No. 21-1333

IN THE

Supreme Court of the United States

————

REYNALDO GONZALEZ, ET AL.,

Petitioners,

v.

GOOGLE LLC,

Respondent.

————

**On Writ of Certiorari to
the United States Court of Appeals
for the Ninth Circuit**

————

**BRIEF OF AMICUS CURIAE CITP TECH
POLICY CLINIC
IN SUPPORT OF NEITHER PARTY**

————

JEREMY S. SPIEGEL
Counsel of Record
LAW OFFICE OF JEREMY SPIEGEL
ONE SOUTH BROAD ST. SUITE 1500,
PHILADELPHIA, PA 19107
(215) 609 3154

spiegel@jeremyspiegellaw.com

Counsel for Amicus Curiae

December 6, 2022

i
TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page

TABLE OF AUTHORITIES

TABLE OF AUTHORITIES

Page

STATEMENT OF INTEREST OF AMICUS CURIAE[1]

CITP Tech Policy Clinic is part of Princeton University's Center for Information Technology Policy ("CITP"), which works to better understand and improve the relationship between technology and society. The Clinic provides nonpartisan research, analysis, and commentary in the public interest. This brief is the product of a clinic project and reflects the views of researchers in computer science and policy who study the behavior of online platforms and the effect that they have on users and society at large.[2]

---

[1] This brief is respectfully submitted in support of neither party. Pursuant to Rule 37.6, counsel for amicus states that no counsel for a party authored this brief in whole or in part, and that no person or entity other than the amicus or its counsel made a monetary contribution to the preparation or submission of this brief. Petitioner and Respondent have filed blanket letters of consent with the Clerk of the Court.

[2] Nia Brazzell (Emerging Scholar), Klaudia Jaźwińska (Emerging Scholar), Varun Rao (Graduate Student in Computer Science), and Mihir Kshirsagar (Clinic Lead), contributed to the brief. This brief reflects the independent views of the researchers and does not represent the views of Princeton University.

SUMMARY OF ARGUMENT

This brief explains how online platforms make third-party content available to a particular audience. We hope this analysis will aid the Supreme Court's reading of the scope of the statutory immunity that Section 230 of the Communications Decency Act grants to providers of interactive computer services.

The abundance of information available on the internet has led to the greater importance of recommendation engines in connecting users to third-party content. Advances in machine learning and computational power have enabled large online platforms to develop sophisticated algorithms that curate a highly personalized content feed for users. These algorithms analyze information about content, contextual information, and user data, to deliver customized information to users and aid in content discovery.

In contrast to the modern dynamically generated internet experience, the picture of internet in 1996 is largely characterized by users requesting specific content from websites or having specific content pushed to them by other users. Section 230 is an artifact of this earlier era, however, and the treatment of online platforms as neutral distribution channels for static content does not translate as well to the current era. As a result, we respectfully submit that the Supreme Court should consider the salient features of recommendation engines in determining what platform conduct to immunize.

ARGUMENT

I. ONLINE PLATFORMS PLAY AN ACTIVE ROLE IN SHAPING AND PERSONALIZING THE CONTENT DELIVERED TO USERS.

Since the passage of the Communications Decency Act in 1996, the internet has become firmly entrenched in American society. We use the internet to conduct commerce and access a wide range of news, information, and entertainment. Online platforms such as Google, Twitter, Facebook, and TikTok play a critical role in society by shaping how third-party content is made available to users. But Section 230 of the Communications Decency Act is often seen as an obstacle to holding such platforms accountable for how they may improperly influence users.

Section 230 immunizes providers of "interactive computer service[s]" from claims that seek to treat the service provider "as the publisher or speaker of any information provided by another information content provider." 47 U.S.C. 230 §(c)(1). As several commentators have noted, this provision is part of a statutory scheme designed to encourage service providers that distribute content to moderate potentially offensive materials without the fear of liability for errors in judgment or lapses in procedures that are inevitable at the scale of the volume of interactions on the internet.[3] Of course, there are constitutional and other defenses available to platforms for claims concerning their promotion of

---

[3] Jeff Kosseff, *A User's Guide to Section 230, and a Legislator's Guide to Amending It (or Not)*, 37 Berkeley Tech. L.J. 2 (2022).

third-party content even without the Section 230 immunity. *New York Times Co. v Sullivan*, 376 U.S. 254 (1964).

