

CASE NO. 23-356

**IN THE UNITED STATES COURT OF APPEALS
FOR THE SECOND CIRCUIT**

EUGENE VOLOKH, RUMBLE CANADA, INC.,
AND LOCALS TECHNOLOGY, INC.,
Plaintiffs-Appellees,

v.

LETITIA JAMES IN HER OFFICIAL CAPACITY AS
ATTORNEY GENERAL OF NEW YORK,
Defendant-Appellant.

On Appeal from the United States District Court
for the Southern District of New York, No. 23-356
(Hon. Andrew L. Carter, Jr.)

**BRIEF OF NETCHOICE AND CHAMBER OF PROGRESS
AS *AMICI CURIAE* IN SUPPORT
OF PLAINTIFFS-APPELLEES AND IN SUPPORT OF AFFIRMANCE OF
THE JUDGMENT**

Jess Miers
Legal Advocacy Counsel
Chamber of Progress
1390 Chain Bridge Road,
#A108
McLean, VA 22101

Nicole Saad Bembridge
Counsel of Record
Christopher J. Marchese
Carl Szabo
Paul D. Taske
NetChoice
1401 K St NW, Suite 502
Washington, D.C. 20005
(443) 254-6330

September 26, 2023

RULE 26.1 CORPORATE DISCLOSURE STATEMENT

Amicus NetChoice is a 501(c)(6) District of Columbia organization. It has no parent corporation and no publicly held corporation owns 10% or more of its stock.

Amicus Chamber of Progress is a 501(c)(6) Virginia organization. It has no parent corporation and no publicly held corporation owns 10% or more of its stock.

TABLE OF CONTENTS

RULE 26.1 CORPORATE DISCLOSURE STATEMENTi
TABLE OF AUTHORITIES iii
INTEREST OF *AMICI CURIAE* 1
SUMMARY OF ARGUMENT3
ARGUMENT4
 I. NEW YORK CANNOT REGULATE SPEECH BY SIMPLY
 CALLING IT 'CONDUCT'4
 A. Defining, Publishing, and Enforcing Community Guidelines is
 Speech.....6
 B. Sharing and Receiving Content on Social Media is Speech.....11
 II. THE ONLINE HATE SPEECH LAW WILL NOT AID EFFORTS
 TO IDENTIFY AND REMOVE HARMFUL CONTENT ONLINE.....13
 III. TO HALT CONTINUING POLITICAL EFFORTS TO
 INFLUENCE ONLINE DISCOURSE, THIS COURT SHOULD
 AFFIRM.....20
CONCLUSION26
CERTIFICATE OF COMPLIANCE.....27
CERTIFICATE OF SERVICE28

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Ark. Educ. Television Comm’n v. Forbes</i> , 523 U.S. 666 (1998).....	6
<i>Brown v. Entm’t Merchs. Ass’n</i> , 564 U.S. 786 (2011)	25
<i>Garnier v. O’Connor-Ratcliff</i> , 41 F.4th 1158 (9th Cir. 2022).....	14
<i>Herbert v. Lando</i> , 441 U.S. 153 (1979).....	29
<i>Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Boston</i> , 515 U.S. 557 (1995).....	6, 7
<i>Machleder v. Diaz</i> , 801 F.2d 46 (2d Cir. 1986).....	6
<i>Miami Herald v. Tornillo</i> , 418 U.S. 241 (1974).....	6
<i>NetChoice, LLC v. Att’y Gen. Fla.</i> , 34 F.4th 1196 (11th Cir. 2022)	passim
<i>NetChoice, LLC v. Bonta</i> , No. 22-cv-08861-BLF, 2023 U.S. Dist. LEXIS 165500 (N.D.Cal. 2023).....	2, 11, 27
<i>NetChoice, LLC v. Griffin</i> , No. 5:23-CV-05105, 2023 U.S. Dist. LEXIS 154571 (W.D.Ark. 2023)	1, 13, 15, 28
<i>NetChoice, LLC v. Paxton</i> , 573 F. Supp. 3d 1092 (W.D. Tex. 2021)	1, 7
<i>Pac. Gas & Elec. Co. v. Pub. Utilities Comm’n of California</i> , 475 U.S. 1 (1986)	12, 13
<i>Packingham v. North Carolina</i> , 582 U.S. 98 (2017).....	13, 14, 25
<i>Reno v. ACLU</i> , 521 U.S. 844 (1997).	13, 14, 25
<i>Sorrell v. IMS Health Inc.</i> , 564 U.S. 552 (2011).....	7
<i>United States v. Playboy Ent. Grp., Inc.</i> , 529 U.S. 803 (2000).....	31
<i>Volokh v. James</i> , No. 22-cv-10195, 2023 U.S. Dist. LEXIS 25196, (S.D.N.Y. 2023).....	11
<i>Wooley v. Maynard</i> , 430 U.S. 705 (1997)	16
Statutes	
Cal. Civ. Code § 1798.99.31	4
Conn. Gen. Stat. Sec. 42-515 § 10.....	22
Fla. Stat. §§ 106.072	4

