (July 22, 2020) 24

Oliver L. 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. Ass'n for Computing Mach. on Hum.-Comput. Interaction 466 (2021)..... 24-25

Press Release, U.S. Att'ys Off. E.D.N.Y., *Social Media Influencer Douglass Mackey Convicted of Election Interference in 2016 Presidential Race* (Mar. 31, 2023)..... 18

Rebekah Herrick et al., *Gender and Race Differences in Mayors' Experiences of Violence*, Ctr. for Am. Women & Politics (2022) 16

Russian Active Measures Campaigns and Interference in the 2016 U.S. Election, Vol. 2: Russia's Use of Social Media with Additional Views, S. Rep. No. 116-290, at 6 (2020)..... 18

Sherman James, *John Henryism and the health of African Americans*, Culture, Med. & Psychiatry 163, 178 (1994)..... 13

Sylvia Albert et al., *As a Matter of Fact: The Harms Caused by Election Disinformation*, Common Cause Educ. Fund 12-29 (Oct. 2021) 20

The US is Bracing for Complex, Fast-Moving Threats to Elections This Year, FBI Director Warns, AP News (Feb. 29, 2024) 20

Tiffany Hsu, *Misinformation Swirls in Non-English Languages Ahead of Midterms*, New York Times (Oct. 12, 2022) 19

Todd C. Helmus, *Artificial Intelligence, Deepfakes, and Disinformation: A Primer*, RAND Corp. (July 2022)..... 21

Zach Montellaro, *America's Election Chiefs Are Worried AI is Coming For Them*, Politico (Mar. 11, 2024)..... 20-21

IDENTITY AND INTEREST OF *AMICUS CURIAE*

Formed in 1963, the Lawyers' Committee for Civil Rights Under Law is a nonpartisan, nonprofit organization using legal advocacy to pursue racial justice. The Lawyers' Committee fights inside and outside the courts to ensure that Black people and other people of color have the voice, opportunity, and power to make the promises of American democracy real.

The Lawyers' Committee's Digital Justice Initiative works at the intersection of racial justice, technology, data, and privacy. The Initiative combats online hate and discrimination targeting Black Americans and other communities of color. It regularly participates in cases involving online discrimination, election disinformation, and privacy. *See, e.g.*, *NetChoice, LLC v. Bonta*, No. 23-2969 (9th Cir. 2023); *NetChoice, LLC v. Paxton*, No. 22-555 (2023), *Moody v. NetChoice*, LLC, No. 22-277 (2023); *Gonzalez v. Google LLC*, 598 U.S. 617 (2023); *Vargas v. Facebook, Inc.*, No. 21-16499 (9th Cir. 2023).[1]

[1] Lawyers' Committee files this brief with the consent of all parties. No counsel for a party authored this brief in whole or in part. No person or entity, other than amicus curiæ, its members, or its counsel, made any monetary contribution to the preparation or submission of this brief. *See* Fed. R. App. P. 29(a)(4)(E).

INTRODUCTION

Race-based hate and harassment, and election disinformation targeted at Black people and other people of color, are pervasive on social media. They silence the voices of Black people online, cause people to disengage from online discourse, and affect the ability of Black people to engage in democratic processes. California Assembly Bill 587 ("AB 587") is a simple transparency measure with the potential to shed light on the prevalence of these issues and empower Black people and other people of color to make informed choices about which social media platforms to use.

AB 587's requirements are minimal. Under it, social media companies must disclose their terms of service and whether they prohibit certain categories of content (like "Hate speech or racism," "Disinformation or misinformation," or "Harassment"). And those companies must provide periodic reports about the amount of content that they act on pursuant to those policies. These are things that, to some extent, most social media companies *already* do. Contrary to X's assertions, *see, e.g.*, Appellant Br. at 1, AB 587 does not require social media companies to prohibit any specific type of conduct or to engage in any type of content moderation efforts.

Yet, AB 587's simple disclosure provisions may have powerful effects on the online experiences of Black people and other people of color. Access to information about social media platforms' policies and practices empowers Black people and other people of color to make informed choices about which social media platforms to use. Such access also allows civil society organizations to identify harmful online practices, educate the public about them, and advocate against them. AB 587 would also shed light on discriminatory practices, including whether automated content moderation decisions have an unfair or discriminatory impact on content created by Black content creators and other creators of color.

