

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

X CORP.,

Plaintiff,

-against-

LETITIA JAMES, in her official capacity, as
Attorney General of New York,

Defendant.

**COMPLAINT FOR
DECLARATORY AND
INJUNCTIVE RELIEF**

Case No. 25-cv-5068

JURY TRIAL DEMANDED

Plaintiff X Corp., by and through its attorneys, Cahill Gordon & Reindel LLP, alleges for its complaint against the above-named Defendant, as follows:

NATURE OF THE ACTION

1. Plaintiff X Corp. brings this action challenging the constitutionality and legal validity of certain provisions within New York Senate Bill S895B¹ that force social media companies (collectively, the “covered companies”) to divulge—under threat of significant financial penalties and lawsuits for injunctive relief—highly sensitive information about whether, and if so how, their platforms define and moderate (i) hate speech or racism, (ii) extremism or radicalization, (iii) disinformation or misinformation, (iv) harassment, and (v) foreign political interference. *See* N.Y. Gen. Bus. Law §§ 1102(1)(c), 1102(1)(d)(i), 1102(1)(e), and 1103 (only insofar as it applies to the prior three provisions) (collectively, the “Content Category Report Provisions”).

¹ S895B is codified at Article 42 (Sections 1100–1104) of the New York General Business Law. For the Court’s convenience, S895B is provided in its entirety as Exhibit 1.

2. The Content Category Report Provisions violate the First Amendment of the United States Constitution and Article I, Section 8, of the New York Constitution. On their face, the provisions compel disclosure of highly sensitive and controversial speech that is both fully protected by the First Amendment and that is disfavored by the State of New York. The law thus impermissibly interferes with the First Amendment-protected editorial judgments of companies such as X Corp. to remove, demonetize, or deprioritize such speech on their platforms. *See Moody v. NetChoice, LLC*, 603 U.S. 707, 729–30, 744 (2024). That the Content Category Report Provisions attempt to tilt the marketplace of ideas by utilizing *public* pressure to achieve this aim does not save them, as the government “cannot do indirectly what [it] is barred from doing directly.” *Nat’l Rifle Ass’n of Am. v. Vullo*, 602 U.S. 175, 190 (2024); *see also NetChoice, LLC v. Bonta*, 113 F.4th 1101, 1118 (9th Cir. 2024) (striking down law on First Amendment grounds that “deputizes covered businesses into serving as censors for the State”).²

3. S895B is based on a nearly identical California law—California Assembly Bill No. 587 (“AB 587”) ³—that was successfully challenged by X Corp. in federal court in California on identical grounds to the challenge brought here. *See X Corp. v. Bonta*, No. 2:23-cv-01939-WBS-AC (E.D. Cal.). After lengthy litigation about the constitutionality of California’s AB 587, on September 4, 2024, the Ninth Circuit Court of Appeals granted a preliminary injunction preventing enforcement of the portions of the law that are analogous to the Content Category Report Provisions of S895B challenged here. The Ninth Circuit held that X Corp. was likely to prevail

² Unless otherwise indicated, emphases in quotes are added and internal citations and quotations are omitted.

³ AB 587 is provided as Exhibit 2. *See also, e.g., Ex. 3 at 220 (June 6, 2024 New York State Assembly Floor Debates, 2023–24 Reg. Sess. (N.Y. 2024), available at <https://www2.assembly.state.ny.us/write/upload/transcripts/2023/6-6-24.pdf#search=%226789%22>)* (bill sponsor Assemb. Member Grace Lee noting that S895B is “identical” to AB 587).

on its facial First Amendment challenge of those provisions. *See X Corp. v. Bonta*, 116 F.4th 888 (9th Cir. 2024).

4. The thoughtful, reasoned decision of the Ninth Circuit explained that AB 587’s “Content Category Report provisions likely compel non-commercial speech and are subject to strict scrutiny, under which they do not survive,” because they, among other things, require social media companies to “recast [their] content-moderation practices in language prescribed by the State” and “opin[e] on whether and how certain controversial categories of content should be moderated.” *Id.* at 898, 901; *see also id.* at 899 n.8 (“No matter how a social media company chooses to moderate such content, the company will face backlash from its users and the public. That is true even if the company decides not to define the enumerated categories, because they will draw criticism for under-moderating their community.”).

5. The Ninth Circuit further held that the “Content Category Report provisions compel every covered social media company to reveal its policy opinion about contentious issues, such as what constitutes hate speech or misinformation and whether to moderate such expression,” and thus “raise the same First Amendment issues for every covered social media company.” *Id.* at 899.

6. Shortly after the Ninth Circuit’s ruling, the California Attorney General, Robert Bonta, agreed to a settlement that provided for a permanent injunction preventing enforcement of the portions of the law that were preliminarily enjoined by the Ninth Circuit and further provided for payment by the State of California of certain of X Corp.’s fees and expenses.

7. On March 4, 2025, pursuant to that settlement, the United States District Court for the Eastern District of California judicially declared that the Content Category Report Provisions of AB 587 violate the First Amendment of the U.S. Constitution, both facially and as-applied to X Corp. *See Ex. 4 at 2 (Order and Final Judgment and Permanent Injunction, X Corp. v. Bonta, No.*

2:23-cv-01939-WBS-AC (E.D. Cal. Mar. 4, 2025), Dkt. No. 59). The Final Judgment and Permanent Injunction also required the California Attorney General to pay X Corp. \$345,576 to cover its fees and costs incurred in connection with its prosecution of AB 587. *Id.* at 3.