N.Y. Gen. Bus. Law § 394-ccc4, 10
 Tex. Bus. & Com. Code § 324.0554
Other Authorities
 Anthony Tellez, *‘Mascara,’ ‘Unalive,’ ‘Corn’: What Common Social Media
 Algospeak Words Actually Mean*, Forbes (Jan. 31, 2023).....23
 Ashutosh Bhagwat, *Do Platforms Have Editorial Rights?*, 1 J. Free Speech L.
 97 (2021)7
 Caitlin Oprysko, *‘Who the hell elected you?’: Cruz blasts Twitter CEO*,
 Politico (Oct. 28, 2020).....24
 Dana DiFilippo, *N.J. legislators propose punishing social media companies
 for kids’ online addiction*, New Jersey Monitor (Feb. 22, 2023).29
 Don't Delete Art, Manifesto14
 Eric Goldman & Jess Miers, *Online Account Terminations/Content Removals
 and the Benefits of Internet Services Enforcing Their House Rules*, 1 J. Free
 Speech L. 191 (2021)8, 10
 Eric Goldman, *The Constitutionality of Mandating Editorial Transparency*,
 73 Hastings Law Journal 1203 (2022)31
 Facebook, *What happens when I report something to Facebook? Does the
 person I report get notified?*19
 FPJ Web Desk, *Improved AI helps reduce hate speech; India makes second-
 highest request for user data*, Free Press Journal 17, 20
 Governor Kathy Hochul, *Governor Hochul Signs Landmark Legislative
 Package to Strengthen Gun Laws and Protect New Yorkers at 12:30–13:07*
 (June 6, 2022).....21
 HuntingNet, Forum Rules9
 Jennifer Pinsof, *This Texas Bill Would Systematically Silence Anyone Who
 Dares to Talk About Abortion Pills*, Electronic Frontier Foundation (March
 13, 2023)31
 Kyooeun Jang, Lulia Pan, Nicol Turner Lee, *The fragmentation of online child
 safety regulations*, Brookings Institute (Aug. 14, 2023)25
 Meta Transparency Center, *Hate Speech Policy Details* 18, 19
 Meta, *Community Standards Enforcement Report (Q1-Q2 2023)*.....16
 Meta, *Facebook Community Standards, Hate Speech*.....8

Meta, Promoting Safety and Expression.....16

News Release, Kathy Hochul, Governor Hochul Signs Landmark Legislative Package to Strengthen Gun Laws and Protect New Yorkers (June 6, 2022).27

News Release, Office of Governor Gavin Newsom, Governor Newsom Signs Nation-Leading Social Media Transparency Measure (Sept. 13, 2022)27

News Release, Office of the Texas Governor, Governor Abbott Signs Law Protecting Texans from Wrongful Social Media Censorship (Sept. 9, 2021)26

News Release, Ron DeSantis, Governor Ron DeSantis Signs Bill to Stop the Censorship of Floridians by Big Tech (May 24, 2021)26

Nihal Krishan, *Anti-cancel culture platforms Rumble and Locals combine to take on Big Tech, Yahoo!* (Oct. 26, 2021)10

On YouTube Kids, Startling Videos Slip Past Filters, N.Y. Times (Nov. 4, 2017).....20

Praveen Paramasivam, *Facebook says it has spent \$13 billion on safety and security efforts since 2016*, Reuters (Sept. 21, 2021)17

Press Release, Anna M. Kaplan, Senator, New York State Senate, Lawmakers Push for Better Enforcement of Social Media Standards Against Hate and Misinformation (May 5, 2021).....29

Press Release, David Carlucci, Senator, New York State Senate, Senator David Carlucci Introduces Legislation to Combat Hate Speech on Social Media (Jan. 21, 2020)28

Ravelry, Policy: Do Not Post In Support of Trump or His Administration10

Rebecca Kern, *Push to rein in social media sweeps the states*, Politico (July 1, 2022)25

Roblox, Community Standards.....9

Sam Bestvater et al., *#BlackLivesMatter Turns 10*, Pew Rsch. Ctr. (Jun. 29, 2023).....13

Snap, Community Guidelines Explainer Series, Hateful Content, Terrorism, and Violent Extremism16

Spandana Singh, *Everything in Moderation* at 12–16, New Am.: Open Tech. Inst.19

TikTok, Community Principles, Safety and Civility16

Tim Cushing, *NJ Legislator Wants State's Cops To Be The New Beneficiaries Of Hate Crime/Bias Laws*, TechDirt (Oct. 13, 2015)30

Tommi Gröndahl, Luca Pajola, Mika Juuti, Mauro Conti, N. Asokan, *All You Need is "Love": Evading Hate-speech Detection*, arXiv:1808.09115 (2018)23

Trust and Safety Models, GitHub23

X Help Center, Report abusive behavior18

X, Help Center, Hateful References and Slurs and Tropes9

INTEREST OF *AMICI CURIAE*¹

NetChoice is a national trade association of online businesses that works to protect free expression and promote free enterprise online. NetChoice’s member organizations have an interest in this proceeding, which arises out of a state’s effort to undermine online services’ editorial freedom to moderate lawful content. The availability of an open internet—free from fragmented, state-level regulation—is critical to NetChoice’s members. For this reason, NetChoice is litigating over government-imposed restrictions on online speech and commerce. *See, e.g., NetChoice, LLC v. Att’y Gen. Fla.*, 34 F.4th 1196 (11th Cir. 2022), cert. pending, No. 22-277 (Sept. 21, 2022); *NetChoice, LLC v. Colmenero* [formerly Paxton], 49 F. 4th 439 (5th Cir. 2022), cert. pending, No. 22-555 (Dec. 15, 2022); *NetChoice, LLC v. Griffin*, No. 5:23-CV-05105. 2023 U.S. Dist. LEXIS 154571 (W.D.Ark. 2023); *Netchoice, LLC v. Bonta*, No. 22-cv-08861-BLF, 2023 U.S. Dist. LEXIS 165500 (N.D.Cal. 2023).

