

NO. 24-271

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

X CORP.,

PLAINTIFF-APPELLANT,

v.

ROBERT BONTA, in his official capacity as Attorney General of California,

DEFENDANT-APPELLEE.

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

**BRIEF OF AMICUS CURIAE ELECTRONIC FRONTIER
FOUNDATION IN SUPPORT OF PLAINTIFF-APPELLANT
AND REVERSAL**

David Greene
Aaron Mackey
ELECTRONIC FRONTIER FOUNDATION
815 Eddy Street
San Francisco, CA 94109
Email: davidg@eff.org
Telephone: (415) 436-9333
Fax: (415) 436-9993

Counsel for Amicus Curiae

CORPORATE DISCLOSURE STATEMENT

Pursuant to Rule 26.1 of the Federal Rules of Appellate Procedure, amicus states that they do not have a parent corporation and that no publicly held corporation owns 10% or more of their stock.

Dated: February 21, 2024

By: /s/ David Greene
David Greene

TABLE OF CONTENTS

CORPORATE DISCLOSURE STATEMENT	i
TABLE OF AUTHORITIES.....	iv
STATEMENT OF INTEREST	1
INTRODUCTION.....	2
ARGUMENT	2
I. The Content Moderation Systems Targeted by the Law Are Editorial Processes Protected by the First Amendment.	2
A. Content Moderation Is an Inherently Expressive Editorial Process.	3
1. Content Moderation Bears the Hallmarks of a Typical Editorial Process.	3
2. Social Media Platforms Have Rules, Standards and Guidelines about What Content They Want and Don’t Want on Their Sites.....	7
3. Social Media Follows A Long Line of Traditional Media That Have Curated Content Written By Others.....	11
B. Content Moderation is Protected by the First Amendment Just as Other Editorial Process Are.	12
C. Government Involvement in Content Moderation Raises Serious First Amendment Concerns Because Content Moderation Is Particularly Vulnerable to Government-Prompted Chilling Effects...	17
II. AB587’S SPEECH COMPULSIONS MUST BE SUBJECT TO STRICT SCRUTINY, NOT THE MORE PERMISSIVE <i>ZAUDERER</i> STANDARD.	20
A. Speech That Is Part of and Related to the Editorial Process Is Not Commercial Speech Nor Otherwise Subject to Diminished First Amendment Protection	20
B. <i>Zauderer</i> Does Not Apply to Intrusions on Editorial Privacy	22

C.	Speech Compulsions that are Informal Censorship Schemes are Subject to Exacting Scrutiny.....	24
D.	<i>Zauderer</i> Does Not Apply Because Content Moderation Processes are Inherently Subjective and Nonfactual.....	25
E.	There is a Significant Public Interest in Promoting Content Moderation Transparency	29
III.	GOVERNMENTS CAN LOOK TO NON-EDITORIAL REGULATIONS TO BENEFIT SOCIAL MEDIA USERS	30
	CONCLUSION	32
	CERTIFICATE OF COMPLIANCE	33
	CERTIFICATE OF SERVICE.....	34

TABLE OF AUTHORITIES

Cases

<i>303 Creative LLC v. Elenis</i> , 600 U.S. 570 (2023).....	14
<i>American Beverage Association v. City & County of San Francisco</i> , 916 F.3d 749 (9th Cir. 2019)	21
<i>Application of Consumers Union of U. S., Inc.</i> , 495 F. Supp. 582 (S.D.N.Y. 1980).....	23
<i>Bantam Books, Inc. v Sullivan</i> , 372 U.S. 58 (1963).....	24, 25
<i>Bursey v. United States</i> , 466 F.2d 1059 (9th Cir. 1972)	23
<i>Central Hudson Gas & Elec. Corp. v. Pub. Serv. Comm’n of New York</i> , 447 U.S. 557 (1980).....	21
<i>Children’s Health Def. v. Facebook Inc.</i> , 546 F. Supp. 3d 909 (N.D. Cal. 2021).....	15
<i>City of Cincinnati v. Discovery Network, Inc.</i> , 507 U.S. 410 (1993).....	22
<i>City of Lakewood v. Plain Dealer Publishing Co.</i> , 486 U.S. 750 (1988).....	21
<i>Dex Media W., Inc. v. City of Seattle</i> , 696 F.3d 952 (9th Cir. 2012)	22
<i>Herbert v. Lando</i> , 441 U.S. 153 (1974).....	24
<i>Huber v. Biden</i> , No. 21-CV-06580-EMC, 2022 WL 827248, (N.D. Cal. Mar. 18, 2022).....	15
<i>Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. Of Boston</i> , 515 U.S. 557 (1995).....	14
<i>In re Consumers Union of the U.S., Inc.</i> , 32 Fed. R. Serv. 2d 1373 (S.D.N.Y. 1981).....	23
<i>In re Grand Jury Proceedings</i> , 863 F.2d 667 (9th Cir. 1988)	23

<i>Kenneally v. Suzuki Motor Co., Ltd.</i> , No. M-8-85, 1994 WL 48840 (S.D.N.Y. Feb. 10, 1994).....	23
<i>Los Angeles v. Preferred Communications, Inc.</i> , 476 U.S. 488 (1986).....	13
<i>Manhattan Community Access Corp. v. Halleck</i> , 139 S. Ct. 1921 (2019).....	13
<i>Miami Herald Publishing Co. v. Tornillo</i> , 418 U.S. 241 (1974).....	13, 16
<i>Miller v. California</i> , 413 U.S. 15 (1973).....	4
<i>Moody v. NetChoice, LLC</i> , 216 L. Ed. 2d 1313 (Sept. 29, 2023).....	13, 16
<i>Nat’l. Ass’n. of Wheat Growers v. Bonta</i> , 85 F.4th 1263 (9th Cir. 2023)	21, 25
<i>Nat’l. Cable Assn. v. FCC</i> , 555 F.3d 996 (D.C. Cir. 2009)	30
<i>NetChoice LLC v. Attorney General, Florida</i> , 34 F.4th 1196 (11th Cir. 2022)	16
<i>NetChoice, LLC v Paxton</i> , 49 F.4th 439 (5th Cir. 2022)	16
<i>NetChoice, LLC v. Paxton</i> , 216 L. Ed. 2d 1313 (Sept. 29, 2023).....	13, 16
<i>New York Times Co. v. Sullivan</i> , 376 U.S. 254 (1964).....	14, 22
<i>O’Handley v. Padilla</i> , 579 F. Supp. 3d 1163 (N.D. Cal. 2022),	15
<i>O’Handley v. Weber</i> , 62 F.4th 1145 (9th Cir. 2023)	15
<i>Packingham v. North Carolina</i> , 137 S. Ct. 1730 (2017).....	15
<i>Pittsburgh Press Co. v. Pittsburgh Comm’n on Human Relations</i> , 413 U.S. 376, (1973).....	21
<i>Prager Univ. v. Google LLC</i> , 951 F.3d 991 (9th Cir. 2020)	15

<i>Shoen v. Shoen</i> , 48 F.3d 412 (9th Cir. 1995)	24
<i>Smith v. California</i> , 361 U.S. 147 (1959).....	22
<i>Stormans, Inc. v. Weisman</i> , 794 F.3d 1064 (9th Cir. 2015)	21
<i>Stossel v. Meta Platforms, Inc.</i> , 634 F. Supp. 3d 743 (N.D. Cal. 2022)	15
<i>Trans Union Corp. v. FTC</i> , 245 F.3d 809 (D.C. Cir. 2001)	30
<i>Trans Union LLC v. FTC</i> , 295 F.3d 42 (D.C. Cir. 2002)	30
<i>Woodhull Freedom Foundation v. United States</i> , 72 F.4th 1286 (D.C. Cir. 2023).....	17
<i>Zauderer v. Office of Disciplinary Counsel</i> , 471 U.S. 626 (1985).....	<i>passim</i>