AB 587 is a modest transparency measure that allows Black people and other people of color—who are often deliberate targets of online disinformation and harassment—the freedom of informed choice. Its disclosure provisions protect their free speech and voting rights and their right to be free from discrimination. This Court should affirm.

ARGUMENT

Black people and other people of color regularly experience online hate and harassment and are the targets of election disinformation. Hate, harassment, and election disinformation profoundly hinder Black

people and other people of color's participation in online discourse and democratic institutions.

Accordingly, the Lawyers' Committee offers three arguments in support of affirmance. *First*, affirmance would sustain the benefits of a minimally burdensome transparency law and enable future transparency efforts. *Second*, by requiring social media companies to publish whether they prohibit hate, harassment, misinformation, and disinformation, AB 587 would empower Black people and other people of color to make informed decisions about which social media platforms to use, safeguard their ability to engage in online discourse, and uphold their right to vote. *Third*, AB 587 would assist the public in evaluating the effect of automated content moderation systems on content created and shared by Black content creators and other content creators of color.

I. AB 587's Disclosure Requirements Would Improve the Online Experiences of Black People and Other People of Color with Minimal Burden on Social Media Companies.

AB 587 is a simple transparency provision. It does not require social media companies to engage in any content moderation efforts. It does not require social media companies to prohibit particular categories of content or change their terms of service. It does not require social media

companies to alter how they run their platforms in any way. Rather, it requires companies to publish existing policies and reports on enforcement of those policies—disclosures that many social media companies *already make*.

Nonetheless, mandating these disclosures from all social media companies may have a significant impact on the online experiences of Black people and other people of color. "Publicity is justly commended as a remedy for social and industrial diseases. Sunlight is said to be the best of disinfectants; electric light the most efficient policeman." Louis D. Brandeis, *What Publicity Can Do,* HARPER'S WEEKLY, at 10 (Dec. 20, 1913). [2] AB 587 would be the sunlight exposing online hate and disinformation targeted at Black people and other people of color on social media platforms. Reversing the district court's Order and enjoining AB 587 could threaten future efforts to hold social media platforms accountable and make it impossible to establish additional transparency measures nationwide.

[2] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7874150/ (last accessed March 20, 2024).

AB 587 imposes minimal burdens on social media companies. The law requires social media companies to publish their terms of service and a statement of whether their terms of service defines several categories of content, including, (A) Hate speech or racism, (B) Extremism or radicalization, (C) Disinformation or misinformation, (D) Harassment, and (E) Foreign political interference, and (F) Controlled substance distribution. *See* Cal. Bus. & Prof. Code § 22677(a)(3). If their terms of service define those categories of conduct, then the social media company must include the current definition. *Id.* AB 587 also requires social media companies to publish data on the application of their content moderation policies, including the amount of flagged, "actioned," removed, demonetized, deprioritized, or banned content and how often that content was viewed and shared. *Id.* § 22677(a)(5).

The social media companies covered under AB 587 are more than capable of complying with AB 587. The law only covers social media companies with annual gross revenues of $100,000,000 or more, *id.* § 22680, and most of these companies already make disclosures that could be adapted to satisfy AB 587. Almost every major social media company already publishes terms of service or community guidelines,

including definitions of prohibited categories of content. *See, e.g.*, Meta, *Hate Speech*;[3] TikTok, *Safety and Civility* (Mar. 2023).[4] Indeed, many of these companies regularly publish enforcement reports with similar data that California hopes to require of them. For example, Meta's "Community Standards Enforcement Report" details the amount of hate speech content that Meta acted on. *See* Meta, *Community Standards Enforcement Report: Hate Speech*. [5] Likewise, TikTok's Community Guidelines Report breaks down the policy distributions of removals, including the amount of content removed because it fell afoul of their "Safety & Civility" standards, which prohibit Hate Speech and Hateful Behaviors. *See* TikTok, *Community Guidelines Enforcement Report*.[6]

AB 587's disclosure provisions would help ensure that all people are treated equally on social media platforms, create accountability, and highlight racial inequity. Although the disclosure provisions do not

---

[3] https://transparency.fb.com/policies/community-standards/hate-speech/ (last accessed Mar. 17, 2024).

[4] https://www.tiktok.com/community-guidelines/en/safety-civility/ (last accessed Mar. 17, 2024).

[5] https://transparency.fb.com/reports/community-standards-enforcement/hate-speech/facebook/ (last accessed March 11, 2024).