Chamber of Progress is a tech-industry coalition devoted to a progressive society, economy, workforce, and consumer climate. Chamber of Progress backs public policies that will build a fairer, more inclusive country in which the tech

¹ Fed. R. App. P. 29 Statement: No counsel for either party authored this brief in any part. No person or entity other than *amicus* made a monetary contribution to its preparation or submission. All parties have consented to the filing of this brief.

industry operates responsibly and fairly, and in which all people benefit from technological leaps. Chamber of Progress seeks to protect Internet freedom and free speech, promote innovation and economic growth, and empower technology customers and users. In keeping with that mission, Chamber of Progress believes that allowing a diverse range of app-store models and philosophies to flourish will benefit everyone—consumers, store owners, and application developers. Chamber of Progress’s work is supported by its corporate partners, but its partners do not sit on its board of directors and do not have a vote on, or veto over, its positions. Chamber of Progress does not speak for individual partner companies, and it remains true to its stated principles even when its partners disagree.

SUMMARY OF ARGUMENT

The First Amendment prohibits the government from censoring, compelling, or otherwise abridging private speech and media. Yet in passing the “social media networks; hateful conduct prohibited” law, New York joins a growing number of states trying to evade the First Amendment’s constraints to influence what lawful speech appears online.

To evade constitutional scrutiny, the Online Hate Speech Law purports to regulate “hateful *conduct*,” but its definition of conduct, “the use of a social media network to vilify, humiliate, or incite violence against a group or a class of persons,” only includes speech. This is because the *only* way to “use” a social media network is to create, share, consume, or exercise editorial discretion over speech. The First Amendment’s protection for speech does not vanish when the government unilaterally renames it “conduct.”

Large services that host user-generated content, like NetChoice and Chamber of Progress’s members and partners, face a constant battle against malicious actors, including spammers, scammers, and users peddling hateful content. To combat this content—and maintain a viable communications forum—these companies invest in state-of-the-art content moderation systems. Though well-intentioned, the Online Hate Speech Law’s “reporting” and “response” requirements will not improve these

systems' ability to identify and remove offensive content. At times, they may undermine existing efforts to combat hateful content.

Persistent political efforts to regulate online speech illustrate confusion about the First Amendment's application to social media services. Though these efforts differ in form, their shared goal is to grant the state influence over private editorial standards. But expression is protected no matter what medium it appears on, and a finding for New York would lead to absurd consequences. One state might require reporting, responses, and disclosures about lawful content perceived by some users to "vilify" or "humiliate." And other states might enforce similar statutes against "hateful" criticism of the police or LGBTQ+ advocacy, leading to a 50-state patchwork of editorial requirements based on which messages local politicians disfavor. Without confirmation from this Court that the First Amendment prohibits the Online Hate Speech Law, political efforts to interfere with online expression will continue to proliferate in the Second Circuit.

ARGUMENT

I. NEW YORK CANNOT REGULATE SPEECH BY SIMPLY CALLING IT 'CONDUCT'

The First Amendment prohibits the government from censoring, compelling, or otherwise abridging speech and media. That this fundamental guarantee cannot be tossed aside in the context of online expression "would be too obvious to mention if it weren't so often lost or obscured in political rhetoric." *NetChoice, LLC v. Att'y*

Gen. Fla., 34 F.4th 1196, 1204 (11th Cir. 2022). Yet in passing the “social media networks; hateful conduct prohibited” law, New York joins a spike of states trying to evade the First Amendment’s constraints to regulate social media. N.Y. Gen. Bus. Law § 394-ccc; *see also* Cal. Civ. Code § 1798.99.31(a)(1) (requiring social media companies to disclose “whether minors may witness . . . potentially harmful *conduct*” on social media).²

To avoid First Amendment scrutiny, the Online Hate Speech Law purports to regulate “hateful conduct” on social media networks, rather than speech, but tellingly, its definition of “hateful conduct” *only* includes speech. Social media networks are, by definition, networks engaged in speech—specifically, networks engaged in publishing third-party speech. They exist solely to connect speakers, to publish, distribute and curate users’ content, and to “engage in some speech of their own.” *Att’y Gen. Fla.*, 34 F.4th at 1204. By design, then, New York’s law targets constitutionally protected publishers and expression based on content—not conduct. This Court should affirm to clarify the First Amendment protects speech even when the government renames it “conduct.”

² *See also*, Tex. Bus. & Com. Code § 324.055(2)(b) (banning social media services from removing content based on the “viewpoint” it expresses); Fla. Stat. §§ 106.072 (2022); 501.2041(2)(c) (banning social media services from removing content from “registered political candidates” and from changing their community guidelines more than once per month).

A. Defining, Publishing, and Enforcing Community Guidelines is Speech

Private media outlets have a well-established right to “exercise[] editorial discretion in the selection and presentation” of speech on their own services. *Ark. Educ. Television Comm’n v. Forbes*, 523 U.S. 666, 674 (1998). The government “cannot substitute its judgment for that of the press”; “the choice of material . . . and treatment of public issues and public officials—whether fair or unfair—constitute the exercise of editorial control and judgment.” *Machleder v. Diaz*, 801 F.2d 46, 55 (2d Cir. 1986); *Miami Herald v. Tornillo*, 418 U.S. 241, 258 (1974). And this right is not “restricted to the [traditional] press.” *Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Boston*, 515 U.S. 557, 574 (1995). It applies equally to advocacy groups, “business corporations,” and “ordinary people engaged” in any kind of “expression” and “dissemination of information,” because that is “speech within the meaning of the First Amendment.” *Id.*; *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 570 (2011). Accordingly, social media services’ “decisions about whether, to what extent, and in what manner to disseminate third party-created content to the public are [also] editorial judgments protected by the First Amendment.” *AG, Fla.*, 34 F.4th at 1213.