8. The Content Category Report Provisions challenged in this case are a carbon copy of the provisions of AB 587 that were preliminarily enjoined by the Ninth Circuit and permanently enjoined as part of the settlement of *X Corp. v. Bonta*, No. 2:23-cv-01939-WBS-AC (E.D. Cal.).

9. The New York legislature cited the existence of California’s AB 587 as support for S895B.⁴ But inexplicably, not only did the New York legislature fail to revisit S895B after the Ninth Circuit’s decision, it “*refuse[d]*” to even have a “discussion” with X Corp. about “potential changes” to S895B after the Ninth Circuit’s ruling. In rejecting an overture from X Corp. to discuss such changes, the bill sponsors stated that they “decline the request on behalf of your client, X, to discuss amendments to [S895B]” because, in the view of the sponsors, X’s “owner, Elon Musk” used the X platform to promote content that, in the view of the bill sponsors, “threatens the foundations of our democracy”—indicating that S895B’s passage was tainted by viewpoint discriminatory motives. Ex. 5 at 1–3 (Letter from Bill Sponsors Sen. Brad Hoylman-Sigal and Assemb. Member Grace Lee to X Corp. (Sept. 9, 2024)).⁵

⁴ Ex. 3 at 220 (*June 6, 2024 New York State Assembly Floor Debates*, 2023–24 Reg. Sess. (N.Y. 2024), available at <https://www2.assembly.state.ny.us/write/upload/transcripts/2023/6-6-24.pdf#search=%226789%22>) (citing *X Corp. v. Bonta*, 2023 WL 8948286 (E.D. Cal. Dec. 28, 2023) (finding that AB 587 likely did not violate the First Amendment), *rev’d and remanded*, 116 F.4th 888 (9th Cir. 2024)); *see also X Corp.*, 116 F.4th at 902 (noting that the “district court performed, essentially, no analysis” on the question of whether the Content Category Report Provisions compelled commercial speech).

⁵ Available at <https://www.nysenate.gov/sites/default/files/admin/structure/media/manage/filefile/a/2024-09/twitter-letter.pdf>.

10. The Content Category Report Provisions of S895B violate the First Amendment for all of the reasons that the Ninth Circuit concluded that the Content Category Report Provisions of California’s AB 587 likely violated the First Amendment.

11. The Content Category Report Provisions also violate the immunity provided to interactive computer service providers such as X Corp. under 47 U.S.C. § 230(c)(2)(A), which prohibits liability “on account” of “any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.” The Content Category Report Provisions contravene that immunity by imposing monetary fines and suits for injunctive relief on X Corp. if it “restrict[s] access” to content on its platform in a way that, in the unfettered discretion of Defendant, “misrepresents” information in a Terms of Service Report. *See* § 1103(1)(b)(iii). This directly contravenes the immunities afforded under Section 230. *Compare, e.g., Doe v. Internet Brands, Inc.*, 824 F.3d 846, 851–52 (9th Cir. 2016) (Section 230 protects the ability to “self-regulate offensive third party content without fear of liability”), *with* Ex. 6 (*ADL, Senator Brad Hoylman-Sigal, and AM Grace Lee Demand Social Media Companies Stop Hiding Hate*, nysenate.gov (May 23, 2023))⁶ (“[W]e cannot rely on social media companies to consistently self-regulate.”).

12. In pursuing this action, X Corp. seeks declaratory relief and injunctive relief on the grounds that S895B violates (i) X Corp.’s and the other covered companies’ free speech rights under the First Amendment of the United States Constitution and Article I, Section 8, of the New

⁶ Available at <https://www.nysenate.gov/newsroom/press-releases/2023/brad-hoylman-sigal/adl-senator-brad-hoylman-sigal-and-am-grace-lee>.

York Constitution and (ii) the immunity afforded to X Corp. and the other covered companies under 47 U.S.C. § 230(c)(2)(A).

13. In pursuing this action, X Corp. seeks to vindicate the deprivation of constitutional rights under color of state statute, ordinance, regulation, custom, and/or usage. X Corp. is also entitled to attorneys' fees and costs if it prevails on any of its Section 1983 claims. *See* 42 U.S.C. § 1988.

PARTIES

14. Plaintiff X Corp. is a corporation organized and existing under the laws of the State of Nevada, with its principal place of business in Bastrop, Texas. X Corp. provides the X service, which is a real-time, open, public conversation platform, where people can see every side of a topic, discover news, share their perspectives, and engage in discussion and debate. X allows people to create, distribute, and discover content and has democratized content creation and distribution. X allows users to create and share ideas and information instantly through various product features, including public posts.

15. S895B and the Content Category Report Provisions therein apply to X Corp. because it is a "social media company," as defined in the statute—i.e., an "entity that owns or operates one or more social media platforms," *see* § 1100(4)—and has generated more than \$100 million in gross revenue in the preceding calendar year. *See* § 1104 (exempting social media companies that generated less than \$100 million in gross revenue during the preceding calendar year).