Several large online platforms now rely on sophisticated recommendation engines to connect users to third-party content. Advances in machine learning and computational power allow service providers to curate a highly personalized content feed for users. Indeed, for platforms such as YouTube, it is reported that 70% of the total viewership time is driven by videos suggested by the site's recommendation algorithm.[4] These algorithms analyze the content (and associated metadata), contextual information, and user data–either voluntarily provided or collected through observations and inferences, to deliver customized information to users and aid in content discovery.

We sketch four potential approaches the Supreme Court could draw under the statute to identify when online platforms are not liable for the third-party content they deliver to users.

First, immunity is only available if the conduct involves content moderation systems that service providers use to remove, flag, or screen potentially objectionable material. This would include manual or automated techniques for identifying such material. We call this the "filtering" approach to immunity.

Second, at the opposite end of the spectrum, service providers are immunized from liability for any action

---

[4] Joan E. Solsman, *YouTube's AI is the Puppet Master over Most of What You Watch,* (Jan. 10, 2018, 10:05 a.m. PT) CNET, https://www.cnet.com/tech/services-and-software/youtube-ces-2018-neal-mohan/.

that involves the platform deciding what third-party content to make available to the user. We call this the "editorial" approach to immunity.

Third, service providers lose immunity if they make a material contribution to the creation or development of any illegal content. *Fair Housing Council of San Fernando Valley v. Roommates.Com, LLC*, 521 F.3d 1157 (9th Cir. 2008). This is the "material contribution" approach to immunity.

Fourth, the service provider is immunized when it treats third-party content "neutrally," or the same as any other third-party content. *Gonzalez v. Google*, 2 F.4th 871 (9th Cir. 2021). The Second Circuit describes it as when, "algorithms take the information provided by Facebook users and 'match' it to other users—again, materially unaltered—based on objective factors applicable to any content, whether it concerns soccer, Picasso, or plumbers." *Force v. Facebook, Inc.,* 934 F.3d 53, 70 (2d Cir. 2019). That is, the platform has not changed how it operates to promote the specific third-party content at issue. This is the "content neutral" approach to immunity.

Our goal in sketching these approaches is to motivate the discussion of the technical issues below. We do not recommend any specific approach to follow. Next, we describe our understanding of how these technical systems work in practice before discussing the implications for how these different lines might be drawn.

A. Content Moderation Systems

Professor James Grimmelman aptly defines content moderation systems as "the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse."[5] Content moderation systems generally have a process to review third-party content for objectionable material shared on online platforms including social media, discussion forums, search results, or email servers. A platform's definition of objectionable material is supplied by internal guidelines or mandated by law in the relevant jurisdictions. Reviews may take place before or after content is published. In the latter case, it may be done proactively by the platform or reactively in response to users' inputs.

The process for reviewing third-party content for objectionable material may take place through automated or manual approaches. Manual approaches involve humans who review the user-generated content and decide whether it violated the platform's guidelines. Automated approaches involve a combination of manually defined rules and heuristics as well as machine learning-based approaches to identify content that is deemed problematic. Typically, platforms employ a combination of the two approaches, where the automated systems assist human moderators in their review. The review process considers information that it extracts about the content as well as contextual information about the user generating the content.

Spam filters are one example of a widely deployed content moderation system. They are used to detect

---

[5] James Grimmelman, *The Virtues of Moderation*, 17 Yale L.J. & Tech 47 (2015).

unsolicited emails. Spam filters work by analyzing the content of the communication as well as information about the source of communication. They also consider how other users have interacted with the same content.

Another widely deployed example are content moderation tools platforms use to identify child sexual abuse material ("CSAM"). Platforms identify CSAM by hashing images or videos to generate unique digital fingerprints and then comparing those to hashes of known CSAM.[6] Platforms also proactively attempt to detect if users are searching for prohibited content. They also provide tools for other users to report suspicious content.

More recently, platforms have begun to moderate more value-laden issues by treating "low quality" or "borderline" content differently from "high quality" or "authoritative" content in the recommendations of third-party content to users.[7] Because recommendation engines are important to the discovery of third-party content, platforms use the ability to prioritize or de-prioritize content to aid their moderation practices.