Online services exercise editorial freedom, in part, by setting and publishing community guidelines. They use these written policies to express their values and “convey a message about the type of community the platform seeks to foster.”

NetChoice v. Paxton, 573 F. Supp. 3d 1092, 1108 (N.D.Tex. 2021). Indeed, crafting community guidelines and “other moderation rules” on hate speech, violence, harassment, electoral falsehoods, and nudity, among others, is “replete with ideological choices.” Ashutosh Bhagwat, *Do Platforms Have Editorial Rights?*, 1 J. Free Speech L. 97, 116 (2021) (explaining defining contours of speech policies reflect services’ values and priorities).³ Through “publish[ing] terms of service or community standards specifying the type of content that it will (and won't) allow on its site,” online services “engage in some speech of their own.” *AG, Fla.*, 34 F.4th at 1204.

While community guidelines on “certain types of . . . content have trended towards industry-wide convergence,” there is “no single universal standard for what’s prohibited by house rules.” Eric Goldman & Jess Miers, *Online Account Terminations/Content Removals and the Benefits of Internet Services Enforcing Their House Rules*, 1 J. Free Speech L. 191, 195 (2021) [*hereinafter* “Goldman & Miers, *House Rules*”].⁴ For example, Meta’s community guidelines provide a definition of prohibited hate speech that is *more restrictive* than New York’s in one sense. Meta bans “direct attack[s] against people” based on characteristics included in the Online Hate Speech Law, as well as based on “caste” and “serious disease.”

³ Available at <https://tinyurl.com/2phr9raa>.

⁴ Available at <https://tinyurl.com/6r3wh7f7>.

Meta, Facebook Community Standards, Hate Speech (last visited Sept. 20, 2023).⁵ By contrast, X requires that a user have “intent” to harass or “degrade” to violate their policies. X, Help Center, Hateful References and Slurs and Tropes (last visited Sept. 20, 2023).⁶ These differences reflect the unique speech preferences of the services and their users.

Like a traditional newspaper’s editorial processes, defining, publishing, and enforcing social media “house rules” on a wide range of subject matter is, itself, an act of expression—one that creates distinct communities. Roblox explains that its moderators exert a heavy hand over user-generated content to ensure the company stays an appropriate place for children—banning “flirtatious gestures” and “sexually suggestive avatar bodies and clothing items,” among other things. Roblox, Community Standards (last visited Sept. 20, 2023).⁷ Likewise, HuntingNet fosters a valuable niche platform for its users by excluding pro-animal rights posts,⁸ and Ravelry, an online knitting community, excludes pro-Trump political content that, its users determined, interfered with their enjoyment of the site. HuntingNet, Forum

⁵ Available at <https://tinyurl.com/2386389z>.

⁶ Available at <https://tinyurl.com/yc47rc48>.

⁷ Available at <https://tinyurl.com/yc2ywx3r>.

⁸ Available at <https://tinyurl.com/49mahzet>.

Rules (last visited Sept. 20, 2023); Ravelry, Policy: Do Not Post In Support of Trump or His Administration (last visited Sept. 20, 2023).⁹

Though all large social media services publish community guidelines, many smaller ones do not employ them at all, or do not employ them as to certain subject matter. *See generally*, Goldman & Miers, *House Rules*. This silence, too, conveys a message about what type of community it hopes to foster. “Differences in house rules creates a key point of competitive differentiation as services customize their offerings to the needs of their users.” *Id.* at 195. Rumble, for example, offers a competitive differentiation by promising that users’ posts will be subject to substantially *less-restrictive* content rules than those offered by NetChoice and CHOP’s members. Pl.-App. Br. 11; Nihal Krishan, *Anti-cancel culture platforms Rumble and Locals combine to take on Big Tech*, Yahoo! (Oct. 26, 2021) (explaining “alternative” platforms Rumble and Locals offer “online users a new way to interact and make money without fear of being canceled.”).¹⁰ By offering distinct guidelines, social media companies foster distinct communities with varied interests.

Because defining, publishing, and enforcing “written polic[ies] [is] speech,” not conduct, the government may not burden private services’ freedom to define (or not define), publish (or not publish), and enforce (or not enforce) their guidelines

⁹ Available at <https://tinyurl.com/ymp87fdm>.

¹⁰ Available at <https://tinyurl.com/mrxdbme>.

against users' constitutionally protected speech. *Volokh v. James*, No. 22-cv-10195, 2023 U.S. Dist. LEXIS 25196, *12 (S.D.N.Y. 2023); *see also NetChoice, LLC v. Bonta*, No. 22-cv-08861-BLF, 2023 U.S. Dist. LEXIS 165500 (N.D.Cal. 2023) *24 (law requiring social media services to "provide prominent, accessible, and responsive tools to . . . report concerns" is a regulation on speech, not conduct). Nor can they force them to remark on users' lawful expression just because the state disapproves of it. *Id.* (law that "facially requires a business to express its ideas and analysis about likely harm" to children is a regulation on speech not conduct); *See generally, Pac. Gas & Elec. Co. v. Pub. Utilities Comm'n of California*, 475 U.S. 1 (1986) (the government cannot compel corporate speakers to share political messages they disagree with).

But the Online Hate Speech Law does both. First, it requires a social media network to endorse the state's position on "hateful conduct." A social media network must make a "concise policy readily available and accessible on their website and application" detailing how the network will "respond and address the reports of incidents of hateful conduct on their platform." N.Y. Gen. Bus. Law § 394-ccc(3). Implicit in this language is the requirement that each social media network's definition of "hateful conduct" must, at minimum, include the definition set forth in the law itself.