16. X is a "social media platform," as defined by S895B, because it is a public internet-based service or application with users in New York and (i) "[a] substantial function of the service or application is to connect users in order to allow users to interact socially with each other within

the service or application” and (ii) it allows its users to (a) “construct a public or semipublic profile for purposes of signing into and using the service or application”; (b) “populate a list of other users with whom an individual shares a social connection within the system”; and (c) “create or post content viewable or audible by other users, including, but not limited to, livestreams, on message boards, in chat rooms, or through a landing page or main feed that presents the user with content generated by other users.” § 1100(5).

17. Defendant Letitia James is the Attorney General of the State of New York and is charged with enforcing S895B. *See* § 1103(2). X Corp. sues Attorney General James in her official capacity as the person charged with enforcing S895B and its Content Category Report Provisions.

JURISDICTION

18. This Court has jurisdiction over X Corp.’s federal claims pursuant to 28 U.S.C. §§ 1331 and 1343(a) and 42 U.S.C. § 1983, because X Corp. alleges violations of its rights under the U.S. Constitution and laws of the United States. The Court has jurisdiction over X Corp.’s state claim pursuant to 28 U.S.C. § 1367.

19. This Court has authority to grant declaratory and injunctive relief under the Declaratory Judgment Act, 28 U.S.C. §§ 2201, 2202, and under the Court’s inherent equitable jurisdiction.

VENUE

20. Venue is proper in this Court under 28 U.S.C. §§ 1391(b)(1) and 1391(b)(2) because the Defendant is located, resides, and has offices in this judicial district and in the State of New York, and the violations of X Corp.’s rights are occurring and/or will occur within this judicial district.

FACTUAL ALLEGATIONS

I. S895B’s Statutory Scheme

21. S895B, which applies to “social media companies” (defined as persons or entities owning or operating one or more “social media platforms”), has three main components—(i) a requirement that social media companies publicly post their terms of service, including processes for flagging content and potential actions that may be taken with respect to flagged content (“Terms of Service Requirement”), *see* § 1101; (ii) a requirement that social media companies submit a report to the New York Attorney General, who will in turn disseminate the report publicly, that includes, among other things, whether and, if so, how the company defines and/or moderates hate speech, racism, extremism, radicalization, disinformation, misinformation, harassment, and foreign political interference and certain statistics relating to moderation of those categories of content (“Terms of Service Report”), *see* § 1102; and (iii) a penalty provision, whereby companies may be liable to pay \$15,000 per violation per day and may be sued in court by the New York Attorney General for (a) violating the Terms of Service Requirement, (b) failing to timely submit a Terms of Service Report, or (c) materially omitting or misrepresenting required information in a Terms of Service Report, *see* § 1103.

a. Required Disclosure of Terms of Service (Terms of Service Requirement)

22. S895B’s Terms of Service Requirement mandates that social media companies publicly post, “in a manner reasonably designed to inform all users” of its “existence and contents,” the “terms of service for each social media platform” they own or operate. **§ 1101(1)**. The terms of service must include (i) “contact information,” so users may “ask the social media company questions about” the terms, (ii) a “description of the process that users must follow to flag content, groups, or other users that they believe violate the terms of service,” as well as “the social media

company’s commitments on response and resolution time,” and (iii) a “list of potential actions the social media company may take against an item of content or a user, including, but not limited to, removal, demonetization, deprioritization, or banning.” § 1101(2).

b. Terms of Service Report

23. S895B requires social media companies to submit semiannual Terms of Service Reports, on April 1 and October 1 of each year, to the New York Attorney General, who will then make them publicly available on the Attorney General’s website. *See* § 1102(2)(a).⁷ The Terms of Service Report must include:

- a. A statement of whether, and if so how, the platform’s terms of use define hate speech, racism, extremism, radicalization, disinformation, misinformation, harassment, and foreign political interference, § 1102(1)(c);
- b. The current version of the terms of service and a “complete and detailed description of any changes” to the terms since any previous report, § 1102(1)(a)–(b);
- c. A “detailed description” of the platform’s “content moderation practices,” including, but not limited to:
 - i. Any “existing policies intended to address the categories of content described in [§ 1102(1)(c)],” § 1102(1)(d)(i);
 - ii. How “automated content moderation systems enforce” the platform’s terms of service and “when these systems involve human review,” § 1102(1)(d)(ii);

⁷ Notwithstanding the requirements of Section 1102(2)(a), the first Content Category Report, covering activity within the third quarter of 2025, is due to the New York Attorney General no later than January 1, 2026; and the second Content Category Report, covering activity within the fourth quarter of 2025, is due to the New York Attorney General no later than April 1, 2026. § 1102(2)(b).