Statutes

Allow States and Victims to Fight Online Sex Trafficking Act/Stop Enabling Sex Traffickers Act (“FOSTA”)	17
Cal. Bus. & Prof. Code §§ 22675-22681.....	<i>passim</i>

Other Authorities

<i>A New Policy Against Self-Harm Blogs</i> , Tumblr (Feb. 23, 2012).....	6
A Note to Regulators	19, 20
<i>Acceptable Use Policy</i> , Strava.....	8
<i>Adult Nudity and Sexual Activity</i> , Facebook	4
Ángel Díaz and Laura Hecht-Felella, <i>Double Standards in Social Media Content Moderation</i> (Aug. 4, 2021)	29
Bennett Cyphers & Cory Doctorow, <i>The New ACCESS Act Is a Good Start. Here’s How to Make Sure It Delivers</i> , EFF Deeplinks (June 21, 2021).....	31
<i>Code of Conduct</i> , The High Road	9
<i>Comment on Evaluating the Competitive Effects of Corporate Acquisitions and Mergers</i> , EFF(August 20, 2018).....	31

<i>Community Standards Enforcement Report: Bullying and Harassment</i> , Meta.....	11
<i>Community Standards Enforcement Report: Hate Speech</i> , Meta	11
<i>Compare Community Guidelines</i> , Instagram	7
Corynne McSherry, et. al, <i>Privacy First: A Better Way to Address Online Harms</i> , EFF (Nov. 14, 2023)	31
Danielle Blunt et al., <i>Posting Into The Void</i> , Hacking//Hustling (2020)	17
<i>Defining the Public Interest</i> , X Help Center	6
<i>Digital Well-being</i> , TikTok	6
Eric Goldman, <i>Content Moderation Remedies</i> , 28 Mich. Tech. L. Rev. (2021).....	6
<i>Facebook Treats Punk Rockers Like Crazy Conspiracy Theorists, Kicks Them Offline</i> , EFF	27
<i>GodTube Community Guidelines</i> , GodTube	9
Hannah Bloch-Wehba, <i>Automation in Moderation</i> , 53 Cornell Int’l L.J. (2020)	3
Hannah Denham, <i>Another Fake Video of Pelosi Goes Viral on Facebook</i> , Wash. Post (Aug. 3, 2020)	9
Hibaq Farah, <i>Pro-Palestinian Instagram account locked by Meta for ‘security reasons’</i> , The Guardian (Oct. 26, 2023)	28
<i>Instagram Stands Against Online Bullying</i> , Instagram	6
Jack Shafer, <i>The Op-Ed Page’s Back Pages: A Press Scholar Explains How the New York Times Op-Ed Page Got Started</i> , Slate (Sept. 27, 2010)	12
James Grimmelman, <i>The Virtues of Moderation</i> , 17 Yale J. of Law & Tech (2015)	5
Jillian C. York & David Greene, <i>How to Put COVID-19 Content Moderation Into Context</i> , Brookings’ TechStream (May 21, 2020).....	3
Jillian C. York, <i>Silicon Values: The Future of Free Speech Under Surveillance Capitalism</i> , Verso, (2021).....	26
Karen Gullo & Jillian C. York, <i>Platforms Must Stop Unjustified Takedowns of Posts By and About Palestinians</i> , EFF Deeplinks (Nov. 8, 2023).....	28
Kate Crawford & Tarleton Gillespie, <i>What Is a Flag For? Social Media Reporting Tools and the Vocabulary of Complaint</i> , 18 New Media & Soc’y (2014)	11
Kate Klonick, <i>The New Governors: The People, Rules, And Processes Governing Online Speech</i> , 131 Harv. L. Rev. (2018).....	5, 6, 11

Kevin Anderson, <i>YouTube Suspends Egyptian Blog Activist’s Account</i> , Guardian (Nov. 28, 2007)	26
Malachy Browne, <i>YouTube Removes Videos Showing Atrocities in Syria</i> , N.Y. Times (Aug. 22, 2017)	28
Michael J. Socolow, <i>A Profitable Public Sphere: The Creation of the New York Times Op-Ed Page</i> , Commc’n & Journalism Fac. Scholarship (2010)	12
Mike Masnick, <i>Masnick’s Impossibility Theorem: Content Moderation At Scale Is Impossible To Do Well</i> , Techdirt (Nov. 20, 2019).....	27
<i>Moderator Code of Conduct</i> , Reddit	10
Olivia-Anne Cleary, <i>Facebook Has Banned Awareness Posts That Include the Words ‘Period,’ ‘Vulva’ and ‘Clitoris’ for Being Too Sexual</i> , Glamour (Jun. 9, 2023)	28
Op-Ed, Wikipedia.....	12
<i>Pennysaver</i> , Wikipedia.....	12
<i>Platform Policies</i> , Quora	8
<i>Policy: Terms of Use</i> , Wikimedia	10
Quora	7
<i>Reddiquette</i> , Reddit	10
<i>Role of Administrators and Moderators on Discord</i> , Discord	10
Ryan Mac, <i>Instagram Censored Posts About One of Islam’s Holiest Mosques, Drawing Employee Ire</i> , BuzzFeed News (May 12, 2021).....	18
<i>Sensitive Media Policy</i> , X (January 2024)	7
Seny Kamara et al., <i>Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems</i> , Ctr. for Democracy & Tech. (2021).....	4, 5
<i>Strava Community Standards</i> , Strava.....	8
<i>Terms and Conditions</i> , Shabbat.com.....	9
The Santa Clara Principles	18, 19
<i>Trust and Safety Fundamentals</i> , “Policy Development: What is Policy and Why Does it Matter,” Trust & Safety Professionals Association.....	4, 5, 7
<i>Violent and Graphic Content</i> , Meta Transparency Center	6
<i>Website Terms and Conditions of Use and Agency Agreement</i> , Rumble.....	9
<i>YouTube Community Guidelines Enforcement</i> , Google.....	11

STATEMENT OF INTEREST¹

Amicus curiae submits this brief to provide additional information and argument on two issues raised by this appeal that are within *amicus*'s field of expertise: the nature of content moderation as practiced by social media and other interactive websites as constitutionally protected editorial and curatorial expression; and, the weighty First Amendment concerns raised by governmental interference with such curatorial expression. *Amicus curiae* does so from the perspective of the Internet users who read and contribute to such sites, rather than the sites themselves as represented by plaintiff.

The Electronic Frontier Foundation (EFF) is a member-supported, nonprofit civil liberties organization that has worked for over 30 years to protect free speech, privacy, security, and innovation in the digital world. With approximately 30,000 members, EFF represents the interests of technology users in court cases and broader policy debates surrounding the application of law to the Internet and other technologies. EFF was a coauthor of the Santa Clara Principles on Transparency and Accountability in Content Moderation and frequently files briefs in cases addressing online intermediary content moderation, including in this Court in

¹ No counsel for a party authored this brief in whole or in part, and no person other than *amicus* or its counsel has made any monetary contributions intended to fund the preparation or submission of this brief. The parties have consented to the filing of this brief.