Second, forcing "social media network[s]" to create a "clear and easily

accessible” Report & Response Mechanism for the state’s definition of “prohibited” “conduct” compels them to broadcast the state’s ideological stance on speech on their own services. N.Y. Gen. Bus. Law § 394-ccc(3). But private services have the right to choose what to say *and* “what not to say,” and they cannot be compelled to “propound political messages with which they disagree.” *Pac. Gas & Elec. Co.* 475 U.S. at 16.

B. Sharing and Receiving Content on Social Media is Speech

When users share and receive speech on social media, they engage in speech, not conduct. *See generally, NetChoice, LLC v. Griffin*, No. 5:23-CV-05105. 2023 U.S. Dist. LEXIS 154571, *51 (social media age-verification law which restricts users’ ability to share and receive speech on social media likely violates the First Amendment). People rely on social media to engage with social and political movements, to share news and art, and to participate in communities. Indeed, among the “most important places . . . for the exchange of views” today are “the ‘vast democratic forums of the Internet’ in general . . . and social media in particular.” *Packingham v. North Carolina*, 582 U.S. 98, 104 (2017) (quoting *Reno v. ACLU*, 521 U.S. 844, 868 (1997)). Black Lives Matter, for example, “stands as a model of a new generation of social movements intrinsically linked to social media.” Sam Bestvater et al., *#BlackLivesMatter Turns 10*, Pew Rsch. Ctr. (Jun. 29, 2023).¹¹ And

¹¹ Available at <https://tinyurl.com/2kvm4yph>.

“from local county supervisors and state representatives to the President of the United States, elected officials across the country increasingly rely on social media both to promote their campaigns and . . . to communicate with constituents and seek their input in carrying out their duties as public officials.” *Garnier v. O’Connor-Ratcliff*, 41 F.4th 1158, 1163 (9th Cir. 2022), cert. granted, 143 S. Ct. 1779 (2023). In its open letter asking large social media services to make community guidelines less restrictive for “objectionable” content, the artists’ coalition Don’t Delete Art wrote: “Social media is a critical avenue for artistic exposure and expression; in 2020, it replaced art fairs as the third most successful way for galleries to sell art.” Don’t Delete Art, Manifesto (March 3, 2023) (last visited September 20, 2023).¹² “In short, social media users employ these websites to engage in a wide array of activity on topics ‘as diverse as human thought.’” *Packingham*, 582 U.S. at 105 (quoting *Reno*, 521 U.S. at 852). These activities are the core of what the First Amendment protects.

To evade constitutional scrutiny, the Online Hate Speech Law purports to regulate “hateful *conduct*,” but its definition of conduct, “the use of a social media network to vilify, humiliate, or incite violence against a group or a class of persons,” only includes speech. This is because the *only* way to “use” a social media network is to create, share, consume, or exercise editorial discretion over speech. When users

¹² Available at <https://tinyurl.com/mn6ncptj>.

share or receive information, ideas, interests, and other forms of expression on social media, they engage in speech. *Griffin*, No. 5:23-CV-05105, 2023 U.S. Dist. LEXIS 154571, *51 (social media age-verification law which restricts users’ ability to share and receive speech on social media likely violates the First Amendment). Likewise, when a user or online service “publishes terms of service or community standards . . . adds addenda or disclaimers to certain posts (say, warning of misinformation or mature content), or publishes its own posts,” they engage in speech. *AG, Fla.*, 34 F.4th at 1204.

By requiring online services to create a policy and mechanism for users to report certain lawful content, the Online Hate Speech Law compels them to express the state’s own message (here, disdain or disapproval) for that content. But speech compulsions are as constitutionally suspect as speech *restrictions*, and, as the district court correctly noted, New York cannot “constitutionally require citizens to display the state’s” message when that message is offensive to their convictions. *Volokh*, No. 22-cv-10195, 2023 U.S. Dist. LEXIS 25196; *Wooley v. Maynard*, 430 U.S. 705, 713 (1997). To ensure the First Amendment’s protections are applied with full force online, this Court should affirm the judgment.

II. THE ONLINE HATE SPEECH LAW WILL NOT AID EFFORTS TO IDENTIFY AND REMOVE HARMFUL CONTENT ONLINE

Though New York’s general goal to limit the spread of harmful content online is one shared by NetChoice and Chamber of Progress’s members and partners, the Online Hate Speech Law will not aid efforts to identify and remove this kind of speech.¹³ Services that host user-generated content—especially the world’s most popular ones—face a constant battle against spammers, scammers, and users peddling hate speech. Meta, Community Standards Enforcement Report (Q1-Q2 2023) (explaining that Facebook has removed 2.7 billion posts for “spam,” 28 million posts for “hate speech,” and 27 million “violent and graphic” posts in the first two quarters of 2023, alone).¹⁴ To combat malicious content—and maintain a viable forum for a diverse user base—these companies invest in state-of-the-art content moderation tools to quickly identify and remove content which violates their guidelines. FPJ Web Desk, *Improved AI helps reduce hate speech; India makes second-highest request for user data*, Free Press Journal (Nov. 23, 2022) (explaining investment in AI content moderation technology is enhancing accuracy of detection)¹⁵; *see also* Praveen Paramasivam, *Facebook says it has spent \$13 billion*

¹³ For some examples of NetChoice and Chamber of Progress’s members and partners’ policies on hate speech, *see* Meta, Promoting Safety and Expression (last visited Sept. 20, 2023), <https://tinyurl.com/v5cmcv6>; TikTok, Community Principles, Safety and Civility (last visited Sept. 20, 2023), <https://tinyurl.com/35ptfnbf>; and Snap, Community Guidelines Explainer Series, Hateful Content, Terrorism, and Violent Extremism (last visited Sept. 20, 2023), <https://tinyurl.com/26tss2xa>.