- iii. How the “company responds to user reports of violations of the terms of service,” § 1102(1)(d)(iii); and
 - iv. How the “company would remove individual pieces of content, users, or groups that violate the terms of service, or take broader action against” individuals or groups of users “that violate the terms of service,” § 1102(1)(d)(iv);
- d. Information about “content that was flagged by the social media company as content belonging to any of the categories described in” § 1102(1)(c), including:
- i. The “total number of flagged items of content,” § 1102(1)(e)(i)(A);
 - ii. The “total number of actioned items of content,” § 1102(1)(e)(i)(B);
 - iii. The “total number of actioned items of content that resulted in action taken by the social media company against the user or group of users responsible for the content,” § 1102(1)(e)(i)(C);
 - iv. The “total number of actioned items of content that were removed, demonetized, or deprioritized” by the company, § 1102(1)(e)(i)(D);
 - v. The “number of times actioned items of content were viewed by users,” § 1102(1)(e)(i)(E);
 - vi. The “number of times actioned items of content were shared, and the number of users that viewed or heard the content before it was actioned,” § 1102(1)(e)(i)(F); and
 - vii. The “number of times users appealed social media company actions taken on” the platform and the “number of reversals of social media company actions on appeal disaggregated by each type of action,” § 1102(1)(e)(i)(G);
- e. All of the information required by § 1102(1)(e)(i) “disaggregated” by:

- i. The “category of content, including any relevant categories described in [§ 1102(1)(c)],” § **1102(1)(e)(ii)(A)**;
- ii. The “type of content” (e.g., “posts, livestreams, comments, messages, profiles of users, or groups of users”), § **1102(1)(e)(ii)(B)**;
- iii. The “type of media of the content” (e.g., “text, images, livestreams, and videos”), § **1102(1)(e)(ii)(C)**;
- iv. How “the content was flagged” (e.g., “flagged by company employees or contractors, flagged by artificial intelligence software, flagged by community moderators, flagged by civil society partners, and flagged by users”), § **1102(1)(e)(ii)(D)**; and
- v. How “the content was actioned” (e.g., “actioned by company employees or contractors, actioned by artificial intelligence software, actioned by community moderators, actioned by civil society partners, and actioned by users”), § **1102(1)(e)(ii)(E)**.

24. In this case, X Corp. challenges only those provisions of the Terms of Service Report that require certain disclosures and statistics concerning the statute’s specified categories of controversial content—specifically, §§ 1102(1)(c), 1102(1)(d)(i), and 1102(1)(e), which are referred to here as the “Content Category Report Provisions.”

25. The Content Category Report Provisions are an impermissible attempt by the State to inject itself into the content-moderation editorial process. By requiring detailed disclosures about how certain controversial content is moderated, the State is impermissibly compelling social media companies to make disclosures about how their editorial processes work. By compelling social media companies to make politically charged disclosures about content moderation, the

State is impermissibly trying to generate public controversy about content moderation in a way that will pressure social media companies, such as X Corp., to restrict, limit, disfavor, or censor certain constitutionally protected content on X that the State dislikes.

26. These forced disclosures violate the First Amendment of the U.S. Constitution and Article I, Section 8, of the New York Constitution.

c. Penalties

27. S895B also sets forth a penalty scheme under which social media companies may be fined \$15,000 per violation per day, and may be enjoined in any court of competent jurisdiction by the New York Attorney General if the company (i) “fails to post terms of service in accordance with” Section 1102, (ii) “fails to timely submit” a Terms of Service Report, or (iii) “materially omits or misrepresents required information in” a Terms of Service Report. **§ 1103(1)(a)–(b).** S895B’s penalty provision further instructs that the court shall, “[i]n assessing the amount of a civil penalty pursuant to” Section 1103(1)(a), “consider whether the social media company has made a reasonable, good faith attempt to comply with the provisions of this article.” **§ 1103(1)(c).**

II. S895B and its Content Category Report Provisions Impermissibly Substitute the Government’s Judgment About Content Moderation for That of Covered Social Media Platforms

28. The topics that S895B forces social media platforms to speak about against their will are highly controversial and politically charged. Those categories were taken word-for-word from California’s AB 587. *Compare* §§ 1102(1)(c), 1102(1)(d)(1), and 1102(1)(e), *with* Cal. Bus. & Prof. Code §§ 22677(a)(3), (a)(4)(A), and (a)(5). The legislative history of AB 587 makes clear that those categories of content were selected largely because they are the most controversial. As the Assembly Reports from the California Legislature’s Committee on Privacy and Consumer

Protection make clear, the categories of content covered by the statute are the most difficult to define and moderate because they are “fraught with political bias”:

As online social media become increasingly central to the public discourse, the companies responsible for managing social media platforms are faced with a complex dilemma regarding content moderation, i.e., how the platforms determine what content warrants disciplinary action such as removal of the item or banning of the user. In broad terms, there is a general public consensus that certain types of content, such as child pornography, depictions of graphic violence, emotional abuse, and threats of physical harm are undesirable, and should be mitigated on these platforms to the extent possible. *Many other categories of information, however, such as hate speech, racism, extremism, misinformation, political interference, and harassment [i.e., the categories that are the focus of California’s AB 587 and New York’s S895B], are far more difficult to reliably define, and assignment of their boundaries is often fraught with political bias. In such cases, both action and inaction by these companies seems to be equally maligned: too much moderation and accusations of censorship and suppressed speech arise; too little, and the platform risks fostering a toxic, sometimes dangerous community.*

Ex. 7 (CAL. ASSEMBLY COMM. ON PRIV. AND CONSUMER PROT., AB 587, 2021–22 Sess., at 4 (Apr. 22, 2021)).

29. By requiring X and the other covered platforms to disclose whether and, if so, how they regulate the controversial and difficult-to-define categories of content targeted by Section 1102(1)(c), the Content Category Report Provisions impermissibly attempt to pressure the covered platforms to adopt—and regulate—these categories of content, even if X Corp. would prefer to categorize content differently.