---

[6] *Fighting Child Sexual Abuse Online* https://protectingchildren.google/ (last visited Dec. 5, 2022); *CSAM Detection Technical Summary* https://www.apple.com/child-safety/pdf/CSAM_Detection_Technical_Summary.pdf (last visited Dec. 5, 2022).

[7] *Information Quality & Content Moderation* https://blog.google/documents/83/information_quality_content_moderation_white_paper.pdf/ (last visited Dec. 5, 2022).

B. Recommendation Systems

Users constantly encounter recommender systems on the internet. These systems are used to suggest relevant content to users and help them navigate the abundance of data available on the internet. Rather than provide information specifically requested by users, services such as TikTok or YouTube increasingly use recommendation engines to determine the bulk of the content users encounter on their platforms.

The core task of recommendation systems is to show users items that are of interest to them. When users buy products from Amazon, stream shows on Netflix, listen to music on Spotify, scroll through a social media feed, or perform myriad other tasks online, recommender systems make suggestions based on their behavior. Typically, recommender systems are designed to learn about users' interests and serve them relevant content to keep them engaged on the platform.

The platform can decide whether to show organic recommendations for third-party content to display to users, or it can show sponsored content from advertisers.

For organic recommendations, the recommendation engine typically relies on three sources of signals to determine what content to show. The first is based on an analysis of the content (including metadata). The second is the context of the user (*e.g.*, their location, time of day, device). The third is information about the user. This last category is derived from inputs voluntarily provided by the user, such as their subscriptions to specific channels or prior "likes," as

well as information observed or inferred about the user's behavior.

Platforms can also inject sponsored ads into the user's feed or channel. These ads are content that an advertiser pays a platform to promote to relevant users. Advertising can be used to boost organic content or to run commercial messages to drive a specific marketing objective. Ads are shown to users as they browse through their feed or actively search for content.

For paid advertising, most large platforms rely on an auction mechanism for advertisers to bid on delivering an ad to the most relevant user. For example, several billion ad auctions occur each day across Meta properties (such as Facebook or Instagram) for a variety of ads including products and services, politics, and social issues, as well as employment, credit, and housing.[8] The ad auction is based on inputs from three parties: the advertiser, the platform, and the user.

An advertiser who wishes to promote their services to Facebook users is first allowed to select among several different ad campaign objectives based on the goal of the business–such as building awareness, generating traffic, or driving engagement, leads, or sales.[9] Next, they define a target audience through

---

[8] *About Ad Auctions*, Meta Business Help Center, https://www.facebook.com/business/help/430291176997542 (last visited Dec. 5, 2022).

[9] *An Update and Changes to Meta Ads Manager Objectives,* Meta Business Help Center, https://www.facebook.com/business/help/325793898950394 (last visited Dec. 5 2022).

selection tools that the platform provides of user interests, user location,[10] and personally identifiable information including name, date of birth, phone number and email addresses.[11] The advertiser also sets the budget and ad delivery schedule. Finally, the advertiser defines the ad creative content which includes the ad text, image(s), and destination page link. The decision of what content to show to a particular individual is left to the platform's discretion through an ad auction The ad auction determines the most relevant ad to show a user at a given point in time. At a high level, the winning ad is based on the following calculation:

Total Value = Advertiser Bid x Estimated Action Rate + User Value

To explain each of these terms, the "advertiser bid" is the amount set by platform (based on the budget set by the advertiser) for showing an ad. The "estimated action rate" is the probability that showing an ad to a user leads to the desired outcome for the advertiser. Finally, the "user value" is a measure of the creative quality of an ad for a specific user. Meta does not explain how it calculates this measure in more detail.

Other platforms employ similar auction-based systems to place ads in the user feeds. For Google, the main factors that govern which ads users see and their order of appearance include: (a) the advertiser's

---

[10] *About Detailed Targeting.* Meta Business Help Center, https://www.facebook.com/business/help/182371508761821 (last visited Dec. 5 2022).