[6] https://www.tiktok.com/transparency/en-us/community-guidelines-enforcement-2023-3/ (last accessed March 11, 2024).

require companies to alter how they run their platforms, the provisions do compel companies to consider the impact of their actions and whether corporate incentives are aligned with the public welfare. When companies are required to examine and disclose harm they cause, they are more apt to address those harms voluntarily or proactively or to take responsibility after the fact. The sanction for failure is their customers' opprobrium.

AB 587 also would allow consumers to "make informed decisions about where they consume and disseminate news and information." Order, at 5-6. Consumers may not want to use a platform in which hate, harassment, and disinformation are prevalent, and AB 587 would allow them to determine whether a platform prohibits that content and how well they enforce their policies. In this way, AB 587 is no different than laws requiring nutrition labels, which allow consumers to choose whether they want to eat food with fewer calories, avoid food with ingredients to which they are allergic, or choose organic or locally sourced products.

AB 587 also will facilitate the collection and publication of data that are critical for civil society and advocacy organizations. The data will make it easier for watchdogs to identify online hate, harassment, and

disinformation, educate the public about the pervasiveness of those problems, and advocate against them. Systemic discrimination, in particular, is difficult to discern without comprehensive, systematic data. Advocacy groups and stakeholders can use such data to push for necessary changes and for the government to create more robust and fair regulations. The data also can inform policy decisions that foster an online environment that is safe and inclusive for all people.

The benefits of transparency laws do not run one way. Such laws foster trust between social media companies and their users. In turn, users are more likely to trust and engage with platforms they perceive to be fair and non-discriminatory. For instance, financial institutions require disclosure of fees and interest rates. Greater transparency helps prevent fraudulent activities and promotes trust between the banks and their customers.

Black people and other people of color deserve the benefits and protections provided by robust transparency regulations. Reversal of the district court's Order would imperil those protections as well as future efforts to establish transparency measures.

II. AB 587 Empowers Black People and Other People of Color to Avoid Platforms that Fail to Curb Hate and Harassment and Election Disinformation.

Online hate and harassment have a profound chilling effect on the free expression of groups that have experienced longstanding discrimination. In fact, as discussed below, *that is often the intent* of such speech. Online disinformation and voter intimidation targeted at Black people and other people of color similarly impair the ability of communities to support and advocate for the representatives of their choice. AB 587 would require that social media companies publish whether they prohibit these types of harmful content, allowing Black people and other people of color to make informed choices about which social media platforms they use.

AB 587 would empower Black people and other people of color to avoid platforms on which hate, harassment, and disinformation are permitted or pervasive. In this way, AB 587 would have an immediate positive effect on the online experiences of Black people and other people of color.

A. Online hate and harassment chill the free speech of Black people and other people of color.

There is an epidemic of online hate and harassment. Roughly four in ten Americans have been targeted by online harassment. *See* Anti-Defamation League & Ctr. for Tech. & Soc'y, *Online Hate and Harassment: The American Experience 2022*, 10 (2022 Report);[7] Anti-Defamation League & Ctr. for Tech. & Soc'y, *Online Hate and Harassment: The American Experience 2023* (2023 Report);[8] Emily A. Vogels, *The State of Online Harassment*, PEW RSCH. CTR. 4 (Jan. 13, 2021).[9] More than one in four of all adults have experienced severe online harassment in the form of physical threats, sustained harassment, stalking, sexual harassment, doxing, or swatting. 2022 Report, *supra*, at 10; Vogels, *supra*, at 4.

No group is immune, but Black people and other people of color experience a different quality and quantity of hate. *See, e.g.*, Vogels,

---

[7] https://www.adl.org/sites/default/files/pdfs/2022-09/Online-Hate-and-Harassment-Survey-2022.pdf (last accessed March 20, 2024).

[8] https://www.adl.org/sites/default/files/pdfs/2023-06/Online-Hate-and-Harassmen-2023_0.pdf (last accessed March 20, 2024).

