UNIVERSITY OF CALIFORNIA, LOS ANGELES **UCLA**

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

EUGENE VOLOKH
GARY T. SCHWARTZ PROFESSOR OF LAW

SCHOOL OF LAW
405 HILGARD AVE.
LOS ANGELES, CA 90095-1476
(310) 206-3926
volokh@law.ucla.edu

November 30, 2021

Dear Subcommittee Chairman Doyle, Ranking Member Latta, Committee Chairman Pallone, Ranking Member Rodgers, and Members of the Subcommittee:

Many thanks for inviting me to testify on "Holding Big Tech Accountable: Targeted Reforms to Tech's Legal Immunity." I'm afraid I only received the invitation yesterday, so I had to prepare this testimony in a rush; my apologies if the result is incomplete or in some ways mistaken.

I understand the Subcommittee has been considering many different proposals, but I was asked to focus on five:

1. The Justice Against Malicious Algorithms Act.
2. The SAFE TECH Act.
3. The Preserving Constitutionally Protected Speech Act
4. The Protecting Americans Against Dangerous Algorithms Act.
5. The Civil Rights Modernization Act.

My plan is mostly to offer an evenhanded analysis of these proposals, focusing (in the interests of brevity) on possible nonobvious effects. I will also include my personal views on some of the proposals, but I will try to keep them separate from the objective analysis.

I. The Justice Against Malicious Algorithms Act

JAMAA would sharply limit interactive computer services' immunity for *personalized recommendations*, for instance for YouTube's recommendations of videos that come up alongside a video that you select. (YouTube recommends such videos in large part based on your search history.)

If the recommended material proves to be—for instance—defamatory, then under the bill YouTube could be liable for damages, since defamation often involves "severe emotional injury." (The Act would be limited to recommendations that "materially contributed to a physical or severe emotional injury to any person.") Likewise with Twitter or Facebook recommending posts based on your past interests, and more.

On the other hand, interactive services would remain immune for *unpersonalized recommendations*—for instance, recommendations of material based on its general popularity, uninfluenced by whether you've shown an interest in such material. And interactive services would be practically protected from liability for *recommending mainstream media material*: That material is less likely to be defamatory or otherwise injurious, and in any event mainstream media organizations have deep pockets, so the computer services could require that those organizations agree to indemnify the services in case of a lawsuit.

JAMAA would thus create a strong incentive for

- YouTube, Facebook, Twitter, etc. to stop recommending user-generated content that they think you would find especially interesting and
- instead to start recommending (1) generic popular material or (2) mainstream media content.

This strikes me as a bad idea. Users benefit from seeing recommendations for things they are especially likely to enjoy: If you like hip-hop, for instance, you'd presumably want to see recommendations for the most popular hip-hop video and not for the most popular material of any genre (which this week might be, say, Adele or Taylor Swift). Indeed, the more personalized the recommendations are, the more you're likely to enjoy them. Why pressure platforms to shift to generic material?

And the public also benefits, I think, from being able to see user-generated conduct and not just professionally produced mainstream media content. The established professional material already has a huge advantage, because of its existing marketing muscle. Why extend that privilege further by making it risky for platforms to recommend user-generated content (even when their algorithms suggest that such content might be exactly what you would most enjoy), and safe to recommend the professional material?

II. Preserving Constitutionally Protected Speech Act

This bill contains several different provisions.

A. Enabling State Civil Rights Laws That Ban Political Discrimination

The bill would change § 230(c)(2) to provide (in proposed new § 230A(a)(2)) that,

> No provider of an interactive computer service that is a covered company [basically, a large social media platform] shall be held liable on account of ... any action voluntarily taken in good faith to restrict access to or availability of material
>
> that is not constitutionally protected or that the provider has an objectively reasonable belief is obscene, lewd, lascivious, filthy, excessively violent, or harassing.

The current version of (c)(2), on the other hand, closes with:

> that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.

To oversimplify, the bill would make clear that platforms have no federal immunity when they block constitutionally protected material that isn't sexual, violent, or harassing. States could then, if they choose, limit platforms' ability to remove posts and users based on the posts' and users' political ideas (or religious, scientific, and other ideas). The bill wouldn't itself ban such political discrimination, but it would clearly allow states to do so.[1]

Whether such bans on political discrimination by social media platforms are constitutional under the First Amendment, and whether they are a good idea, are difficult questions, which I canvass in a recent article.[2] But the bill would make clear that § 230 doesn't preclude such bans.