¹⁴ Available at <https://tinyurl.com/5xru534v>.

¹⁵ Available at <https://tinyurl.com/msdxun47>.

on safety and security efforts since 2016, Reuters (Sep. 21, 2021).¹⁶ The Online Hate Speech Law’s “reporting” and “response” requirements will not improve these systems’ ability to identify and remove violative content on large social media services. Instead, it will produce unhelpful and redundant data and create significant administrative burdens on existing systems that otherwise frustrate efforts to combat hate speech. Likewise, the “clear and concise policy” requirement may actually hamper existing content moderation efforts by pressuring services to abandon nuance in moderation techniques and to divulge sensitive details about how violative content is identified.

The Online Hate Speech Law’s mandatory reporting mechanism, as applied to large social media networks, is a redundant effort that will not improve content moderators’ ability to identify and remove hateful content. Major social media networks already have extensive hateful conduct policies and readily available content reporting systems. *See, e.g.* Meta Transparency Center, Hate Speech Policy Details (last visited Sept. 20, 2023);¹⁷ Reddit, Reddit Report (last visited Sept. 20, 2023);¹⁸ X Help Center, Report abusive behavior (last visited Sept. 20, 2023).¹⁹ And users already make considerable use of the reporting systems provided to them. In

¹⁶ Available at <https://tinyurl.com/3zt4wkhz>.

¹⁷ Available at <https://tinyurl.com/bdf6m29s>.

¹⁸ Available at <https://tinyurl.com/3ktx8usp>.

¹⁹ Available at <https://tinyurl.com/mr2hwhb2>.

the first two quarters of 2022, alone, Twitter users made over 800,000 reports. *Id.* When “something gets reported,” human and algorithmic systems “review it and take action on anything [that] doesn't follow our Community Standards.” Facebook, What happens when I report something to Facebook? Does the person I report get notified? (last visited Sept. 20, 2023).²⁰ Content reporting systems can be valuable tools, but the inherently subjective nature of these reports makes them an unreliable resource for identifying hate speech, as defined by a service’s community guidelines, online. Because users tend to report what *they* believe should not belong on the platform, rather than what the services’ community guidelines prohibit, the reports are often unhelpful to moderators. For example, Blue Lives Matter and Black Lives Matter are both routinely flagged as hate speech on large social media services – neither of which is content most large services want to take down. The same can be said for content about LGBTQ+ people and reproductive health care.

The vast majority of violative content is removed by AI systems, which “proactively and retroactively detect[] and remove[] the . . . violating content.” Meta Transparency Center, Hate Speech Policy Details;²¹ *see generally*, Spandana Singh, *Everything in Moderation* at 12–16, New Am.: Open Tech. Inst. (July 22, 2019) (explaining how automated tools are used in the content moderation process on large

²⁰ Available at <https://tinyurl.com/2p95ejrt>.

²¹ Available at <https://tinyurl.com/bdf6m29s>.

social media services).²² Though they are continually improving over time, these systems are not perfect—bad actors always adapt to existing proactive security measures. *See generally*, The Free Press Journal, *Improved AI helps reduce hate speech, supra*. This forces services into an interminable cat and mouse game to identify and respond to innovative new means to evade detection. Sapna Maheshwari, *On YouTube Kids, Startling Videos Slip Past Filters*, N.Y. Times (Nov. 4, 2017) (explaining how bad actors found ways to “fool” content moderation algorithms to post disturbing variations of popular children’s cartoons).²³ But identification systems adapt and improve over time, and mandating additional user reporting systems for content a user perceives to “vilify,” “humiliate,” and “incite” —whether or not that content violates services’ community guidelines—will not aid these systems’ development.

The Online Hate Speech Law’s reply requirement imposes significant administrative burdens on services—large and small—that host user-generated content, ratcheting up the cost of publishing speech. If the Attorney General enforces the law to *require* social media networks to “provide a direct response” to each individual complainant “informing them of how the matter is being handled,” as Plaintiffs-Appellees expect, compliance becomes untenable. Pl.-App. Br. 24.

²² Available at <https://tinyurl.com/yhx34m7b>.

²³ Available at <https://tinyurl.com/3mwjab8v>.

Governor Kathy Hochul, *Governor Hochul Signs Landmark Legislative Package to Strengthen Gun Laws and Protect New Yorkers* at 12:30–13:07 (June 6, 2022) (Governor Hochul suggesting that “[o]ur great leader, our Attorney General, will be championing this cause with every power her office can bring in [to enforce the Online Hate Speech Law].”).²⁴ Last year, the United States Court of Appeals for the Eleventh Circuit granted a preliminary injunction on part of Florida’s “anti-Big Tech censorship” law in *NetChoice, LLC v. Att’y Gen. Fla.*, which required large social media services to provide users “detailed notice” before they removed or deprioritized the user’s content. 34 F.4th at 1203. The Court reasoned this requirement was unduly burdensome and that it imposed impossible implementation costs, given the number of posts users publish each day. If the Hateful Conduct Law is also enforced to require direct responses to each report, it would impose similar burdens on large services’ speech and editorial processes. And for smaller services that lack the resources to automate even parts of the process, the reply requirement will goad them to “more aggressively remove content, and/or limit the ability for registered users to participate in [their] comment sections” to avoid the Attorney General’s scrutiny. Pl.-App. Br. 125.

²⁴ Available at <https://tinyurl.com/4vf5atad>.