[9] https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/.

supra, at 21 (more than half of Black people who have experienced harassment online say they were harassed due to their race or ethnicity, compared to 17% of white people); *see also, e.g.*, Amnesty Int'l, *Troll Patrol Findings* (analyzing millions of tweets and finding that Black women were 84% more likely to be mentioned in abusive tweets than white women).[10] Modern online hate reprises historical tropes, deriding Black people and other people of color "as unintelligent, as criminals, and as lazy" via racist traditions like disgust-based metaphors that strip people of color of humanity, and blackface. Brendesha Tynes et al., *Trajectories of Online Racial Discrimination and Psychological Functioning among African American and Latino Adolescents*, 91(5) Child Dev. 1577, 1578 (2020). Once targeted, Black people and other people of color are harassed even more viciously than their white counterparts. Vogels, *supra*, at 24; *see also* Cmty. Rel. Serv., *2022 FBI Hate Crime Statistics*, Dep't. of Just. (Oct. 30, 2023) (race was the most common motivator of hate crimes in 2022, and anti-Black incidents "were

[10] https://decoders.amnesty.org/projects/troll-patrol/findings (last accessed Mar. 5, 2024).

more than three times higher than the next highest racial or ethnic category").[11]

The physical, psychological, and emotional effects of experiencing online hate and harassment are myriad. Routine exposure to abhorrent beliefs about and abuse of those who share their racial community takes a physical toll on Black people and other people of color, resulting in higher rates of diseases such as hypertension, advanced cellular aging, diabetes, and respiratory infection. *See*, *e.g.*, Sherman James, *John Henryism and the health of African Americans*, Culture, Med. & Psychiatry 163, 178 (1994); *see also Ctr. for Fam. Rsch., Skin-deep Resilience*, Univ. of Ga. (Mar. 2021). Experiencing and witnessing discrimination also has negative psychological effects, including stress, depression, and anxiety. *See*, *e.g.*, Alyan Layug, Samiksha Krishnamurthy, Rachel McKenzie, and Bo Feng, *The Impacts of Social Media Use and Online Racial Discrimination on Asian American Mental Health: Cross-sectional Survey in the United States During COVID-19*, Front Public Health, 20 (Jan. 12, 2021)[12] (finding that "individual and

[11]https://www.justice.gov/crs/highlights/2022-hate-crime-statistics.

[12] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7874150/.

vicarious forms of online discrimination experienced during the COVID-19 pandemic were . . . positively associated with [secondary traumatic stress], depression, and anxiety").

The ultimate effect of online hate and harassment is to chill the speech of Black people and other people of color. For many, social media platforms "are the principal sources for knowing current events, checking ads for employment, speaking and listening in the modern public square, and otherwise exploring the vast realms of human thought and knowledge." *Packingham v. North Carolina*, 582 U.S. 98, 107 (2017). But online hate and harassment lead people to preemptively self-censor and withdraw from "the modern public square" for fear of being targeted. Maeve Duggan, *Online Harassment 2017*, Pew Rsch. Ctr. (July 11, 2017).[13] When confronted with online harassment, women in particular are more likely to self-censor or withdraw from online platforms altogether. *See, e.g.*, George Veletsianos et al., *Women Scholars' Experiences with Online Harassment and Abuse: Self-Protection, Resistance, Acceptance, and Self-Blame*, at *14 (2018) (harassment of

[13] https://www.pewresearch.org/internet/2017/07/11/online-harassment-2017-methodology/.

women scholars led to their avoiding certain social media platforms and "turning to silence");[14] Amanda Lenhart, *Online Harassment, Digital Abuse, and Cyberstalking,* Data & Soc'y Rsch. Inst. (Nov. 21, 2016) (41% of women ages 15-29 self-censor).[15] Numerous prominent women of color have withdrawn from online discourse because of online harassment. *See, e.g.*, Jason Guerrasio, *'Star Wars' actress Kelly Marie Tran Left Social Media After Racist and Sexist Trolls Drove Her to Therapy*, Insider (Mar. 3, 2021);[16] Lucina Fisher & Brian McBride, *'Ghostbusters' Star Leslie Jones Quits Twitter After Online Harassment*, ABC News (July 20, 2016).[17]

By design, online hate also discourages Black people and other people of color from running for public office. *See* Dhanaraj Thakur & DeVan Hankerson Madrigal, *An Unrepresentative Democracy: How*

---

[14] https://veletsianos.com/wp-content/uploads/2011/07/harassment_coping_postPrint.pdf (last accessed Mar. 20, 2024).

[15] https://www.datasociety.net/pubs/oh/Online_Harassment_2016.pdf.

[16] https://www.businessinsider.in/entertainment/news/star-wars-actress-kelly-marie-tran-left-social-media-after-racist-and-sexist-trolls-drove-her-to-therapy/articleshow/81318930.cms.

[17] https://abcnews.go.com/Entertainment/ghostbusters-starleslie-jones-quits-twitter-online-harassment/story?id=40698459.