B. Requirement That Users "Knowingly and Willfully Select[] ... Algorithm[s]" for Displaying Content

The bill would strip large platforms of immunity when they "utilize[] an algorithm to amplify, promote, or suggest content to a user unless a user knowingly and willfully selects an algorithm to display such content" (proposed § 230A(c)(3)). Yet everything that computers do, they do via "algorithm[s]."

This means that any platform that amplifies, promotes, or suggests content to a user will have to make sure that the user "knowingly and willfully selects" that platform's "algorithm." This might simply mean that the platform will have to prompt each of its users with a "Click here to select our algorithm for suggesting material to you," and refrain from "amplify[ing], promot[ing], or suggest[ing]" any content to a user until the user clicks. If so, then that should be easy enough for the platform to do—though it's hard to see how it would help anyone.

On the other hand, if such a click isn't enough to count as "a user knowingly and willfully select[ing] an algorithm," then it's hard to know what platforms could do by way of suggesting content. Would they have to provide a choice of at least two different algorithms, so the user's action counts as truly "select[ing]"? Would they have to explain in detail each of the algorithms, so that it counts as "knowingly and willfully select[ing]"? Would they have to do something else? And what benefit would that provide to the user? It's hard to know given the current language.

C. Requirement That Platforms Provide Explanations and Appeals in Case of Removal

The bill would also provide (sec. 201, emphasis added) that

[1] It's possible that even the existing 47 U.S.C. § 230(c) doesn't stop states from banning platforms from removing posts based on the posts' political views, see Adam Candeub & Eugene Volokh, *Interpreting 47 U.S.C. § 230(c)(2)*, 1 J. Free Speech L. 175 (2021), https://www.journaloffreespeechlaw.org/candeubvolokh.pdf. But right now that's just a possibility, on which courts are divided.

[2] Eugene Volokh, *Treating Social Media Platforms as Common Carriers?*, 1 J. Free Speech L. 377 (2011), http://www.law.ucla.edu/volokh/pubaccom.pdf.

> Each covered company shall implement and maintain *reasonable and user-friendly* appeals processes for decisions about content on such covered company's platforms....
>
> For any content a covered company edits, alters, blocks, or removes, the covered company shall— ...
>
> *clearly* state why such content was edited, altered, blocked, or removed, including by citing the specific provisions of such covered company's content policies on which the decision was based ....

Sec. 201 seems to require only transparency and an appeal process, without any substantive criteria for what platforms may or may not remove; in that respect, these requirements would presumably be quite limited in scope. But the bill doesn't explain what counts as "reasonable" appeals, "user-friendly" appeals, or "clearly stat[ing]." For instance, say a platform says "we removed the material because it was pornographic / hateful / misleading / supportive of violence." Is that clear enough, or would the platform have to provide more details on where it draws the line between pornography and art? Would the platform have to explain why it views a statement as "hateful" or "supportive of violence," when the statement also has other possible meanings? Would the platform have to explain why it viewed certain controversial material as "misleading"?

Likewise, the bill states that any appeal must "provide an opportunity for [the] user to present reasons why the covered company's action should not have been taken, including demonstrating inconsistent application of such company's specific content policy at issue." Would the platform then need to "clearly state" why it views this material as deserving removal when it didn't remove past material, as to which the rules were supposedly "inconsistent[ly] appli[ed]"? How much expense, litigation, or deterrence to removal the proposal would yield depends heavily on how terms such as "clearly state" end up being interpreted.

The provisions would apparently be enforced only by the FTC (sec. 203(a)) or by state attorneys general or other executive officials (sec. 203(b)), and not by private litigants. But, as noted in Part II.A, the bill would free states to (1) ban political discrimination by social media platforms and to (2) let private litigants sue over such discrimination. If some states do that, then the transparency requirements would help the private litigants marshal evidence that they were indeed discriminated against based on their political views.

D. "Conservative"/"Liberal" Accounts

The provision in sec. 202(a)(4)-(5) requiring that platforms disclose "the number of [content enforcement] decisions related to conservative content and conservative accounts" and "to liberal content and liberal accounts" are likely unconstitutionally vague. There is no established definition of "conservative" and "liberal," and it's hard to imagine how such a definition could be developed in a way that is clear enough for a legal rule.[3]

III. SAFE TECH Act

This bill contains several different provisions; let me focus on some of the less obvious ones.