Prager Univ. v. Google LLC, 951 F.3d 991 (9th Cir. 2020), and *O’Handley v. Weber*, 62 F.4th 1145 (9th Cir. 2023), and in the Eleventh Circuit in *NetChoice LLC v. Attorney General, Florida*, 34 F.4th 1196 (11th Cir. 2022).

INTRODUCTION

The TOS Report and penalty provisions of AB587² are an unconstitutional intrusion into the editorial freedom that all publishers, including social media platforms and other websites, enjoy. Government interference into social media content moderation in any form raises serious First Amendment, and broader human rights, concerns. Courts must generally closely scrutinize such interventions, and laws such as AB587 that single out certain disfavored categories of protected speech for differential treatment must survive strict First Amendment scrutiny; The district court erred in applying less demanding scrutiny. This Court should reverse the district court’s dismissal.

ARGUMENT

I. THE CONTENT MODERATION SYSTEMS TARGETED BY THE LAW ARE EDITORIAL PROCESSES PROTECTED BY THE FIRST AMENDMENT.

As a threshold matter, this Court should affirm that the social media platforms and other websites subject to AB587 have the First Amendment right to curate the speech of others that they publish on their sites.

² Assembly Bill 587, codified at Cal. Bus. & Prof. Code §§ 22675-22681.

A. Content Moderation Is an Inherently Expressive Editorial Process.

1. Content Moderation Bears the Hallmarks of a Typical Editorial Process.

Social media platforms, at least from their point of mass adoption, have rarely published all legal speech submitted to their sites. Instead, they engage in an editorial or curatorial process commonly known as content moderation: the use of policies, systems, and tools to decide what user-generated content or accounts to publish, remove, amplify, or manage.³ Large-scale, outsourced content moderation first emerged in the early 2000s.⁴

Platforms moderate content in phases that are typical of many editorial processes: they define permissible and impermissible content; detect content that may violate their policies or the law; evaluate that content to determine whether it in fact violates their policies or the law; take an enforcement action against violative content; allow users to appeal or otherwise seek review of content moderation decisions that they believe are erroneous; and educate users about

³ See Hannah Bloch-Wehba, *Automation in Moderation*, 53 Cornell Int'l L.J. 41, 42, 48 (2020).

⁴ Jillian C. York & David Greene, *How to Put COVID-19 Content Moderation Into Context*, Brookings' TechStream (May 21, 2020), <https://www.brookings.edu/articles/how-to-put-covid-19-content-moderation-into-context/>.

content moderation policies and their enforcement.⁵ In each phase, platforms make editorial judgments about what content they wish to allow or forbid on their services, or how to display or arrange it.

During the definitional phase, some platforms develop a content policy, *i.e.*, a set of rules about what content is and is not allowed on their platforms, speech that a platform believes is detrimental to its users and the public, its business interests, its editorial preferences, or all of these.⁶ These policies typically address constitutionally protected speech: for example, many platforms ban legal, non-obscene sexual content, even though such speech enjoys First Amendment protection, *see Miller v. California*, 413 U.S. 15 (1973).⁷ Content moderation is an iterative process: platforms may have significant internal discussion and debate, conduct internal and external research, and continuously reformulate their content

⁵ Seny Kamara et al., *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*, Ctr. for Democracy & Tech. 9–11 (2021), <https://cdt.org/wp-content/uploads/2021/08/CDT-Outside-Looking-In-Approaches-to-Content-Moderation-in-End-to-End-Encrypted-Systems-updated-20220113.pdf>.

⁶ *Id.* at 9; *Trust and Safety Fundamentals*, “Policy Development: What is Policy and Why Does it Matter,” Trust & Safety Professionals Association, <https://www.tspa.org/curriculum/ts-fundamentals/policy/policy-development/> (last visited Feb. 20, 2024).

⁷ *See, e.g., Adult Nudity and Sexual Activity*, Facebook, <https://transparency.fb.com/policies/community-standards/adult-nudity-sexual-activity/> (last visited Feb. 20, 2024).

policies.⁸

Many platforms publish some portion of their content policies, typically as “community standards,” but others do not, in order to maintain flexibility and for other reasons.⁹ Some sites may publish a high-level policy that is supplemented and fleshed out by constantly evolving internal memoranda, enforcement guidelines, directions, decision trees, policy guidance, and supporting information.¹⁰ Smaller platforms might not adopt formal policies at all—particularly, for example, platforms like blogs, where the small community size could make it manageable for a single blogger to moderate the platform via ad hoc decision-making.¹¹

The evaluation phase often involves deliberation and debate to determine whether particular content violates its policies. Some platforms detect potentially

⁸ See Kate Klonick, *The New Governors: The People, Rules, And Processes Governing Online Speech*, 131 Harv. L. Rev. 1598, 1631-35 (2018); Kamara, *supra* n.4, at 8.

⁹ *Trust and Safety Fundamentals*, “Policy Development: What is Policy and Why Does it Matter,” *supra* n.5.

¹⁰ *Id.*

¹¹ See James Grimmelman, *The Virtues of Moderation*, 17 Yale J. of Law & Tech 42, 73 (2015) (“[T]he larger a community is, the better it is at competing with external alternatives, but the more internal moderation it requires As a community grows, it becomes easier for individuals and groups to resist a norm. This breakdown makes it harder to use social norms to moderate large communities. A group of twenty can operate by unspoken consensus in a way that a group of twenty thousand cannot.”).

violating content only after it is posted; others screen some or all content ex ante.¹²

In the enforcement phase, a platform must decide what, if any, action to take with respect to violative content. That is not a binary decision about whether to take down content or allow it to remain on a service, but also includes whether to change the manner or place in which content is displayed or to add the platform's own affirmative speech.¹³ For example, depending on the nature of the violative content, a platform might choose to add its own content warning, public service announcement, health or safety resources, or "fact-checking" information to provide context and support for its users.¹⁴

Content moderation is typically a continually evolving and iterative process

¹² Klonick, *supra* n.7, at 1635.

¹³ See Eric Goldman, *Content Moderation Remedies*, 28 Mich. Tech. L. Rev. 1, 23–39 (2021) (describing various enforcement options).

¹⁴ *Id.* at 26-27, 30-31. See also, e.g., *Instagram Stands Against Online Bullying*, Instagram, <https://about.instagram.com/community/anti-bullying> (last visited Feb. 20, 2024) (warnings before posting "potentially offensive" comments); *Digital Well-being*, TikTok, <https://www.tiktok.com/safety/en-gb/well-being/> (last visited Feb. 20, 2024) (digital well-being and media literacy resource guide); *Violent and Graphic Content*, Meta Transparency Center, <https://transparency.fb.com/policies/community-standards/violent-graphic-content/> (last visited Feb. 20, 2024) (warning labels for graphic content); *Defining the Public Interest*, X Help Center, <https://help.twitter.com/en/rules-and-policies/public-interest> (last visited Feb. 20, 2024) (public click-through notice added to posts by government official that violate content policies and would otherwise be taken down, but are left up under X's public interest exception); *A New Policy Against Self-Harm Blogs*, Tumblr (Feb. 23, 2012), <https://staff.tumblr.com/post/18132624829/self-harm-blogs> (last visited Feb 20, 2024) (proposal for adding PSAs to self-harm-related search results).

whereby the evaluation and enforcement phases in one case feeds into a site's reevaluation of its rules and process.¹⁵

2. Social Media Platforms Have Rules, Standards and Guidelines about What Content They Want and Don't Want on Their Sites.

There is wide variety in these editorial processes.