Disinformation and Online Abuse Hinder Women of Color Political Candidates in the United States, Ctr. for Democracy & Tech. 8 (2023).[18] In a study of the 2020 congressional election, women of color seeking office were five times more likely than other candidates to experience online abuse related to their gender and racial identities. *Id.*; *see also* Rebekah Herrick et al., *Gender and Race Differences in Mayors' Experiences of Violence*, Ctr. for Am. Women & Politics (2022) (study of U.S. mayors showed women of color experienced more threats than other groups). Even if it does not actually stop women of color from succeeding in politics, "hate and risk should not be the inherent cost of any woman participating in politics, regardless of her race, creed, sexual or gender identity, or any other defining quality of who she is." Herrick et al., *supra* (quoting U.S. Representative Rashida Tlaib).

In short, Black people and other people of color experience significant amounts of online hate and harassment on social media platforms. These experiences have negative physical, psychological, and

[18] https://cdt.org/insights/an-unrepresentative-democracy-how-disinformation-and-online-abuse-hinder-women-of-color-political-candidates-in-the-united-states/ (last accessed March 20, 2024).

social consequences, and ultimately chill the free speech of affected communities.

B. Online disinformation suppresses the votes of Black people and other people of color.

Misinformation (false information) likewise is rampant on social media platforms. And, as discussed below, election disinformation (intentionally false information) often is targeted specifically at Black people and other people of color, with the intent of suppressing their vote. Whereas the means used to intimidate Black voters were once "guns, torches, burning crosses, and other dire methods perpetuated under the cover of white hoods," 150 years later individuals "carry out electoral terror using telephones, computers, and modern technology adapted to serve the same deleterious ends." *National Coalition on Black Civic Participation v. Wohl*, 498 F. Supp. 3d 457, 464 (S.D.N.Y. 2020). "Because of the vastly greater population [those individuals] can reach instantly with false and dreadful information, contemporary means of voter intimidation may be more detrimental to free elections than the approaches taken for that purpose in past eras." *Id.* These means of voter intimidation merit "swift and effective [legislative] relief." *Id.*

That wrongdoers will use social media to suppress the vote of Black people and other people of color is more than a theoretical possibility. In 2016, a prominent social media personality fraudulently promoted images and tweets encouraging voters to cast their ballots online or via text message. This included a fake advertisement, purportedly on behalf of the Hillary Clinton campaign, showing a Black woman holding an "African Americans for Hillary" sign and encouraging voters to "Avoid the Line" and "Vote from Home." Press Release, U.S. Att'ys Off. E.D.N.Y., *Social Media Influencer Douglass Mackey Convicted of Election Interference in 2016 Presidential Race* (Mar. 31, 2023).[19] Likewise, a Senate investigation confirmed that the use of disinformation on social media to suppress Black votes was a pillar of Russia's strategy to interfere in the 2016 presidential election. *See Russian Active Measures Campaigns and Interference in the 2016 U.S. Election, Vol. 2: Russia's Use of Social Media with Additional Views*, S. Rep. No. 116-290, at 6 (2020) (Senate Report) ("No single group of Americans was targeted by [the Internet Research Agency (IRA)] information operatives more than

[19] https://www.justice.gov/usao-edny/pr/social-media-influencer-douglass-mackey-convicted-election-interference-2016.

African-Americans."). The IRA operated as Russia's private troll factory, using Instagram accounts with names like "@Blackstagram_," which gained over 100,000 followers, and X and Facebook accounts targeted at Black voters that disseminated messages like "Don't Vote at All," "Why Would We Be Voting," or "Our Votes Don't Matter." *Id.* at 92, 105, 119.

Misinformation on social media in non-English languages, particularly Spanish, was also rampant in the 2020 and 2022 elections. Tiffany Hsu, *Misinformation Swirls in Non-English Languages Ahead of Midterms*, New York Times (Oct. 12, 2022). [20] Other recent disinformation campaigns include spreading falsehoods that "told [voters] it was dangerous to vote . . . [because] the government would record who people voted for and they might face consequences," Ashley Nerbovig, *Michigan's Immigrant Communities Hit With Misinformation on Closed Platforms*, Detroit Free Press (Nov. 27, 2020); [21] that immigration officers would monitor voting, Blake Peterson, *ICE, Dispelling Rumors, Says It Won't Patrol Polling Places*, ProPublica (Nov.