A. Stripping Immunity from Paid Hosting Services (e.g., WordPress), Platforms That Share Ad Revenue with Creators (e.g., YouTube), and Platforms That Subsidize New Content

The bill would deny immunity to providers that have "accepted payment to make the speech available or, in whole or in part, created or funded the creation of the speech" (sec. 2(1)(A)(iii)).

This would threaten liability for any service that charges to provide hosting—for instance, blogging platforms such as WordPress or hosting services such as Amazon Web Services. After all, they "accept[] payment to make the speech available," which is unsurprising since they're in business to make money. Advertising-supported free services (which generally make money by selling access to their users, and their users' data) would still be immune, so the market would be strongly pushed in that direction.

This section would also threaten liability on any service that shares its advertising revenue with creators, for instance as YouTube does. After all, by letting providers of popular videos monetize those videos, YouTube would be "in part[] . . . fund[ing] the creation of the speech." (The providers will likely have created the videos in expectation of making money from them on YouTube, and the money they make would help fund future videos.) Creators would thus be less likely to earn money from their works, unless they're earning so much as to make it worth the platform's while to run the risk of liability in exchange for a share of the proceeds.

And the section would threaten liability whenever any providers provide grants to support local journalism or other such projects (something like the Google News Initiative[4]), since there the providers will have again "in part[] . . . funded the creation of the speech." Providers would thus become less likely to directly or indirectly support journalism and other expression.

[3] *Cf. Hynes v. Mayor & Council of Oradell*, 425 U.S. 610 (1976) (striking down as unconstitutionally vague a requirement that door-to-door political solicitors register with the city before soliciting "for a Federal, State, County or Municipal political . . . . cause," because "it is not clear what is meant by" that phrase).

[4] https://newsinitiative.withgoogle.com/info/innovation-challenges/.

B. Turning Immunity Into an Affirmative Defense: Likely a Largely Ineffective Modification

The bill specifies (sec. 2(1)(A)(iv)) that a provider that seeks immunity from liability would assert the immunity as an "affirmative defense," and thus "shall have the burden of persuasion, by a preponderance of the evidence," that it "is a provider or user of an interactive computer service and is being treated as the publisher or speaker of speech provided by another information content provider." But this change is unlikely to do much.

The burden of proof (or of persuasion) is important in *factual* disputes. If there's a close call on the evidence, who has the burden of proof might matter. But the placement of the burden generally doesn't much affect *legal* questions, such as how a statute is to be interpreted or how a legal claim is characterized.

Whether a defendant is a provider or user of an interactive computer service can be a factual question, but one on which the facts will rarely be close. Twitter is clearly a provider of an interactive computer service, wherever you place the burden of persuasion.

And whether a defendant "is being treated as the publisher or speaker of speech provided by another information content provider" by the plaintiff's Complaint is a legal question, which the burden of persuasion is unlikely to affect. Thus, for instance, the seminal decision in *Zeran v. America Online, Inc.* concluded, as a matter of law, that plaintiff's "complaint treats AOL as a publisher" and that plaintiff's lawsuit is therefore preempted. 129 F.3d 327, 333 (4th Cir. 1997). On the other hand, *Fair Housing Council of San Fernando Valley v. Roommates.com* concluded, as a matter of law, that § 230 didn't apply, because plaintiff's claim was that a roommate advertising website's practices made it "become much more than a passive transmitter of information provided by others; it becomes the developer, at least in part, of that information." 521 F.3d 1157, 1166 (9th Cir. 2008). In neither case did the burden of persuasion matter.

Nor would making clear that § 230 immunity is an affirmative defense preclude early motions to dismiss. Indeed, courts have already recognized that § 230 immunity is an affirmative defense, yet allowed such motions. "Preemption under the Communications Decency Act is an affirmative defense, but it can still support a motion to dismiss if the statute's barrier to suit is evident from the face of the complaint." *Klayman v. Zuckerberg*, 753 F.3d 1354, 1357 (D.C. Cir. 2014); *see also Ricci v. Teamsters Union Loc. 456*, 781 F.3d 25, 28 (2d Cir. 2015); *Marshall's Locksmith Serv. Inc. v. Google, LLC*, 925 F.3d 1263, 1267 (D.C. Cir. 2019). Nor would the statement about burden of persuasion displace the principle that "Section 230 immunity, like other forms of immunity, is generally accorded effect at the first logical point in the litigation process," *Nemet Chevrolet, Ltd. v. Consumeraffairs.com, Inc.*, 591 F.3d 250, 254 (4th Cir. 2009), usually the motion to dismiss.