Content moderation policies differ from platform to platform, with some choosing to address the disfavored “categories of content” set forth in § 22677(a)(3), among many other categories, and addressing them in starkly different ways.¹⁶

Quora, a site where any user can ask questions “and get answers from people who have been there and done that” and which sees itself as a place to “take refuge from misinformation and incendiary arguments,”¹⁷ prohibits “posting irrelevant answers or comments”; “targeted insults or profanity directed at private individuals based on personal attributes, such as physical appearance”; “hate speech,” including “content that dehumanizes or calls for violence, exclusion, or segregation

¹⁵ *Trust and Safety Fundamentals*, “Why Does Policy Change,” *supra* n.8.

¹⁶ *Compare Community Guidelines*, Instagram, <https://help.instagram.com/477434105621119> (last visited Feb. 20, 2024) (prohibiting nudity except in the context of breastfeeding, birth-related moments, health-related situations, in paintings or sculptures, or as an act of protest), *with Sensitive Media Policy*, X (January 2024), <https://help.twitter.com/en/rules-and-policies/media-policy> (permitting “consensually produced and distributed” adult nudity and sexual behavior) (last visited Feb. 20, 2024).

¹⁷ Quora, <https://www.quora.com/about> (last visited Feb. 20, 2024).

of protected classes” and Holocaust and Armenian Genocide denial; and “encouraging, glorifying, or promoting” harmful activities such as “self-harm (including eating disorders) and animal cruelty.”¹⁸

Strava, a social media platform for athletes, has Community Standards that prohibit the posting of content that is “harassing, abusive, or hateful, discriminatory, or that advocates violence.”¹⁹ One of Strava’s main features is for cyclists and runners to share their routes, called “segments”; but Strava’s Community Standards allow only “good segments” created with “common sense.”²⁰ The Community Standards also require all users to be “inclusive and anti-racist.”²¹

Rumble, a video sharing alternative to YouTube, prohibits both videos and comments on a number of criteria, including a bar on content that “Promotes, supports or incites individuals and/or groups which engage in violence or unlawful acts, including but not limited to Antifa groups and persons affiliated with Antifa, the KKK and white supremacist groups and or persons affiliated with these

¹⁸ *Platform Policies*, Quora, <https://help.quora.com/hc/en-us/articles/360000470706-Platform-Policies> (last updated Jun. 2023).

¹⁹ *Acceptable Use Policy*, Strava, https://www.strava.com/legal/acceptable_use (updated Feb. 15, 2024) (last visited Feb. 20, 2024).

²⁰ *Strava Community Standards*, Strava, <https://www.strava.com/community-standards> (last visited Feb. 20, 2024).

²¹ *Id.*

groups.”²²

The policies of many sites reflect their niche focuses. The High Road, a firearms discussion platform, requires that all posts be “related to firearms or ‘Right to Keep and Bear Arms’ (RKBA) issues” and explicitly prohibits users from engaging in “discussions relating to the preparation for possible societal breakdown” or “foreign invasion.”²³ GodTube, a Christian video platform, prohibits users from promoting any “beliefs or teachings contrary to those of Christianity as articulated by the historic creeds...”²⁴; while Shabbat.com, “the world’s largest Jewish social network,” prohibits users from posting any content that is “missionizing, Christian or otherwise.”²⁵

Implementation also varies widely. Platforms make different judgment calls about whether particular content violates their content policies, even if those policies are similar.²⁶ Some platforms structure their services so that some content

²² *Website Terms and Conditions of Use and Agency Agreement*, Rumble, <https://rumble.com/s/terms> (last visited Feb. 20, 2024).

²³ *Code of Conduct*, The High Road, <https://www.thehighroad.org/index.php?pages/code-of-conduct/> (last visited Feb. 20, 2024).

²⁴ *GodTube Community Guidelines*, GodTube, <https://www.godtube.com/terms-of-use.html> (last visited Feb. 20, 2024).

²⁵ *Terms and Conditions*, Shabbat.com, <https://www.shabbat.com/terms> (last visited Feb. 20, 2024).

²⁶ See, e.g., Hannah Denham, *Another Fake Video of Pelosi Goes Viral on Facebook*, Wash. Post (Aug. 3, 2020),

moderation decisions are made by a broader community of users. Reddit and Discord rely on certain users to moderate content through the practice of “community moderation.”²⁷ Reddit users manage and create thousands of communities, called “subreddits.” Although Reddit has an overriding content policy, a moderator makes the decisions within each community as guided by Reddit’s “Moderator Code of Conduct.”²⁸ Discord employs a similar model.²⁹ Each site thereby empowers some users to remove and down-rank other users’ speech if that speech is against the community’s rules.³⁰

Most platforms, even those which do not employ community moderation, allow users to report or “flag” content they believe violates the platforms’ rules or

<https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fake-video-facebook/> (reporting that TikTok, Twitter and YouTube removed a doctored video of Rep. Nancy Pelosi, while Facebook allowed it to remain with a label).

²⁷ See *Moderator Code of Conduct*, Reddit, <https://www.redditinc.com/policies/moderator-code-of-conduct> (effective Sept. 25, 2023) (last visited Feb. 20, 2024); *Role of Administrators and Moderators on Discord*, Discord, <https://discord.com/safety/360044103531-role-of-administrators-and-moderators-on-discord> (last visited Feb. 20, 2024); see also *Policy: Terms of Use*, Wikimedia, https://foundation.wikimedia.org/wiki/Policy:Terms_of_Use (last visited Feb. 20, 2024).

²⁸ Reddit, *supra* n.26.

²⁹ Discord, *supra* n.26.

³⁰ See, e.g., *Reddiquette*, Reddit, <https://reddit.zendesk.com/hc/en-us/articles/205926439-Reddiquette> (last visited Feb. 20, 2024).

standards.³¹ Indeed, such flags may account for a large amount of content moderation actions. In the third quarter of 2023, YouTube removed 8.1 million videos, 322,513 of which were flagged by users.³² During that same period, 12.2% of the 8.3 million posts Facebook actioned for bullying and harassment was reported by users, while 5.2% of the 9.6 million posts they actioned for hate speech was reported by users.³³

Some platforms allow users to appeal content moderation decisions, while others do not.³⁴

3. Social Media Follows A Long Line of Traditional Media That Have Curated Content Written By Others

The typical newspaper, the editorial rights of which are unquestioned, is, like social media, a mix of original writing and content created by others. For newspapers, this includes syndicated and wire service articles, advertisements,

³¹ See generally Kate Crawford & Tarleton Gillespie, *What Is a Flag For? Social Media Reporting Tools and the Vocabulary of Complaint*, 18 *New Media & Soc'y* 410 (2014).

³² *YouTube Community Guidelines Enforcement*, Google, <https://transparencyreport.google.com/youtube-policy/removals> (last visited Feb. 20, 2024).

³³ *Community Standards Enforcement Report: Bullying and Harassment*, Meta, <https://transparency.fb.com/reports/community-standards-enforcement/bullying-and-harassment/facebook/> (last visited Feb. 20, 2024); *Community Standards Enforcement Report: Hate Speech*, Meta, <https://transparency.fb.com/reports/community-standards-enforcement/hate-speech/facebook/> (last visited Feb. 20, 2024).

