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
INFORMATION SCIENCE SCHOLARS
AS *AMICI CURIAE*
IN SUPPORT OF RESPONDENT

Michael S. Kwun
Counsel of Record
KWUN BHANSALI LAZARUS LLP
555 Montgomery Street
Suite 750
San Francisco, CA 94111
(415) 630-2350
mkwun@kblfirm.com

January 19, 2023

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Cases

Statutes

Other Authorities

INTEREST OF *AMICI CURIAE*[1]

Amici curiae are information science scholars. They submit this brief to provide the Court with a historical and technical perspective on recommender systems and their relationship to this case. Titles and affiliations are included for identification purposes only and do not indicate institutional endorsements.

Paul Resnick is the Michael D. Cohen Collegiate Professor of Information at the School of Information at the University of Michigan. He received the 2010 ACM (Association for Computing Machinery) Software Systems Award for his work on the GroupLens Collaborative Filtering Recommender System. He is an ACM Fellow, member of the ACM SIGCHI (Special Interest Group on Computer–Human Interaction) Academy, past chair of the steering committee for the ACM Conference on Recommender Systems, and was the co-chair of the Platform for Internet Content Selection at the World Wide Web Consortium (W3C) in the 1990s.

Jon Kleinberg is the Tisch University Professor of Computer Science and Information Science at Cornell University. He is an ACM Fellow and a member of the National Academy of Sciences and the National Academy of Engineering. In 2005, he was awarded a MacArthur Foundation Fellowship for his work revealing the deep structure of complex networks such

---

[1] No counsel for a party has authored this brief in whole or in part. No such counsel or a party made a monetary contribution intended to fund this brief's preparation or submission. No one other than *amici curiae* and their counsel has made such a monetary contribution.

as genomes or computer networks, and creating new methods to extract the information embedded in them.

Karen Levy is an Associate Professor of Information Science at Cornell University, and an associate member of the faculty at Cornell Law School. She has been a New America National Fellow, a postdoctoral fellow at New York University School of Law's Information Law Institute, and a fellow at the Data and Society Research Institute.

Thorsten Joachims is a Professor of Computer Science and Information Science at Cornell University. He is an ACM Fellow, AAAI (Association for the Advancement of Artificial Intelligence) Fellow, and a member of the ACM SIGIR (Special Interest Group on Information Retrieval) Academy.

Maarten de Rijke is a Distinguished University Professor of Artificial Intelligence and Information Retrieval at the University of Amsterdam, and the scientific director of the national Innovation Center for Artificial Intelligence. He is a member of the Royal Netherlands Academy of Arts and Sciences and a member of the ACM SIGIR Academy.

ChengXiang Zhai is a Donald Biggar Willett Professor in Engineering in the Department of Computer Science at the University of Illinois at Urbana-Champaign. He is an ACM Fellow and a member of the ACM SIGIR Academy, which in 2021 awarded him the Gerard Salton Award for his foundational work in information retrieval.

Jure Leskovec is an Associate Professor of Computer Science at Stanford University, where he is a member of the InfoLab and the AI Lab. He is the recipient of an Alfred P. Sloan Fellowship, and was awarded the 2015 Lagrange Prize for his work on social and information networks.

SUMMARY OF ARGUMENT

In enacting the Communications Decency Act, Congress found that the internet and the services it enables "represent an extraordinary advance in the availability of educational and informational resources to our citizens." 47 U.S.C. § 230(a)(1). Yet access to all the world's information has created a need for ways to find the *right* information.[2] Section 230 encourages that process by providing certain immunities to interactive computer services when they provide users with access to information from others on their servers. *See* 47 U.S.C. § 230(c).

Section 230 provides that one type of interactive computer service is an "access software provider," which it defines to include providers of tools that filter, pick, choose, search, subset, or organize content. *Id.* § 230(f)(2), (f)(4). Recommender systems do just that, by helping to identify content that might serve a user's needs. These systems have long sought to provide personalized services, for example by relying on the interests of a particular user, their past actions, or the preferences of other users who appear to have similar interests. Work on such systems predates the Communications Decency Act, and modern recommender systems are based on similar principles. Because recommender systems filter, pick, and choose content to recommend to a user, the providers of the recommendations are access software providers.

---

[2] *Cf.* BARRY SCHWARTZ, THE PARADOX OF CHOICE: WHY MORE IS LESS (2004).