[11] *How to Format Customer Lists,* https://www.facebook.com/business/help/2082575038703844 (last visited Dec. 5, 2022).

bid; (b) the ad quality (the relevance and usefulness of the ad and destination page to the user as measured by Google); and (c) the ad context, which is determined by the search terms the user has entered, the user's location at the time of the search, the type of device they're using (for example, mobile or desktop), the time of the search, the nature of the search terms, other ads and search results that show on the page, and other user signals and attributes.[12]

In addition, some platforms have begun to offer advertisers the ability to boost specific organic content in a user's feed by modifying the recommendation. Spotify, for example, offers a "discovery mode" tool that allows the advertiser to partner with them to promote certain songs recommended to users.[13]

C. Transparency Incentives

Online platforms are not transparent about the practical impact of their content moderation and recommender systems. There are limited incentives to share that information with the public if the platforms are not held accountable for their decisions.

For content moderation, the goal of the platform is to create an appealing space in which users can

---

[12] *How the Google Ads Auction Works*, https://support.google.com/google-ads/answer/6366577?hl=en (last visited Dec. 5, 2022).

[13] *Discovery Mode*, https://artists.spotify.com/en/discovery-mode (last visited Dec. 5, 2022).

interact.[14]   At times, the platform benefits from disclosing its moderation practices to show users that it is behaving responsibly. And, in some cases, users can monitor whether certain kinds of content are blocked or flagged. But generally, users have limited insight into how platforms are using moderation tools.

For recommender systems, the goal is to drive engagement with the content on the platform. When a user is served relevant content, they spend more time on the platform. This engagement can be monetized by selling space to advertisers. But the platforms benefit from keeping secret the details of how their algorithms work. And without platform disclosures, external observers struggle to document systematically how and why content has been promoted (or demoted) across the myriad personalized feeds.

II. IMPLICATIONS OF THE TECHNICAL FEATURES OF RECOMMENDATION SYSTEMS FOR STATUTORY IMMUNITY.

In Section I, we outlined four different approaches to granting statutory immunity under Section 230. We now discuss how these different approaches might fare based on the technical attributes of content moderation and recommender systems.

The most direct implication is that the "content neutral" approach is difficult to square with the how platforms decide what content to make available to

---

[14] Tarleton Gillespie, *Platforms are Not Intermediaries*, 2 Geo L. Tech Rev. 198, 201 (2018).

users. If, content neutrality is taken at a high level of generality, as suggested by the Second Circuit in *Force,* to imply that a platform is immune if it uses the same tools across different types of content, then most actions taken by a platform to recommend organic or sponsored content would be immunized. At a more granular level, however, the platform does promote content differently based on its analysis of the relevancy of that content to users. In other words, different content will be promoted differently based on the attributes of that content.

Another implication is that the "editorial" approach would immunize a wide variety of conduct. For example, a provider would be immunized from any allegation that it helped steer users to certain kinds of third-party content. It would also be immunized from any allegation that its personalization of the content delivery was discriminatory or otherwise problematic. But as discussed above, platforms use complex algorithmic tools to decide who sees content. Especially in a world where it is difficult for outside observers to examine the recommendation algorithms, the editorial approach grants significant power to the online platforms to operate without accountability.

The "filtering" approach suggests a relatively bright line the Court might draw between decisions to flag, remove or retain content (immunized) and decisions to recommend content (subject to further investigation). But one challenge is how to categorize the treatment of decisions by platforms to demote or deprioritize "borderline content." In this case, the platforms are essentially using the means of recommendation engines for content moderation ends.

Finally, the "material contribution" approach could capture potential problems with personalization that are associated with improperly targeted content. For example, if the platform's algorithm was responsible for delivering sponsored content in a discriminatory manner, or if the platform offered filters that facilitated illegal conduct, then the platform could lose its immunity. Such an approach could also be used to distinguish instances where the platform responds to an information request from a user, such as when a user conducts a search, but withholding that automatic immunity if the content is recommended to the user in a way that contributes to the illegality of the content.

CONCLUSION

The technical features of content moderation and recommendation systems have an important role in determining what content is made available to users. As the Court considers what platform conduct to immunize, we hope the salient commonalities and differences highlighted in this brief about these systems can aid the Court's decision.

Respectfully submitted,

JEREMY S. SPIEGEL
Counsel of Record

LAW OFFICE OF JEREMY SPIEGEL
ONE SOUTH BROAD ST. SUITE 1500,
PHILADELPHIA, PA 19107
(215) 609 3154

spiegel@jeremyspiegellaw.com

*Counsel for Amicus Curiae CITP
Tech Policy Clinic*