³⁴ Klonick, *supra* n.7, at 1648.

wedding, engagement, and birth announcements, and comics. Opinion pages in particular typically publish a lot of content created by others: opinion pieces, letters to the editor, syndicated editorial cartoons and columns.³⁵

And also like social media sites, print news media vary widely in how selective they are with external content. Pennysavers, for example, local newspapers either entirely or primarily composed of classified advertisements, coupons, life milestone announcements, congratulatory messages, recipes, public notices, and the like, have a long and storied history as a relatively non-selective print news publication.³⁶

B. Content Moderation is Protected by the First Amendment Just as Other Editorial Process Are.

The First Amendment protects the rights of social media services to publish both user speech and their own speech, regardless of whether they curate user

³⁵ See Jack Shafer, *The Op-Ed Page's Back Pages: A Press Scholar Explains How the New York Times Op-Ed Page Got Started*, Slate (Sept. 27, 2010), <https://slate.com/news-and-politics/2010/09/a-press-scholar-explains-how-the-new-york-times-op-ed-page-got-started.html> (describing how the pages opposite newspapers' editorial pages became a forum for outside contributors to express views different from those expressed by the paper's editorial board); Michael J. Socolow, *A Profitable Public Sphere: The Creation of the New York Times Op-Ed Page*, Comm'n & Journalism Fac. Scholarship (2010), https://digitalcommons.library.umaine.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1001&context=cmj_facpub; Op-Ed, Wikipedia, <https://en.wikipedia.org/wiki/Op-ed> (last visited Feb. 20, 2024).

³⁶ *Pennysaver*, Wikipedia, <https://en.wikipedia.org/wiki/Pennysaver> (last visited Sept. 21, 2023).

speech a lot, not at all, or anything in between.

Although this specific issue is currently before the Supreme Court in the *Netchoice* cases,³⁷ the Court has long held that private publishers have a First Amendment right to control the content of their publications, and specifically whether and how to publish things written by others. *See Miami Herald Publishing Co. v. Tornillo*, 418 U.S. 241, 254-55 (1974). *Cf. Manhattan Community Access Corp. v. Halleck*, 139 S. Ct. 1921, 1930 (2019) (reaffirming that “when a private entity provides a forum for speech,” “[t]he private entity may . . . exercise editorial discretion over the speech and speakers in the forum”). *See also Los Angeles v. Preferred Communications, Inc.*, 476 U.S. 488, 494 (1986) (recognizing cable television providers’ First Amendment right to “exercis[e] editorial discretion over which stations or programs to include in [their] repertoire[s]”).

Relevant here, editorial freedom is not held only by publishers who primarily publish their own speech. It also applies to any entity that speaks by curating and publishing the speech of others. The Supreme Court found that the organizers of a parade had a First Amendment right to curate the participants, and thus could not be required to include a certain message, even if the parade was perceived as generally open for public participation, and such selectivity was not

³⁷ *Moody v. NetChoice, LLC*, 216 L. Ed. 2d 1313 (Sept. 29, 2023); *NetChoice, LLC v. Paxton*, 216 L. Ed. 2d 1313 (Sept. 29, 2023).

generally evident in the past. *Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. Of Boston*, 515 U.S. 557, 569–70 (1995). As the *Hurley* Court explained, “a private speaker does not forfeit constitutional protection simply by combining multifarious voices, or by failing to edit their themes to isolate an exact message as the exclusive subject matter of the speech. Nor, under our precedent, does First Amendment protection require a speaker to generate, as an original matter, each item featured in the communication.” *Id.*

Indeed, the Supreme Court’s most powerful pronouncement of freedom of the press, *New York Times Co. v. Sullivan*, 376 U.S. 254 (1964), centered on *The Times* publishing someone else’s unsolicited content, a paid advertisement. The Court found that *The Times*’ role as a host for the speech of others was critical to its decision: newspapers are “an important outlet for the promulgation of information and ideas by persons who do not themselves have access to publishing facilities—who wish to exercise their freedom of speech even though they are not members of the press.” *Id.* at 266.

The Supreme Court consistently applies this general principle of editorial and curatorial freedom in a variety of speech contexts, including online intermediaries; just last term, the Court reaffirmed that the state cannot “alter the expressive content of” or “interfere with” a website designer’s desired message. *303 Creative LLC v. Elenis*, 600 U.S. 570, 596 (2023) (internal quotation marks

and citations omitted).

There is thus no reason to have a different rule for social media, which the Supreme Court has recognized as providing “perhaps the most powerful mechanisms available to a private citizen to make his or her voice heard.”

Packingham v. North Carolina, 137 S. Ct. 1730, 1737 (2017). And indeed, courts have consistently recognized the First Amendment right to editorial freedom for social media platforms that primarily, if not exclusively, publish user-generated content, including numerous trial courts within this circuit.³⁸ As the Eleventh Circuit found in one of the *Netchoice* cases currently under review by the Supreme Court:

Social-media platforms exercise editorial judgment that is inherently expressive. When platforms choose to remove users or posts, deprioritize content in viewers’ feeds or search results, or sanction breaches of their community standards, they engage in First-Amendment-protected activity. Social-media platforms’ content-moderation decisions are, we think, closely analogous to the editorial judgments that the Supreme Court recognized in *Miami Herald*, *Pacific Gas*, *Turner*, and *Hurley*. . . . [¶] All such decisions about what speech to permit, disseminate, prohibit, and

³⁸ See, e.g., *Children’s Health Def. v. Facebook Inc.*, 546 F. Supp. 3d 909, 922-23 (N.D. Cal. 2021), *appeal docketed*, No. 21-1620 (9th Cir. July 21, 2021); *Huber v. Biden*, No. 21-CV-06580-EMC, 2022 WL 827248, at *6 (N.D. Cal. Mar. 18, 2022), *aff’d*, No. 22-15443, 2022 WL 17818543 (9th Cir. Dec. 20, 2022); *O’Handley v. Padilla*, 579 F. Supp. 3d 1163, 1186 (N.D. Cal. 2022), *aff’d sub nom. O’Handley v. Weber*, 62 F.4th 1145 (9th Cir. 2023); *Stossel v. Meta Platforms, Inc.*, 634 F. Supp. 3d 743, 760 (N.D. Cal. 2022). This Court has similarly rejected the argument that social media platforms are state actors that are limited by the First Amendment in their ability to select the speech of others. *Prager Univ. v. Google LLC*, 951 F.3d 991, 995 (9th Cir. 2020).

deprioritize—decisions based on platforms’ own particular values and views—fit comfortably within the Supreme Court’s editorial-judgment precedents.

NetChoice LLC v. Attorney General, Florida, 34 F.4th 1196, 1213-14 (11th Cir. 2022), *cert. granted sub nom, Moody v. NetChoice, LLC*, 216 L. Ed. 2d 1313 (Sept. 29, 2023).

This Court should not follow the contrary decision of the Fifth Circuit in *NetChoice, LLC v Paxton*, which quizzically saw the Supreme Court’s First Amendment jurisprudence as secondary to what the court divined as “the original public meaning of the First Amendment.” 49 F.4th 439, 454-55 (5th Cir. 2022), *cert. granted sub nom, NetChoice, LLC v. Paxton*, 216 L. Ed. 2d 1313 (Sept. 29, 2023). The Fifth Circuit, analyzing whether Texas could compel large social media platforms to publish users’ posts regardless of the user’s viewpoints, wrongly saw social media platforms as mere “conduits” for user speech that, “[u]nlike newspapers, . . . exercise virtually no editorial control or judgment.” *Id.* at 459-60 (quoting *Tornillo*, 418 U.S. at 258). This is inconsistent with the very reason for AB587: to learn more about platforms’ “editorial control or judgment” and guide such judgment with respect to the disfavored categories of speech in section 22677(a)(3).