30. The legislative history of California’s AB 587 (upon which S895B is based) makes this plain. As the California Assembly’s Committee on the Judiciary report on AB 587 conceded, the intent of the law was, in part, to “pressure” social media companies to do more to eliminate the controversial categories of content covered by the Content Category Report Provisions. As the California legislature put it, “if social media companies are forced to disclose what they do in this regard [i.e., how they moderate online content], it may pressure them to become better corporate

citizens by doing more to eliminate hate speech and disinformation.” Ex. 8 (CAL. ASSEMB. COMM. ON JUDICIARY REPORT, AB 587, 2021–22 Sess., at 4 (Apr. 27, 2021)).

31. The government does not have a compelling interest—or any legitimate interest at all—in applying such pressure. Yet that is precisely what S895B will do and that is precisely what it is intended to do.

32. For example, X Corp. does not currently regulate “hate speech,” “racism,” or “extremism,” which are three types of content that the Content Category Report Provisions force social media companies to publicly address. But X does regulate “hateful conduct,” a category of content that may include content that some people might argue constitutes “hate speech,” “racism,” or “extremism.” See Ex. 9 (*Hateful Conduct*, X, <https://help.x.com/en/rules-and-policies/hateful-conduct-policy> (last visited June 17, 2025)).

33. Similarly, while X does not currently regulate “misinformation” or “disinformation,” see § 1102(1)(c), X does regulate “manipulated, or out-of-context media that may result in widespread confusion on public issues, impact public safety, or cause serious harm.” See Ex. 10 (*Authenticity*, X, <https://help.x.com/en/rules-and-policies/authenticity> (last visited June 17, 2025)).

34. While the law does not force X to adopt and regulate “hate speech,” “racism,” “extremism,” “misinformation,” or “disinformation,” it attempts to pressure X to do so in an insidious and impermissible way. Put differently, by requiring public disclosure of the content-moderation policies covered platforms adhere to with respect to the highly controversial topics that the State has identified, the State is seeking nothing less than increasing public pressure on

covered platforms like X to regulate and therefore limit such content, notwithstanding that just such decisions are not for the State to make.

35. S895B and its Content Category Report Provisions are, in short, an attempt by the State to impermissibly frame the debate about content moderation in a way that pressures social media companies to regulate constitutionally protected content that the State finds objectionable.

36. Social media content moderation is an inherently controversial undertaking. How social media companies apply their content-moderation rules to particular posts is a sensitive and controversial exercise that is subject to rigorous debate and controversy. Deciding what content should appear on a social media platform is a question that engenders considerable debate among reasonable people about where to draw the correct proverbial line.

37. This is not a role that the government may play. “However imperfect the private marketplace of ideas,” a “worse proposal,” as here, is “the government itself deciding when speech [is] imbalanced, and then coercing speakers to provide more of some views or less of others.” *Moody*, 603 U.S. at 733.

38. The Content Category Report Provisions are not saved by the fact that the State is using them to apply pressure on covered platforms indirectly, rather than directly. As the Supreme Court has made clear, when it comes to First Amendment rights, “a government official cannot do indirectly what she is barred from doing directly.” *Vullo*, 602 U.S. at 190; *see also Smith v. Cal.*, 361 U.S. 147, 154 (1959) (government efforts to impose a system that encourages “self-censorship” by booksellers violates the First Amendment because it results in censorship “compelled by the State . . . affecting the whole public”); *NetChoice, LLC*, 113 F.4th at 1118 (striking down law on First Amendment grounds that “deputizes covered businesses into serving as censors for the State”).

FIRST CAUSE OF ACTION

(Declaratory Relief and Preliminary and Permanent Injunctive Relief for Violations of the First Amendment of the United States Constitution (42 U.S.C. § 1983) and Article I, Section 8, of the New York Constitution—Facial and As-Applied)

39. X Corp. realleges and incorporates herein by reference each and every allegation set forth above.

40. The Content Category Report Provisions violate the First Amendment of the United States Constitution and Article I, Section 8, of the New York Constitution, both facially and as-applied to X Corp., by compelling X Corp. and the other covered social media companies to divulge publicly, under the threat of significant financial penalty and civil suit, their “policy opinion about contentious issues, such as what constitutes hate speech or misinformation and whether to moderate such expression.” *X Corp.*, 116 F.4th at 899; *see also id.* at 901–02 (“The Content Category Report Provisions would require a social media company to convey the company’s policy views on intensely debated and politically fraught topics, including hate speech, racism, misinformation, and radicalization, and also convey how the company has applied its policies.”). The First Amendment protects “both the right to speak freely and the right to refrain from speaking at all.” *Wooley v. Maynard*, 430 U.S. 705, 714 (1977).