[20] https://www.nytimes.com/2022/10/12/business/media/midterms-foreign-language-misinformation.html.

[21]https://www.freep.com/story/news/politics/elections/2020/11/27/michigan-election-misinformation-immigrant-whatsapp/6393693002/.

2, 2018);[22] that scantron ballots had "swapped sensors," so that "[i]f you are intending on voting for Joe Biden, you must bubble in Trump and vice versa," Ashley Nerbovig, *WhatsApp, KaKaoTalk make it easy to spread misinformation, especially among immigrant communities*, Detroit Free Press (Nov. 27, 2020);[23] and numerous others. *See generally* Sylvia Albert et al., *As a Matter of Fact: The Harms Caused by Election Disinformation*, Common Cause Educ. Fund 12-29 (Oct. 2021); Ian Vandewalker, *Digital Disinformation and Voter Suppression*, Brennan Ctr. for Just. (Sept. 2, 2020).

United States' officials are bracing for an acceleration of those threats in the 2024 election, fueled by generative AI and social media. *See The US is Bracing for Complex, Fast-Moving Threats to Elections This Year, FBI Director Warns*, AP News (Feb. 29, 2024).[24] "A false call from a secretary of state telling poll workers they aren't needed on

[22] https://www.propublica.org/article/ice-dispelling-rumors-says-it-wont-patrol-polling-places.

[23]https://www.usatoday.com/story/news/politics/elections/2020/11/27/immigrant-communities-2020-election-misinformation-messaging-apps/6436984002/.

[24] https://apnews.com/article/fbi-election-interference-wray-2024-campaign-ai-a0c4a95c818839b18f919c6d648c4dcf.

Election Day. A fake video of a state election director shredding ballots before they're counted. An email sent to a county election official trying to phish logins to its voter database." Zach Montellaro, *America's Election Chiefs are Worried AI is Coming for Them*, Politico (Mar. 11, 2024).[25] Widely available generative AI products have made all these possible. Those products allow individuals quickly and cheaply to create and spread "deepfake" images, videos, and audio of events that never occurred and recordings of people saying things they never said. *See generally* Todd C. Helmus, *Artificial Intelligence, Deepfakes, and Disinformation: A Primer*, RAND Corp. (July 2022).[26]

Already there are examples of malicious actors using generative AI to target Black people with disinformation and spread election disinformation. *See, e.g.*, Daniel I. Weiner & Lawrence Norden, *Regulating AI Deepfakes and Synthetic Media in the Political Arena*, Brennan Ctr. for Just. (Dec. 5, 2023) (describing how deepfake audio recordings spread before Slovakia's October 2023 election). In March

[25] https://www.politico.com/news/2024/03/11/secretary-state-ai-election-misinformation-00146137.

[26] https://www.rand.org/pubs/perspectives/PEA1043-1.html (last accessed Mar. 20, 2024).

2024, BBC discovered that Donald Trump supporters had created and shared AI-generated deepfake images portraying Trump with Black people. *See* Marianna Spring, *Trump Supporters Target Black Voters With Faked AI Images*, BBC News (Mar. 4, 2024).[27] Fake news accounts have used AI-generated fake content to target Black celebrities with salacious disinformation. *See* Kat Tenbarge, *Fake News YouTube Creators Target Black Celebrities With AI-Generated Misinformation*, NBC News (Jan. 30, 2024).[28] And one individual created a fake robocall that used Joe Biden's voice to tell voters not to vote in the 2024 New Hampshire primary, but instead to "save your vote for the November election." *See* Ali Swenson & Will Weissert, *New Hampshire Investigating Fake Biden Robocall Meant to Discourage Voters Ahead of Primary*, AP News (Jan. 22, 2024).[29]

---

[27] https://www-bbc-com.cdn.ampproject.org/c/s/www.bbc.com/news/world-us-canada-68440150.amp.

[28] https://www.nbcnews.com/tech/misinformation/ai-deepfake-fake-news-youtube-black-celebrities-rcna133368.

[29] https://apnews.com/article/new-hampshire-primary-biden-ai-deepfake robocallf3469ceb6dd613079092287994663db5.

Online misinformation or disinformation are threats to the voting rights of Black people and other people of color—a right that is "of the most fundamental significance under our constitutional structure." *Ill. State Bd. of Elections v. Socialist Workers Party*, 440 U.S. 173, 184 (1979). Requiring social media companies to disclose whether they prohibit disinformation or misinformation and statistics of content moderation decisions taken pursuant to those policies is an important step toward safeguarding that "fundamental political right." *Yick Wo v. Hopkins*, 118 U.S. 356, 370 (1886).