C. Government Involvement in Content Moderation Raises Serious First Amendment Concerns Because Content Moderation Is Particularly Vulnerable to Government-Prompted Chilling Effects.

Because governments have outsized influence to manipulate content moderation systems for their own political goals, courts must be steadfast in applying appropriately demanding First Amendment scrutiny to all government interference in content moderation.

Platforms have often aggressively over-moderated user speech that was not prohibited by their content policies when attempting to minimize legal or reputational risks arising from government regulation or criticism. For example, many platforms responded to the enactment of the Allow States and Victims to Fight Online Sex Trafficking Act/Stop Enabling Sex Traffickers Act (“FOSTA”) by removing content by sex workers and sex worker advocates that FOSTA did not actually prohibit.³⁹ Government pressure to remove terrorist content from platforms has also led to over-removals of speech. For example, in 2021, Instagram removed posts about one of Islam’s holiest mosques, Al-Aqsa, because its name is contained within the name of an organization the company had designated as a

³⁹ See Danielle Blunt et al., *Posting Into The Void*, Hacking//Hustling (2020), <https://hackinghustling.org/wp-content/uploads/2020/09/Posting-Into-the-Void.pdf>; *Woodhull Freedom Foundation v. United States*, 72 F.4th 1286, 1299-1305 (D.C. Cir. 2023) (describing FOSTA’s very limited prohibitions).

terrorist group.⁴⁰

Government involvement in private companies' content moderation processes, regardless of form or degree, raises serious human rights concerns. The revised Santa Clara Principles, of which *amicus curiae* Electronic Frontier Foundation is a co-author, specifically scrutinize "State Involvement in Content Moderation," and affirm that "state actors must not exploit or manipulate companies' content moderation systems to censor dissenters, political opponents, social movements, or any person."⁴¹ "Special concerns are raised by demands and requests from state actors (including government bodies, regulatory authorities, law enforcement agencies and courts) for the removal of content or the suspension of accounts."⁴²

And while the Santa Clara Principles include standards for transparency around content moderation processes that may seem similar to those mandated by AB587, the standards are expressly *not* to be mandated by government. The Santa Clara Principles specifically state, "This second iteration of the Santa Clara Principles has been developed to support companies to comply with their

⁴⁰ Ryan Mac, *Instagram Censored Posts About One of Islam's Holiest Mosques, Drawing Employee Ire*, BuzzFeed News (May 12, 2021), <https://www.buzzfeednews.com/article/ryanmac/instagram-facebook-censored-al-aqsa-mosque>.

⁴¹ The Santa Clara Principles, <https://santaclaraprinciples.org> (last visited Feb. 20, 2024).

⁴² *Id.*

responsibilities to respect human rights and enhance their accountability, and to assist human rights advocates in their work. **They are not designed to provide a template for regulation.**⁴³ In a Note to Regulators, the Principles explain that its standards do not readily scale or account for the variations among online services:

The Santa Clara Principles seeks to set standards. Some services will appropriately meet these standards. Some will appropriately meet only some of them, while others will and should exceed them. Where any particular service falls will depend on many factors—number of users, capitalization, age, focus of service, editorial priorities, user priorities—that will vary from service to service. While companies should design their services with due process in mind from the beginning, companies must have some flexibility as to how they implement the Santa Clara Principles, from their inception, and then evolving over time as the service matures. The Santa Clara Principles are thus best seen as touchstones against which any company’s practices can be evaluated and compared, not as dictates.

To maintain this necessary flexibility, governments should resist legal mandates that would be prohibitively expensive or practically impossible to meet. Such mandates discourage new entrants into the field and thus discourage innovation and competition. Even among well-established services, there are no metrics that readily correspond to a required level of compliance.⁴⁴

The Principles also discuss other obstacles to employing them as governmental mandates: the potential for political exploitation, the variation among regional and national legal systems that govern these inherently

⁴³ *Id.* (emphasis in original).

⁴⁴ A Note to Regulators, <https://santaclaraprinciples.org/regulators/> (last visited Feb. 20, 2024).

international services, and the constantly evolving landscape of available services.⁴⁵

Moreover, the Santa Clara Principles focus on providing information to users about why their own speech might be moderated. AB587, in contrast, focuses on providing information to the state itself about why certain user speech was *not moderated* — part of a larger scheme by the state to influence what user content is published.

II. AB587’S SPEECH COMPULSIONS MUST BE SUBJECT TO STRICT SCRUTINY, NOT THE MORE PERMISSIVE *ZAUDERER* STANDARD.

Amicus agrees with the arguments presented by X that the speech compelled by AB587, as a content-based restriction on speech, must be subject to strict scrutiny, but writes separately to emphasize how the compelled commercial speech doctrine cannot be applied per se to compulsions that intrude upon the protected editorial process described above.

A. Speech That Is Part of and Related to the Editorial Process Is Not Commercial Speech Nor Otherwise Subject to Diminished First Amendment Protection

The very speech AB587 compels, details about editorial decision making, is not commercial speech and thus is not subject to the diminished First Amendment protections of either *Zauderer v. Office of Disciplinary Counsel*, 471 U.S. 626

⁴⁵ *Id.* (emphasis in original).

(1985) or *Central Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n of New York*, 447 U.S. 557 (1980).⁴⁶

Social media sites, like newspapers, “are in the business of expression.” *City of Lakewood v. Plain Dealer Publishing Co.*, 486 U.S. 750, 761 (1988). They can thus not be regulated as commercial enterprises like “soda vendors.” *Id.*

It does not matter that some social media platforms, though not necessarily all, monetize the speech they publish: “the degree of First Amendment protection is not diminished merely because the newspaper or speech is sold, rather than given away.” *Id.* at 756 n.5 “If a newspaper's profit motive were determinative, all aspects of its operations—from the selection of news stories to the choice of editorial position—would be subject to regulation if it could be established that they were conducted with a view toward increased sales. Such a basis for regulation clearly would be incompatible with the First Amendment.” *Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations*, 413 U.S. 376, 385, (1973).

⁴⁶ Although courts occasionally refer to *Zauderer* scrutiny as akin to “rational basis” scrutiny, *see, e.g., American Beverage Association v. City & County of San Francisco*, 916 F.3d 749, 767 (9th Cir. 2019) (Nguyen, J., concurring), *Zauderer* is actually more demanding. While rational basis scrutiny is satisfied if the restriction on speech is rationally related to a legitimate governmental purpose, *Stormans, Inc. v. Weisman*, 794 F.3d 1064, 1084 (9th Cir. 2015), *Zauderer* requires both a stronger governmental interest—“substantial,” the strength typically associated with intermediate scrutiny—and tailoring of means to achieve that interest to avoid undue burden on the speaker. *See Nat'l Ass'n of Wheat Growers v. Bonta*, 85 F.4th 1263, 1275 (9th Cir. 2023).

See also New York Times, 376 U.S. at 266 (“That the Times was paid for publishing the advertisement is as immaterial in this connection as is the fact that newspapers and books are sold.”); *Smith v. California*, 361 U.S. 147, 149–50 (1959) (“It is, of course, no matter that the dissemination” of “books and other forms of the printed word” “takes place under commercial auspices.”)

Nor is an editorial publication’s First Amendment protection diminished simply because it includes advertisements. *See City of Cincinnati v. Discovery Network, Inc.*, 507 U.S. 410, 423 (1993).