41. Because the “Content Category Report Provisions compel every covered social media company to reveal” those opinions, they “raise the same First Amendment issues for every covered social media company.” *X Corp.*, 116 F.4th at 899; *see also, e.g., Moody*, 603 U.S. at 744 (quoting *Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Boston*, 515 U.S. 557, 570 (1995) (social media platforms “present[] a curated and edited compilation of [third party] speech,” which “is itself protected speech”)); *id.* at 729–730 (“A private party’s collection of third-party content into a single speech product (the operators’ ‘repertoire’ of programming) is itself expressive, and

intrusion into that activity must be specially justified under the First Amendment.”); *Miami Herald Pub. Co. v. Tornillo*, 418 U.S. 241, 258 (1974) (editorial control and judgment protected from government regulation by First Amendment).⁸

42. “The Content Category Report Provisions compel non-commercial speech and are subject to strict scrutiny.” *X Corp.*, 116 F.4th at 899. The “Content Category Reports are not commercial speech” because they “require a company to recast its content-moderation practices in language prescribed by the State, implicitly opining on whether and how certain controversial categories of content should be moderated.” *Id.* at 901. Accordingly, “few indicia of commercial speech are present in the Content Category Reports.” *Id.*

43. The Content Category Reports are not commercial speech because, e.g., they (i) “do not satisfy the ‘usual[] defin[ition]’ of commercial speech—i.e., ‘speech that does no more than propose a commercial transaction,’” *id.* (quoting *United States v. United Foods, Inc.*, 533 U.S. 405, 409 (2001)); and (ii) they “fail to satisfy at least two of the three *Bolger* factors,” since they “are not advertisements” and “a social media company has no economic motivation in their content,” *id.*; see also *Bad Frog Brewery, Inc. v. N.Y. State Liquor Auth.*, 134 F.3d 87, 96–97 (2d Cir. 1998) (quoting *Bolger v. Youngs Drug Prods. Corp.*, 463 U.S. 60, 66 (1983) (“core . . . commercial speech” is “speech which does no more than propose a commercial transaction”)); *id.* at 97 (citing *Bolger*, 463 U.S. at 66–67) (“Outside this so-called ‘core’ lie various forms of speech

⁸ The Content Category Report Provisions violate Article I, Section 8, of the New York Constitution for all of the same reasons that they violate the First Amendment of the United States Constitution. See, e.g., *Konikoff v. Prudential Ins. Co. of Am.*, 234 F.3d 92, 105 (2d Cir. 2000) (“[t]he protection afforded by the guarantees of free press and speech in the New York Constitution is often broader than the minimum required by the First Amendment”); *C.R. Corps v. Pestana*, 2022 WL 2118191, at *3 (S.D.N.Y. June 13, 2022) (“Article I, Section 8 [of the New York Constitution] generally offers broader protection than the First Amendment”); *Themed Restaurants, Inc. v. Zagat Surv., LLC*, 781 N.Y.S.2d 441, 449 (Sup. Ct. 2004) (“the free speech guarantee of the New York State Constitution is even more stringent than that of the First Amendment”), *aff’d*, 801 N.Y.S.2d 38 (App. Div. 1st Dep’t 2005).

that combine commercial and noncommercial elements,” depending on factors such as “whether the communication is an advertisement,” “whether the communication makes reference to a specific product,” and “whether the speaker has an economic motivation for the communication,” but none of these factors alone “would render the speech in question commercial,” and even the “presence of all three” would not be dispositive for a finding that the speech is commercial).

44. In addition, each of the content categories about which X Corp. and the other covered social media companies are compelled to speak—hate speech, racism, extremism, radicalization, disinformation, misinformation, harassment, and foreign political interference (§ 1102(1)(c))—is “anything but an ‘uncontroversial’ topic.” *Nat’l Inst. of Fam. & Life Advoc. v. Becerra* (“*NIFLA*”), 585 U.S. 755, 769 (2018). The same is true about how such categories of speech are defined and moderated on X. *X Corp.*, 116 F.4th at 899 n.8 (“No matter how a social media company chooses to moderate such content, the company will face backlash from its users and the public. That is true even if the company decides not to define the enumerated categories, because they will draw criticism for under-moderating their community.”). The deep-seated controversy inherently tied to each topic in Section 1102(1)(c) makes crystal clear that the disclosures compelled by the Content Category Report Provisions are not the sort of purely factual and uncontroversial disclosures that fit within the narrow exception to strict scrutiny set forth under *Zauderer v. Off. of Disciplinary Couns. of Sup. Ct. of Ohio*, 471 U.S. 626 (1985).

45. Because the Content Category Report Provisions “are content-based,” they are, at the very least, “subject to strict scrutiny.” *X Corp.*, 116 F.4th at 903 (citing *NIFLA*, 585 U.S. at 766); *see also Volokh v. James*, 656 F. Supp. 3d 431, 443 (S.D.N.Y. 2023) (striking down New York’s Hateful Conduct Law under strict scrutiny, finding that the law did not compel commercial speech, since it “compels a social media network to speak about the range of protected speech it

will allow its users to engage (or not engage) in”); *id.* (“Because the Hateful Conduct Law regulates speech based on its content, the appropriate level of review is strict scrutiny.”); *id.* at 445 (“Even though the law does not require social media networks to remove ‘hateful conduct’ from their websites and does not impose liability on users for engaging in ‘hateful conduct’, the state’s targeting and singling out of this type of speech for special measures certainly could make social media users wary about the types of speech they feel free to engage in without facing consequences from the state.”).

46. As the Supreme Court held in *NIFLA*, the “Court’s precedents do not permit governments to impose content-based restrictions on speech without persuasive evidence . . . of a long (if heretofore unrecognized) tradition to that effect.” 585 U.S. at 767; *see id.* at 766 (laws that “compel[] individuals to speak a particular message . . . alte[r] the content of [their] speech” and constitute “content-based regulation[s] of speech”).