III. AB 587 Would Shed Light on the Effects of Social Media Companies' Content Moderation Practices on Content Creators of Color.

Transparency laws like AB 587 help shed light on discrimination resulting from the social media companies' automated content moderation systems. AB 587 requires that social media companies explain how their automated content moderation systems enforce their terms of service, and when these systems involve human review. *See* Cal. Bus. & Prof. Code § 22677(4)(B). These explanations, along with the statistics underlying content moderation practices, *see id.* § 22677(5), will

assist the public in ensuring that automated content moderation systems do not discriminate against Black people and other people of color.

It is well-documented that social media users who are Black or transgender are more likely to have their accounts disciplined when they discuss race or gender issues, respectively, even if they are not violating platform rules. *See*, *e.g.*, Elizabeth Dwoskin et al., *Facebook's Race-Blind Practices Around Hate Speech Came At The Expense Of Black Users, New Documents Show*, Wash. Post (Nov. 21, 2021);[30] Megan McCluskey, *These TikTok Creators Say They're Still Being Suppressed for Posting Black Lives Matter Content*, Time (July 22, 2020);[31] Jessica Guynn, *Facebook While Black: Users Call It Getting 'Zucked,' Say Talking About Racism Is Censored As Hate Speech*, USA Today (Apr. 24, 2019).[32] For example, research has shown that Facebook's algorithm polices comments by Black people and other people of color more aggressively. *See* Oliver L. Haimson et al., *Disproportionate Removals and Differing Content*

---

https://www.washingtonpost.com/technology/2021/11/21/facebook-algorithm-biased-race/.

[31] https://time.com/5863350/tiktok-black-creators/.

[32] https://www.usatoday.com/story/news/2019/04/24/facebook-while-black-zucked-users-say-they-get-blocked-racism-discussion/2859593002/.

Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas, 5 Proc. Ass'n for Computing Mach. on Hum.-Comput. Interaction 466 (2021). But "[t]his is not just a dominant platform problem; this is a digital marketplace issue that has little to do with the size of the platform but rather with the broader digital platform economic structure." *Id.*

Requiring social media companies to explain how automated content moderation systems enforce the companies' terms of service, *see* Cal. Bus. & Prof. Code § 22677(4)(B), and to publish statistics about their content moderation practices, will help the public better understand the effects of social media companies' automated content moderation systems on Black content creators and other creators of color.

CONCLUSION

For the foregoing reasons, this Court should affirm the district court's Order.

DATED this 20th day of March 2024.

HOLLAND & HART LLP
Attorneys for Lawyers' Committee

By: */s/ Michelle Quist*

Jon M. Greenbaum

Edward G. Caspar
Marc P. Epstein
LAWYERS' COMMITTEE FOR
CIVIL RIGHTS UNDER LAW
1500 K St. NW, Suite 900
Washington, DC 20005
(202) 662-8600
jgreenbaum@lawyerscommittee.org
ecaspar@lawyerscommittee.org
mepstein@lawyerscommittee.org

Michelle Quist
Lauren D. Wigginton
HOLLAND & HART LLP
222 South Main Street, Suite 2200
Salt Lake City, Utah 84101
(801) 799-5800
9555 Hillwood Drive, 2nd Floor
Las Vegas, Nevada 89134
(702) 669-4600
mlquist@hollandhart.com
ldwigginton@hollandhart.com

Counsel for *Amicus Curiae*

UNITED STATES COURT OF APPEALS
FOR THE NINTH CIRCUIT
FORM 8. CERTIFICATE OF COMPLIANCE FOR BRIEFS

Instructions for this form:
http://www.ca9.uscourts.gov/forms/form08instructions.pdf

9th Cir. Case Number: No. 24-271

I am the attorney.

This brief contains 4,253 words, including 0 words manually counted in any visual images, and excluding the items exempted by FRAP 32(f). The brief's type size and typeface comply with FRAP 32(a)(5) and (6).

I certify that this brief is an **amicus** brief and complies with the word limit of FRAP 29(a)(5), Cir. R. 29-2(c)(2), or Cir. R. 29-2(c)(3).

Signature */s/ Michelle Quist* **Date** March 20, 2024

31695216_v1