C. Enabling State Civil Rights Laws That Ban Political Discrimination

The bill would modify § 230 to provide that "Nothing in this section shall be construed to limit, impair, or prevent any action alleging discrimination on the basis of any protected class, or conduct that has the effect or consequence of discriminating on the basis of any protected class, under any Federal or State law." Some state legislatures are now discussing making political ideology a "protected class" under state public accommodations laws, and applying those laws to social media platforms.

D.C. law, for instance, already treats political party membership as a protected class, alongside race, sex, religion, and the like.[5] The Montana Constitution more broadly provides that (among other things) no "firm, corporation, or institution shall discriminate against any person in the exercise of his civil . . . rights on account of race, color, sex, culture, social origin or condition, or political or religious ideas."[6] Some local ordinances in cities such as Seattle, Madison, Lansing, Champaign-Urbana, and Ft. Lauderdale likewise ban discrimination in public accommodations based on political ideology (and not just party membership);[7] state legislatures may easily adopt similar laws as well.

As noted above, whether such bans on political discrimination by social media platforms are constitutional under the First Amendment, and whether they are a good idea, are difficult questions.[8] But the bill would make clear that § 230 doesn't preclude such bans.

IV. Protecting Americans Against Dangerous Algorithms Act

This bill would deprive platforms of immunity from

1. claims brought for conspiracy to interfere with civil rights (42 U.S.C. § 1985), failure to prevent conspiracy to interfere with civil rights (42 U.S.C. § 1986), and international terrorism (18 U.S.C. § 2333) when
2. "the claim involves a case in which the interactive computer service used an algorithm . . . to rank, . . . recommend, [or] amplify . . . information . . . provided to a user of the service if the information is directly relevant to the claim."

Responses to a user's "specifically search[ing] for" "information" are excluded, as are services with 10 million or fewer unique monthly visitors and infrastructure companies that provide hosting, domain registration, and the like.

[5] D.C. Code §§ 2-1401.02(25), -1402.31(a).

[6] Mont. Const. art. II, § 4.

[7] *See* Eugene Volokh, *Bans on Political Discrimination in Places of Public Accommodation and Housing*, 15 NYU J. L. & Lib. 709 (forthcoming 2021), http://www.law.ucla.edu/volokh/pubaccom.pdf.

[8] *See* Eugene Volokh, *Treating Social Media Platforms as Common Carriers?*, 1 J. Free Speech L. 377 (2011), http://www.law.ucla.edu/volokh/pubaccom.pdf.

The bill would also exclude recommendations that come from simple and nonpersonalized algorithms—sorting “chronologically or reverse chronologically,” “by average user rating or number of user reviews,” “alphabetically,” “randomly,” and “by views, downloads, or a similar usage metric.” But platforms are unlikely to want to use such simple algorithms in place of their usual, more complex algorithms (which turn on, for instance, a user’s own viewing history), since those more complex algorithms generally increase user engagement and thus platform profit.

A. Pressuring Platforms Not to Recommend Material That Appears Like It May Have Been Put Out by Terrorist Groups

The chief effect of PADAA would be to hold social media platforms liable for recommending material that later turns out to have been put out by foreign terrorist groups (or by people working directly with those groups). The difficulty, of course, is that a platform can’t know with any real certainty whether particular material is indeed put out by (say) Hamas employees or associates, or whether it is instead just constitutionally protected expression of support for Hamas. But because of this uncertainty, platforms would likely internally flag material that appears like it *could* have been put out by foreign terrorist groups, and exclude it from any recommendations they offer.

B. Pressuring Platforms Not to Recommend Pages That Appear to Involve Conspiracies to Interfere with Civil Rights

The bill’s allowing liability for 42 U.S.C. § 1985 violations is likely to have no real effect, because § 1985 basically just covers conspiracies to violate civil rights; to intimidate parties, witnesses, or jurors; to intimidate people to affect federal elections; or to injure people based on their advocacy of federal candidates. A conspiracy requires a specific purpose to promote a shared criminal objective,[9] and platforms are not likely to have any such specific purpose.

But the bill also allows liability for 42 U.S.C. § 1986 violations, and § 1986 imposes liability for failure to prevent *others’* conspiracies:

- “Every person who, having knowledge that any of the wrongs conspired to be done, and mentioned in section 1985 of this title, are about to be committed,”
- “and having power to prevent or aid in preventing the commission of the same, neglects or refuses so to do,”
- “if such wrongful act be committed, shall be liable to the party injured ....”