47. The Content Category Report Provisions also impermissibly substitute the judgment of the government for that of the covered companies as to what content should remain on their platforms. *See Moody*, 603 U.S. at 740 (it is not a “valid, let alone substantial” interest for a state to seek “to correct the mix of speech” that “social-media platforms present”); *id.* at 742 (quoting *Pac. Gas & Elec. Co. v. Pub. Utilities Comm’n of California*, 475 U.S. 1, 20 (1986) (a “State ‘cannot advance some points of view by burdening the expression of others’”)); *id.* at 734 (the “government may not, in supposed pursuit of better expressive balance, alter a private speaker’s own editorial choices about the mix of speech it wants to convey”); *id.* at 719 (“it is no job for government to decide what counts as the right balance of private expression—to ‘un-bias’ what it thinks biased, rather than to leave such judgments to speakers and their audiences. That principle works for social-media platforms as it does for others”); *id.* at 733 (“[h]owever imperfect

the private marketplace of ideas,” a “worse proposal” is “the government itself deciding when speech [is] imbalanced, and then coercing speakers to provide more of some views or less of others”); *see also Wash. Post v. McManus*, 944 F.3d 506, 511–13, 519 (4th Cir. 2019) (striking down state law that required, in an effort to address foreign interference in U.S. elections, “online platforms,” within “48 hours of an ad being purchased,” to “display somewhere on their site the identity of the purchaser, the individuals exercising control over the purchaser, and the total amount paid for the ad,” and declaring the law “a compendium of traditional First Amendment infirmities” that would “chill speech”); *id.* at 515 (“the fact that [the law] is content-based” and “compels certain speech [] poses a real risk of either chilling speech or manipulating the marketplace of ideas”).

48. These concerns are especially pronounced where, as here, the government’s infringement of speech has the potential to impact multiple layers of speech—that is, not only the speaker’s First Amendment rights to make editorial judgments about what speech belongs on a given platform (e.g., a bookstore, movie theater, or here, a social media platform), but also the public’s right to access constitutionally protected speech that the government may not suppress directly. *See, e.g., Smith*, 361 U.S. at 154 (striking down, on First Amendment grounds, ordinance imposing liability on booksellers that would “tend to restrict the public’s access to forms of the printed word which the State could not constitutionally suppress directly . . . self-censorship, compelled by the State, would be a censorship affecting the whole public, hardly less virulent for being privately administered.”).

49. Even worse, the Content Category Report Provisions discriminate against speech based on viewpoint, which renders them *per se* invalid. *See Pleasant Grove City, Utah v. Summum*, 555 U.S. 460, 469 (2009) (finding that content-based restrictions “must satisfy strict

scrutiny,” but “restrictions based on viewpoint are prohibited”). The Content Category Report Provisions discriminate based on viewpoint because they do not compel disclosures about content-moderation practices about categories that are “positive,” but do compel such speech for categories perceived by the State to be negative—they compel disclosures about “hate speech,” but not “kind speech.” *See Iancu v. Brunetti*, 588 U.S. 388, 393 (2019) (quoting *Matal v. Tam*, 582 U.S. 218, 249 (2017) (Kennedy, J., concurring) (explaining that such differential treatment “reflects the Government’s disapproval of a subset of messages it finds offensive” and “is the essence of viewpoint discrimination”)). It is hornbook law that such “ideologically driven attempts to suppress a particular point of view are [at least] presumptively unconstitutional[.]” *Rosenberger v. Rector & Visitors of Univ. of Virginia*, 515 U.S. 819, 830 (1995).

50. The Content Category Report Provisions neither survive strict scrutiny, nor any lesser standard of review.⁹

51. The Content Category Report Provisions are facially invalid under the First Amendment because “a substantial number of [the law’s] applications are unconstitutional, judged in relation to the statute’s plainly legitimate sweep.” *Americans for Prosperity Found. v. Bonta*, 594 U.S. 595, 615 (2021); *see also X Corp.*, 116 F.4th at 904 (finding the Content Category Report

⁹ *See, e.g.*, Ex. 3 at 207–09 (*June 6, 2024 New York State Assembly Floor Debates*, 2023–24 Reg. Sess. (N.Y. 2024), available at <https://www2.assembly.state.ny.us/write/upload/transcripts/2023/6-6-24.pdf#search=%226789%22>) (Assembly Member John McGowan stating “I’m not really sure what the purpose of this bill is other than providing kind of vague government oversight of these social media platforms that could have absolutely chilling effect on free speech”); *id.* at 218 (Assembly Member Edward Flood stating “I’m still not sure . . . why we need to collect this data other than to potentially -- I mean, maybe to me it makes it more difficult for these companies to transact business in New York without really having a true reason for the data other than that our Attorney General may want to use this data against, you know, individuals or companies within New York”); *id.* at 217–20 (Assembly Member Edward Flood stating “I’m not sure why we need the data [called for by S895B]” and that it “doesn’t seem there’s a need for [the law] at all”).

provisions of AB 587 facially invalid under the First Amendment). They are also unconstitutional as applied to X Corp.

52. There is a *bona fide* and actual controversy between X Corp. and Defendant because Defendant is charged with enforcing, and intends to enforce, the Content Category Report Provisions, even though they violate the First Amendment of the United States Constitution and Article I, Section 8, of the New York Constitution, both facially and as-applied to X Corp.