This Court has thus found that “publications like yellow pages directories and newspapers receive full First Amendment protection not only because their content is somehow inextricably intertwined, but because, as a threshold matter, they do not constitute commercial speech” under the Supreme Court’s tests in the first instance. *Dex Media W., Inc. v. City of Seattle*, 696 F.3d 952, 962 (9th Cir. 2012).

B. *Zauderer* Does Not Apply to Intrusions on Editorial Privacy

Moreover, applying *Zauderer* to compelled disclosure of editorial decision-making also conflicts with the long-standing commitment of this Court, and others, to editorial privacy.

Editorial freedom is not simply the right to devise and implement an editorial policy, but also protections for privacy in doing so. Intrusions upon it are

subject to exacting First Amendment scrutiny. This Court has thus overturned on First Amendment grounds a grand jury contempt order that sought to compel disclosure of editorial decision-makers and their motives. *See Bursey v. United States*, 466 F.2d 1059, 1085-86 (9th Cir. 1972), *superseded by statute in non-pertinent part*, *In re Grand Jury Proceedings*, 863 F.2d 667, 669 (9th Cir. 1988). Similarly, in *Application of Consumers Union of U. S., Inc.*, the district court quashed third-party subpoenas that sought the publisher’s research methodology and “procedure by which he formed his conclusions.” 495 F. Supp. 582, 585 (S.D.N.Y. 1980). The court explained that discovery seeking “to examine the reportorial and editorial processes . . . would represent a substantial intrusion on fact gathering and editorial privacy which are significant aspects of a free press.” *Id.* at 586.⁴⁷ And although the Supreme Court declined to create an absolute First Amendment evidentiary privilege when the publisher is the defendant and “there is a specific claim of injury arising from a publication that is alleged to have been knowingly or recklessly false,” it stated that a “law that subjects the editorial process to private or official examination merely to satisfy curiosity or to serve some general end such as the public interest . . . would not survive constitutional scrutiny” under the First Amendment, a standard more searching than *Zauderer*.

⁴⁷ *See also In re Consumers Union of the U.S., Inc.*, 32 Fed. R. Serv. 2d 1373 (S.D.N.Y. 1981); *Kenneally v. Suzuki Motor Co., Ltd.*, No. M-8-85, 1994 WL 48840 (S.D.N.Y. Feb. 10, 1994).

Herbert v. Lando, 441 U.S. 153, 174 (1974).⁴⁸

C. Speech Compulsions that are Informal Censorship Schemes are Subject to Exacting Scrutiny

This Court must also apply strict scrutiny because, as X ably demonstrates, AB587 was intended to be and is operating as an informal censorship scheme to pressure online intermediaries to moderate user speech.

In *Bantam Books, Inc. v Sullivan*, the Supreme Court found that the First Amendment prohibited not only direct censorship demands but also “system[s] of informal censorship” aimed at speech intermediaries. 372 U.S. 58, 71 (1963). The Supreme Court found that “the threat of invoking legal sanctions and other means of coercion, persuasion, and intimidation” against book distributors was enough to violate the First Amendment rights of the publishers of the books that would have otherwise been distributed. *Id.* at 67.

Indeed, the state’s actions around AB587 present an even more compelling case than *Bantam Books*.

In *Bantam Books*, the Rhode Island commission did not have the ability to penalize the book distributor. Nevertheless, it was unconstitutional “informal censorship” of the books when the commission issued notices to book distributors

⁴⁸ Even for nonconfidential journalistic information, this Court requires that the information sought be otherwise unavailable, noncumulative, and “clearly relevant to an important issue in the case,” also a more demanding standard than *Zauderer*. See *Shoen v. Shoen*, 48 F.3d 412, 416 (9th Cir. 1995).

that “certain designated books,” published by plaintiffs, were “objectionable for sale,” and that it was the commission’s “duty to recommend to the Attorney General prosecution of purveyors of obscenity,” and circulated the notices to local police, who visited the distributor “to learn what action he had taken.” *Id.* at 61-63. Predictably, the distributors stopped selling the books. *Id.* at 64. The Supreme Court found that the publishers had a First Amendment remedy against the state commission, even though it was the distributor’s action that directly harmed the publishers’ sales, and the government did not actually seize any books or prosecute anyone. *Id.* at 64 n.6.

AB587, in contrast grants the attorney general punitive powers over online sites to directly enforce its requirements that websites disclose how and whether they will curate certain categories of disfavored speech and then act “consistently” with those representations.

D. *Zauderer* Does Not Apply Because Content Moderation Processes are Inherently Subjective and Nonfactual.

Even if the speech compelled by AB587 were commercial speech, *Zauderer* would not apply because speech compelled is not “factual and uncontroversial.” *Nat’l Ass’n of Wheat Growers v. Bonta*, 85 F.4th 1263, 1275 (9th Cir. 2023).

As discussed above, content moderation is not an exact science, or any type of science at all. Content policies are frequently amorphous and ever-evolving, and their implementation is highly subjective; this is not the type of objective, carved-

in-stone information typically considered “factual” under *Zauderer*.

Indeed, content moderation has long been a highly fraught and puzzling process.

In 2007, YouTube, only two years old at the time, shut down the account of Egyptian human rights activist Wael Abbas after receiving multiple reports that the account featured graphic videos of police brutality and torture.⁴⁹ YouTube’s community standards at the time stated that “[g]raphic or gratuitous violence is not allowed.”⁵⁰ Just one year before, Abbas became the first blogger to receive the Knight International Journalism Award.⁵¹

And government’s attempts to influence content moderation date back just as far: Abbas’s account was restored only after the U.S. State Department communicated with YouTube’s new owner, Google.⁵²

Content moderation remains highly subjective and controversial to this day—which is precisely why it is squarely in the bullseye of efforts to regulate various aspects of it. Even the largest and best-resourced social media companies

⁴⁹ Kevin Anderson, *YouTube Suspends Egyptian Blog Activist’s Account*, Guardian (Nov. 28 2007), <https://www.theguardian.com/news/blog/2007/nov/28/youtubesuspendsegyptianblog>.

⁵⁰ *Id.*

⁵¹ Jillian C. York, *Silicon Values: The Future of Free Speech Under Surveillance Capitalism* 25-27, Verso, (2021).

⁵² *Id.*

struggle with it, often to the frustration of both users and governments. Even when using precise rules or carefully articulated “community standards,” moderated platforms often struggle to draw workable lines between permitted and forbidden speech. They make close editorial calls about which reasonable minds might differ and inevitable mistakes. As is often said, content moderation at scale is impossible to do perfectly, and nearly impossible to do well.⁵³

Platforms thus make a variety of contentious and seemingly erroneous decisions every day, especially with politically charged categories similar to the ones specified in Section 22677(a)(3). In January 2021, Facebook’s updated policy to remove “harmful conspiracy theories” resulted in it disabling a punk rock band’s page because its name, Adrenochrome, is a chemical that was a central part of the QAnon conspiracy theory.⁵⁴ Facebook also recently faced renewed criticism over the negative moderation of posts containing words pertinent to feminine hygiene, including “vulva” and “endometriosis” as too sexual.⁵⁵ YouTube has removed

⁵³ See, e.g., Mike Masnick, *Masnick’s Impossibility Theorem: Content Moderation At Scale Is Impossible To Do Well*, Techdirt (Nov. 20, 2019), <https://www.techdirt.com/articles/20191111/23032743367/masnicks-impossibility-theorem-content-moderation-scale-is-impossible-to-do-well.shtml>.