Beyond not aiding existing content moderation efforts, the response requirement may actually hamper existing systems' effectiveness by requiring services to provide bad actors with a blueprint to circumvent platforms' existing content moderation policies. As Plaintiffs-Appellees note, it isn't at all obvious what type of response will satisfy the law's requirements. Pl.-App. Br. 63. Nor is it clear whether a platform's response can describe their content moderation in broad terms, or whether platforms must provide an exact explanation of how the user's specific "matter" was handled. Given the imprecise language in the "clear and concise" policy and "direct response" requirements, the Attorney General has liberty to enforce these provisions as an editorial transparency requirement—to force companies to reveal detailed information about how hate speech is detected, and the criteria content moderators use when deciding to remove content. This would be a boon to bad actors seeking to evade those same detection measures.

Online services use content moderation tools, in part, as security measures. For the same reason banks do not disclose details of how they detect fraud, online services resist disclosure of the particularities of their editorial procedures, knowing that bad actors can and will uncover creative ways to dodge them. Trust and Safety Models, GitHub (last visited Sep. 20, 2023) (Twitter engineers explaining they did not open source all their trust and safety classifiers because of the "adversarial nature

of this area”).²⁵ For example, just by knowing that a hate speech detector examines each word in a post rather than each character, users can avoid detection for racial epithets by randomly adding or removing spaces between letters. Tommi Gröndahl, Luca Pajola, Mika Juuti, Mauro Conti, N. Asokan, *All You Need is "Love": Evading Hate-speech Detection*, arXiv:1808.09115 at 6 (2018),²⁶ see also Anthony Tellez, ‘Mascara,’ ‘Unalive,’ ‘Corn’: What Common Social Media Algospeak Words Actually Mean, *Forbes* (Jan. 31, 2023) (explaining “algospeak” that is used to evade content moderation, such as “unalive” for suicide, “corn” for pornography, and “mascara” for sexual assault).²⁷ The enforcement powers in the Online Hate Speech could backfire to inadvertently sabotage security measures and, effectively, empower the bad actors this law seeks to address.

III. TO HALT CONTINUING POLITICAL EFFORTS TO INFLUENCE ONLINE DISCOURSE, THIS COURT SHOULD AFFIRM

In a Senate Commerce Committee hearing, Senator Ted Cruz challenged Twitter’s then-CEO Jack Dorsey to explain “Who the hell elected you and put you in charge of what the media are allowed to report and what the American people are allowed to hear?” Caitlin Oprysko, *Who the hell elected you?’: Cruz blasts Twitter CEO*, *Politico* (Oct. 28, 2020).²⁸ Implicit in this rhetorical question is the idea that

²⁵ Available at <https://tinyurl.com/ypc56rp5>.

²⁶ Available at <https://tinyurl.com/2uha3j34>.

²⁷ Available at <https://tinyurl.com/5y24w3cy>.

²⁸ Available at <https://tinyurl.com/56sm33ft>.

elected officials should be in charge of online speech. Thirty years after the internet’s inception, the First Amendment’s protections still “do not vary when a new and different medium for communication appears.” *Brown v. Entm’t Merchs. Ass’n*, 564 U.S. 786, 790 (2011); *see also Reno* 521 U.S. at 870 (establishing First Amendment protection for online media is coextensive with offline media); *see also Packingham*, 582 U.S. at 105 (affirming that the First Amendment applies equally online). Yet widespread confusion among state government officials across the political spectrum appears to remain—states continue to introduce bills which compel, censor, or restrict access to speech on the internet. *See generally*, Kyooeun Jang, Lulia Pan, Nicol Turner Lee, *The fragmentation of online child safety regulations*, Brookings Institute (Aug. 14, 2023) (explaining the “disjointed patchwork of regulation” being created by efforts to require government-issued ID to use social media in Arkansas, Texas, Utah, California, and Louisiana)²⁹; Rebecca Kern, *Push to rein in social media sweeps the states*, Politico (July 1, 2022) (explaining that politicians across the country introduced over 100 bills in 2022 to control what content gets shared on the internet).³⁰

Though these regulatory efforts differ in form, their shared goal is to grant the state influence over private editorial standards. *See, e.g.*, News Release, Ron

²⁹ Available at <https://tinyurl.com/mufsp34k>.

³⁰ Available at <https://tinyurl.com/mryw6392>.

DeSantis, Governor Ron DeSantis Signs Bill to Stop the Censorship of Floridians by Big Tech (May 24, 2021) (announcing passage of SB 7072 to stop “Silicon Valley censorship” of conservatives on social media platforms); News Release, Office of the Texas Governor, Governor Abbott Signs Law Protecting Texans from Wrongful Social Media Censorship (Sept. 9, 2021) (announcing passage of HB20 to stop censorship of “conservative viewpoints and ideas”); News Release, Office of Governor Gavin Newsom, Governor Newsom Signs Nation-Leading Social Media Transparency Measure (Sept. 13, 2022) (announcing passage of bill like Online Hate Speech Law, AB 587, to “protect Californians from hate, harassment and lies spread online.”). New York passed its Online Hate Speech Law after identifying social media services as “unchecked [vehicles] for dangerous and corrosive ideas to spread.” News Release, Kathy Hochul, Governor Hochul Signs Landmark Legislative Package to Strengthen Gun Laws and Protect New Yorkers (at 12:30–13:07) (June 6, 2022). The bill, Governor Hochul said, would allow her office to “expand our work . . . to address this growing threat” from “dangerous and hateful platforms.” *Id.* at 35:50–36:35. Likewise, in its recent order preliminarily enjoining a similar social media law, AB 2273, the United States District Court for the Northern District of California noted that California “does not deny that the end goal of [the law] is to reduce the amount of harmful content”—harmful in the State’s estimation, that is—on the internet. *Bonta*, No. 22-cv-08861-BLF, 2023 U.S. Dist.