53. X Corp. maintains that Content Category Report Provisions are illegal and unconstitutional. Defendant claims otherwise.

54. X Corp. requests a judicial determination regarding the validity of the Content Category Report Provisions to prevent the harm caused by their enactment and enforcement. Such a determination is both necessary and appropriate to avoid the deprivation of X Corp.'s and the other covered companies' constitutional rights, which would occur if the Content Category Report Provisions are applied to X Corp. or any other covered company.

55. Given the violation of the First Amendment of the United States Constitution and Article I, Section 8, of the New York Constitution, X Corp. seeks injunctive relief against enforcement of the Content Category Report Provisions. X and the other covered social media companies would be irreparably harmed if they were forced to comply with the Content Category Report Provisions and have no adequate remedy at law.

SECOND CAUSE OF ACTION

(Declaratory Relief and Preliminary and Permanent Injunctive Relief for Immunity Under and Preemption by 47 U.S.C. § 230(c)(2)(A))

56. X Corp. realleges and incorporates herein by reference each and every allegation set forth above.

57. 47 U.S.C. § 230(c)(2)(A) directly conflicts with, and thus preempts, the Content Category Report Provisions.

58. 47 U.S.C. § 230(e)(3) provides that “[n]o cause of action may be brought and no liability may be imposed under any State or local law that is inconsistent with this section.”

59. “Liability” under Section 230(e)(3) includes being subjected to the kind of financial penalties and suits for injunctive relief authorized by Section 1103.

60. X is an “interactive computer service,” as that term is defined under 47 U.S.C. § 230(f)(2).

61. The Content Category Report Provisions impose liability on X Corp. with regard to how it moderates content on X. This directly contravenes the immunity granted under Section 230(c)(2)(A), which provides that “[n]o provider or user of an interactive computer service shall be held liable on account” of “any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected[.]”

62. The immunity provided by Section 230(c)(2) is very broad. It precludes liability for “any action” taken voluntarily in good faith by an interactive computer service to restrict access to certain content. The Content Category Report Provisions impose such liability, however, if

actions are taken by interactive computer services to restrict access to content without making the disclosures required by the Content Category Report Provisions, to the State's satisfaction.

63. The phrase “any action” in Section 230(c)(2)(A) means “any action,” including any actions taken in good faith to restrict access to content without making the disclosures required by the Content Category Report Provisions, to the State's satisfaction.

64. The Content Category Report Provisions also contravene the immunity afforded to interactive computer service providers such as X Corp. by Section 230(c)(2)(A) because Defendant may penalize X Corp.—through significant monetary fines and suits for injunctive relief—if she determines, in her sole unfettered discretion, that X Corp.'s Content Category Report “materially omits or misrepresents required information[.]” *See* § 1103(b)(iii). Such imposition of financial penalty and injunctive relief constitutes liability for X Corp “restrict[ing] access” to content in a way that, in Defendant's view, is contrary to X Corp.'s promulgated content-moderation policies.

65. There is a *bona fide* and actual controversy between X Corp. and Defendant because Defendant is charged with enforcing, and intends to enforce, the Content Category Report Provisions, even though such enforcement is precluded and preempted by 47 U.S.C. § 230(c)(2)(A).

66. X Corp. maintains that the Content Category Report Provisions are unconstitutional under the Supremacy Clause and thus void as a matter of law. Defendant claims otherwise.

67. X Corp. seeks a declaratory judgment that the Content Category Report Provisions are legally invalid and unenforceable because they are precluded and preempted by 47 U.S.C. § 230(c)(2)(A).

68. In light of the violation of the Section 230(c)(2)(A), X Corp. seeks injunctive relief against enforcement of the Content Category Report Provisions. X Corp. would be irreparably

harm if it were forced to comply with, or litigate, the requirements thereof and has no adequate remedy at law.

PRAYER FOR RELIEF

WHEREFORE, X Corp. respectfully requests that this Court enter judgment in X Corp.'s favor and grant the following relief:

1. A declaration that the Content Category Report Provisions violate the First Amendment of the United States Constitution and Article I, Section 8, of the New York Constitution, both facially and as-applied to X Corp.;
2. A declaration that the civil penalties and injunctive relief imposed by the Content Category Report Provisions are precluded and preempted by 47 U.S.C. § 230(c)(2)(A) and are therefore null and void and have no legal effect;
3. A preliminary and permanent injunction enjoining Defendant and her employees, agents, and successors in office from enforcing the Content Category Report Provisions;
4. An award of fees, costs, expenses, and disbursements, including attorneys' fees, to which X Corp. is entitled pursuant to 42 U.S.C. § 1988 and other applicable law; and
5. Such other and further relief as the Court deems just and proper.

DEMAND FOR JURY TRIAL

Pursuant to Federal Rule of Civil Procedure 38, X Corp. demands a trial by jury in this action of all issues so triable.

Dated: June 17, 2025

By: /s/ Joel Kurtzberg
CAHILL GORDON & REINDEL LLP
Joel Kurtzberg (SBN NY 2835007)
Floyd Abrams (SBN NY 1758184)
Jason Rozbruch (SBN NY 5753637)
32 Old Slip
New York, NY 10005
Phone: 212-701-3120
Facsimile: 212-269-5420
jkurtzberg@cahill.com
fabrams@cahill.com
jrozbruch@cahill.com