[9] Ocasio v. United States, 136 S. Ct. 1423, 1429 (2016)

It's hard to tell for sure, since successful § 1986 claims against private entities are so rare; but in principle, it seems that, under PADAA, once a platform learns of material that appears like it could violate § 1985, it would need to exclude it from any recommendations, or face the risk of liability. Such exclusion from recommendations, after all, may "prevent or aid in preventing the commission" of the conspiracy.

What we should think of such proposals to enlist platforms to police potential foreign terrorist advocacy and potential conspiracies to commit various domestic crimes is a difficult question. On one hand, such proposals may indeed make it harder for conspirators, foreign and domestic, to effectively organize and promote their crimes. On the other hand, they are also likely to lead to cautious platforms suppressing even constitutionally protected advocacy, since the platforms will have only limited information about who is posting material, why they are posting it, how the posters and their readers are likely to interpret it.

V. Civil Rights Modernization Act

This bill would, among other things, allow liability (civil or criminal) for "target[ed]" paid "advertisements" that may violate

- "(A) any Federal, State, or local law, any part of which prohibits discrimination or other adverse action on the basis of a protected class or status [i.e., actual or perceived race, color, ethnicity, religion, national origin, sex (including sexual orientation and gender identity), age, disability, familial status, pregnancy, genetic information, or citizenship or immigration status];"
- "(B) any other Federal law that is administered or enforced, in whole or in part, by the Civil Rights Division of the Department of Justice; or"
- "(C) any Federal, State, or local law that prohibits the dissemination of false or misleading information intended, with respect to an election for public office, to prevent voters from casting their ballots or to prevent voters from voting for the candidate of their choice."

"Targeting" is defined as using technology "to deliver or show a covered advertisement to any particular subset of users who are part of or have a protected class or status."

This would likely cover discriminatorily targeted ads for employment, housing, and the like (under (A)). And it would also potentially hold platforms liable (under (C)) for accepting ads that ultimately prove "false or misleading" in a way "intended ... to prevent voters from casting their ballots or to prevent voters from voting for the candidate of their choice," so long as the ads are targeted based on, for instance, age or familial status—or for that matter on citizenship status, in trying to focus on eligible voters.

This latter provision, related to elections, is potentially dangerous, because it puts platforms in a position where they can be liable for ads containing "misleading information."

Platforms often know little about particular elections, and about whether particular statements about those elections are likely to be seen as false or misleading by a future judge or jury. And if "misleading information intended ... to prevent voters from casting their ballots or to prevent voters from voting for the candidate of their choice" includes information intended to *dissuade* voters from voting or from voting for a candidate (and not just to dupe them into voting, say, at the wrong time or the wrong place[10]), platforms would potentially be still more exposed. Indeed, some jurisdictions do have statutes that purport to ban false statements about election campaigns; lower courts are split on whether such statutes are constitutional.[11]

The simplest way for platforms to avoid the risk of such liability is to require that any political ads as to which any question might arise—e.g., all political ads that criticize rival candidates and thus might be seen as including "misleading information intended ... to prevent voters from voting for the candidate of their choice"—be run in a purely untargeted way (since the bill would only strip away immunity for allegedly false or misleading ads that are targeted). I'm not sure whether this would on balance improve election-related discourse or unduly interfere with it.

* * *

I hope this analysis has been helpful. Thank you again for inviting me, and please let me know if there are any further questions I can answer.

Sincerely Yours,

[redacted]

Eugene Volokh

[10] *See* Eugene Volokh, *Are Douglas Mackey's Memes Illegal?*, Tablet, Feb. 9, 2021, https://www.tabletmag.com/sections/news/articles/douglass-mackey-ricky-vaughn-memes-first-amendment.

[11] For cases upholding such statutes, see *In re Chmura*, 608 N.W.2d 31 (Mich. 2000); *State v. Davis*, 27 Ohio App. 3d 65 (1985). For cases striking them down, see *Susan B. Anthony List v. Driehaus*, 814 F.3d 466 (6th Cir. 2016); *281 Care Comm. v. Arneson*, 766 F.3d 774 (8th Cir. 2014); *Commonwealth v. Lucas*, 472 Mass. 387 (2015); *State ex rel. Public Disclosure Comm'n v. 119 Vote No! Comm.*, 135 Wash. 2d 618 (1998).