LEXIS 165500, *24. There can be no doubt: the underlying goal of social media “transparency” legislation like the Online Hate Speech Law is to give the government influence over online publishers' editorial processes. But social media services are, “for many, the principal sources for knowing current events, and exploring the vast realms of human thought and knowledge,” and users have a right to access these sources, untampered with by the government’s own judgments. *Griffin*, No. 5:23-CV-05105, 2023 U.S. Dist. LEXIS 154571, *51 (quoting *Packingham*).

Allowing the Online Hate Speech Law to stand would greenlight a growing patchwork of political efforts to influence online editorial processes in the Second Circuit. *See, e.g.*, Press Release, David Carlucci, Senator, New York State Senate, Senator David Carlucci Introduces Legislation to Combat Hate Speech on Social Media (Jan. 21, 2020), (announcing bill to require social media networks to remove or block “hate speech” within 24 hours of receiving a visitor complaint and “immediately” notify the complainant of a decision);³¹ Conn. Gen. Stat. Sec. 42-515 § 10(e) (requiring online services to submit public reports about services, products, or features that “pose a heightened risk of harm to minors.”); Press Release, Anna M. Kaplan, Senator, New York State Senate, Lawmakers Push for Better Enforcement of Social Media Standards Against Hate and Misinformation (May 5,

³¹ Available at <https://tinyurl.com/3vtupjjk>.

2021) (announcing bill to “crack down” on hate and misinformation, as defined by the state)³²; *see also* Dana DiFilippo, *N.J. legislators propose punishing social media companies for kids’ online addiction*, *New Jersey Monitor* (Feb. 22, 2023) (announcing bill to sue online services for algorithms that recommend “addictive” content).³³ All these bills—and dozens of others across the country—push social media services to avoid regulatory scrutiny by removing certain lawful speech, allowing the government to “accomplish indirectly via market manipulation what it cannot do through direct regulation—control the available channels for political discussion.” *Washington Post v. McManus*, 944 F.3d 506, 517 (4th Cir. 2019).

But if New York can impose reporting, response, and disclosure requirements on every website that hosts lawful content its politicians disfavor, so too could lawmakers in other states enforce similar statutes against speech that criticizes elected officials, casts law enforcement in a bad light, discusses reproductive healthcare, and supports or opposes gender transitions. *See, e.g.*, Tim Cushing, *NJ Legislator Wants State's Cops To Be The New Beneficiaries Of Hate Crime/Bias Laws*, *TechDirt* (Oct. 13 2015);³⁴ Jennifer Pinsof, *This Texas Bill Would Systematically Silence Anyone Who Dares to Talk About Abortion Pills*, *Electronic Frontier Foundation* (March 13, 2023) (analyzing Texas House Bill 2690 which

³² Available at <https://tinyurl.com/ycxdes3a>.

³³ Available at <https://tinyurl.com/yck23py9>.

³⁴ Available at <https://tinyurl.com/4vdhc4j8>.

would criminalize online speech about abortion-inducing drugs).³⁵ Fifty different state-enforced “clear and concise” explanations of how private services moderate arbitrarily chosen content is an absurd result that would be unthinkable if applied in the context of traditional media: “[t]here is no law that subjects [a newspaper’s] editorial process to private or official examination merely to satisfy curiosity or to serve some general end such as the public interest; and if there were, it would not survive constitutional scrutiny.” *Herbert v. Lando*, 441 U.S. 153, 174 (1979). Online editorial processes are no different. *See generally*, Eric Goldman, *The Constitutionality of Mandating Editorial Transparency*, 73 *Hastings Law Journal* 1203 (2022) (emphasizing that editorial disclosure requirements—like the laws enacted in California, Florida, New York, and Texas—cater to “satisfying curiosity” or pursuing vague public interest objectives).³⁶ Legislatures across the country need a reminder that the government cannot “burden the speech of others in order to tilt public debate in a preferred direction,” whether that speech occurs offline or online. *Att’y Gen. Fla.*, 34 F.4th at 1228 (quoting *Sorrell*, 564 U.S. at 578–79).

³⁵ Available at <https://tinyurl.com/yckvwrf7>.

³⁶ Available at <https://tinyurl.com/yw7rc2nr>.

CONCLUSION

For the reasons stated above, and those offered by Plaintiffs-Appellees, the court should affirm the lower court's judgment issuing a preliminary injunction.

Respectfully submitted,³⁷

DATED: September 26, 2023

/s/ Nicole Saad Bembridge

Nicole Saad Bembridge
Counsel of Record
Christopher J. Marchese
Carl Szabo
Paul D. Taske
NetChoice
1401 K St., Suite 501
Washington, D.C., 20005

³⁷ Credit to Soham G. Mehta, NetChoice's former Google Public Policy Fellow, for significant research support.

CERTIFICATE OF COMPLIANCE

1. This brief complies with the type-volume limitation of Fed. R. App. P. 29(a)(5) and Fed. R. App. P. 32(a)(7)(B) because it contains 5,404 words, excluding the parts exempted by Fed. R. App. P. 32(f).
2. This brief complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type-style requirements of Fed. R. App. P. 32(a)(6) because it has been prepared in a proportionally spaced typeface in Times New Roman, 14-point font.

/s/ Nicole Saad Bembridge
September 26, 2023

CERTIFICATE OF SERVICE

I hereby certify that I electronically filed the foregoing with the Clerk of Court, who will enter it into the CM/ECF system, which will send a notification of such filing to the appropriate counsel.

/s/ Nicole Saad Bembridge

September 26, 2023