⁵⁴ *Facebook Treats Punk Rockers Like Crazy Conspiracy Theorists, Kicks Them Offline*, EFF, <https://www.eff.org/takedowns/facebook-treats-punk-rockers-crazy-conspiracy-theorists-kicks-them-offline> (last visited Sept. 21, 2023).

⁵⁵ Olivia-Anne Cleary, *Facebook Has Banned Awareness Posts That Include the Words ‘Period,’ ‘Vulva’ and ‘Clitoris’ for Being Too Sexual*, Glamour (Jun. 9,

videos documenting atrocities in Syria and elsewhere under its graphic violence policy.⁵⁶

Palestinian rights advocates have complained of unfair and biased content moderation against their views before and during the current Israel-Gaza conflict, including the removal of the Palestinian flag emoji and the Let's Talk Palestine and eye.on.palestine accounts from Instagram.⁵⁷

This pressure to strictly enforce hot-button community standards tends to disproportionately silence already marginalized speakers. A 2021 report documented how the viewpoints of communities of color, women, LGBTQ+ persons, and religious minorities are subject to over-enforcement of social media community standards. Their posts are more frequently taken down en masse, whereas the posts of dominant communities are more commonly subject to subtler

2023), <https://www.glamourmagazine.co.uk/article/facebook-bans-bodyform-posts-with-words-period-vulva-clitoris>.

⁵⁶ Malachy Browne, *YouTube Removes Videos Showing Atrocities in Syria*, N.Y. Times (Aug. 22, 2017), <https://www.nytimes.com/2017/08/22/world/middleeast/syria-youtube-videos-isis.html>; Anderson, *supra* n.49.

⁵⁷ Hibaq Farah, *Pro-Palestinian Instagram account locked by Meta for 'security reasons'*, The Guardian (Oct. 26, 2023), <https://www.theguardian.com/technology/2023/oct/26/pro-palestinian-instagram-account-locked-by-meta-for-security-reasons>; Karen Gullo & Jillian C. York, *Platforms Must Stop Unjustified Takedowns of Posts By and About Palestinians*, EFF Deeplinks (Nov. 8, 2023), <https://www.eff.org/deeplinks/2023/11/platforms-must-stop-unjustified-takedowns-posts-and-about-palestinians>.

forms of content moderation, such as warning labels and temporary demonetization.⁵⁸

E. There is a Significant Public Interest in Promoting Content Moderation Transparency

Amicus agrees with X that AB587 is not sufficiently tailored to a significant governmental interest, and also agree that the state lacks a legitimate government interest in favoring some legal speech on social media over other, as it does in section 22677(a)(3). And as discussed above, governments' role in content moderation should be sharply limited.

But amici do believe that as a general matter there is a significant *public interest* in promoting transparency around social media content moderation so that, in the interests of due process, users have a better understanding of how their speech and the speech they seek to read is moderated. As set forth in the Santa Clara Principles, “Companies should ensure that human rights and due process considerations are integrated at all stages of the content moderation process, and should publish information outlining how this integration is made,” a value discussed further in the Notice principle.⁵⁹

⁵⁸ Ángel Díaz and Laura Hecht-Felella, *Double Standards in Social Media Content Moderation* (Aug. 4, 2021), <https://www.brennancenter.org/our-work/research-reports/double-standards-social-media-content-moderation>.

⁵⁹ <https://santaclaraprinciples.org/>.

III. GOVERNMENTS CAN LOOK TO NON-EDITORIAL REGULATIONS TO BENEFIT SOCIAL MEDIA USERS

Nothing in this brief should be read as broadly exempting online services from regulation. Regulatory measures that do not target the editorial process or are not enacted in retaliation against or to provide levers against disfavored editorial policies and decisions may be acceptable.

This Court should thus be mindful to confine its opinion to compelled editorial disclosures and not all disclosures that may be required of social media companies. Not all policies and practices of social media sites pertain to the sites' editorial policies and practices.

In particular, many data privacy practices do not relate to or intrude upon the editorial process, and will be analyzed differently than the compelled editorial disclosures at issue with AB587. In contrast to laws that regulate how editorial decisions are made, Courts routinely apply intermediate First Amendment scrutiny, and not strict scrutiny, to consumer data privacy laws. *See, e.g., Trans Union Corp. v. FTC (Trans Union I)*, 245 F.3d 809, 818–19 (D.C. Cir. 2001) (upholding FTC rule under Fair Credit Reporting Act requiring opt-in consent to sell marketing lists); *Trans Union LLC v. FTC (Trans Union II)*, 295 F.3d 42, 52–53 (D.C. Cir. 2002) (upholding FTC rule under Gramm-Leach-Bliley Act that restricted sharing and use of consumer information); *Nat'l. Cable Assn. v. FCC*, 555 F.3d 996, 1000–02 (D.C. Cir. 2009) (upholding FCC rule under Telecommunications Act requiring

opt-in consent to disclose call records).

Amicus EFF also supports regulations that benefit users, promote user choice and control by encouraging competition and platform interoperability,⁶⁰ measures that can address perceived platform bias.⁶¹ Many of the same problems that legislators seek to address through editorial interference are better addressed through data privacy legislation.⁶²

⁶⁰ *Comment on Evaluating the Competitive Effects of Corporate Acquisitions and Mergers*, EFF (August 20, 2018), <https://www.eff.org/document/eff-comments-ftc-competition-0>.

⁶¹ Bennett Cyphers & Cory Doctorow, *The New ACCESS Act Is a Good Start. Here's How to Make Sure It Delivers*, EFF Deeplinks (June 21, 2021), <https://www.eff.org/deeplinks/2021/06/new-access-act-good-start-heres-how-make-sure-it-delivers>.

⁶² Corynne McSherry, et. al, *Privacy First: A Better Way to Address Online Harms*, EFF (Nov. 14, 2023), <https://www.eff.org/wp/privacy-first-better-way-address-online-harms>.

CONCLUSION

This court should reverse the district court because that court applied the wrong First Amendment standard. The requirement of AB587 that social media platforms disclose the details of their editorial processes implicates core First Amendment rights. The district court erred in applying diminished First Amendment scrutiny.

Dated: February 21, 2024

By: /s/ David Greene

David Greene

Aaron Mackey
ELECTRONIC FRONTIER
FOUNDATION
815 Eddy Street
San Francisco, CA 94109
Telephone: (415) 436-9333
Fax: (415) 436-9993
davidg@eff.org

*Counsel for Amicus Curiae
Electronic Frontier
Foundation*

CERTIFICATE OF COMPLIANCE

Pursuant to Fed. R. App. P. 32(g), I certify as follows:

1. This Brief of Amicus Curiae Electronic Frontier Foundation in Support of Plaintiff-Appellant and Reversal with the type-volume limitation of Fed. R. App. P. 32(a)(7)(B) because this brief contains 6,774 words, excluding the parts of the brief exempted by Fed. R. App. P. 32(f); and

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 this brief has been prepared in a proportionally spaced typeface using Microsoft Word 365, the word processing system used to prepare the brief, in 14 point font in Times New Roman font.

Dated: February 21, 2024

/s/ David Greene

David Greene

Counsel for Amicus Curiae

CERTIFICATE OF SERVICE

I hereby certify that I electronically filed the foregoing with the Clerk of the Court for the United States Court of Appeals for the Ninth Circuit by using the appellate CM/ECF system on February 21, 2024.

I certify that all participants in the case are registered CM/ECF users and that service will be accomplished by the appellate CM/ECF system.

Dated: February 21, 2024

/s/ David Greene

David Greene

Counsel for Amicus Curiae