Moreover, contrary to Petitioners' position, Section 230 does not draw a distinction between computer systems that rely on explicit user requests for information and those that rely on implicit requests via a user's actions. Many common functions protected by Section 230 are based on implicit requests. In a directly analogous way, YouTube provides users with access to its computer servers when it responds to the signals contained in users' actions, regardless of whether they have formulated explicit search queries. Under the standard understanding of client-server architectures, a computer can act as a server even when a human user is not making explicit requests to it, and acts as an interactive "server" whenever it is receiving requests from a "client" computer program, which in the case of YouTube could be a web browser or a smartphone app.

Finally, search engines, too, provide recommendations. There is fundamentally no distinction between the rankings that search engines perform and the operations that recommendation systems perform: by ranking the search results provided in response to a query, a search engine recommends some results more highly than others. And at times the recommendations in these search rankings may depend on information that isn't part of the user's query. Nothing in Section 230, or in the way these systems are designed, supports distinguishing between liability for recommendations made by search engines and recommendations made by YouTube.

ARGUMENT

I. Recommender systems are a paradigmatic way of filtering, picking, choosing, searching, subsetting, or organizing content that a user may want to view or consume.

Congress enacted Section 230 of the Communications Act of 1934, commonly called Section 230 of the Communications Decency Act,[3] in support of the policy of the United States "to promote the continued development of the Internet and other interactive computer services and other interactive media." 47 U.S.C. § 230(b)(1). Section 230 defines an "interactive computer service" to include an "access software provider" that enables computer users to access internet servers. *Id.* § 230(f)(2). An "access software provider," in turn, is defined to include a provider of software or enabling tools that "filter," "pick," "choose," "search, subset[, or] organize" content. *Id.* § 230(f)(4).

Recommender systems—systems that use algorithms to filter by picking, choosing, searching, subsetting, or organizing content that a user may be more likely to want—have a long and rich history. Even in the 1980s, well before Section 230 was enacted in 1996,[4] it was "already a common experience in mature computer-based messaging communities for people to feel flooded with large quantities of electronic 'junk

---

[3] *See* U.S. Br. 1 n.1.

[4] Communications Decency Act of 1996, Pub. L. No. 104-104, tit. V, § 509, 110 Stat. 133, 137–39.

mail.'"[5] It was similarly recognized at this time that a priority was not to "just reduce the flow of 'junk mail,' but to dramatically increase the amount of useful information that can be exchanged electronically without leading to information overload."[6] In response to these types of concerns, information scientists developed systems to help "filter, sort, and prioritize" messages users had already received, and also to help "find useful messages they would not otherwise have received"—i.e., to recommend useful content to them.[7]

Fishwrap, a "prototype electronic newspaper" that was launched to the MIT community in 1993, was one such system.[8] Fishwrap provided a customized newspaper "with an egocentric [i.e., personalized] window into world affairs, allowing [readers] to receive news from their home town as well as stories of personal interest."[9] The front page for the newspaper included stories that were ranked "based on the number of people who actually accessed the article."[10] Notifications for articles of interest to a particular reader could also appear in a special area of the screen.[11] The system adjusted its news recommendations based on past con-

---

[5] Thomas W. Malone, Kenneth R. Grant, Franklyn A. Turbak, Stephen A. Brobst & Michael D. Cohen, *Intelligent Information-Sharing Systems,* 30 COMM. ACM 390, 390 (1987).

[6] *Id.*

[7] *Id.*

[8] *See* Pascal R. Chesnais, Matthew J. Mucklo & Jonathan A. Sheena, *The Fishwrap Personalized News System,* PROC. IEEE 2ND INT'L WORKSHOP ON COMMUNITY NETWORKING 275 (1995).

[9] *Id.* at 275.

[10] *Id.* at 276.

[11] *Id.* at 277.

duct, placing topics in which a reader had previously shown interest "closer to the top of their paper."[12] "Every click of the mouse can have an impact on the reader's and others' news presentation."[13] These and other systems demonstrated the functionality that recommender systems provided for filtering, picking, choosing, searching, subsetting, and organizing in the years prior to Section 230.

Fundamentally, recommender systems are probabilistic information retrieval models: attempts to rank based on "the probability that a user's information need is satisfied given a particular object."[14] A recommender system uses an algorithm that takes as its input a set of signals that describe aspects of the content, the user, and the context. The system then produces scores for the items in the current context. Those scores are then used to determine what to recommend to the user.

One approach might be to infer the value of unviewed content based on prior data consumption patterns.[15] For example, if Sally has watched a lot of cat videos, recommending more cat videos probably makes sense.[16] Likewise, a recommender system could rely on the intuitive idea that "people who agreed in

---

[12] *Id.* at 280.

[13] *Id.*

[14] Nicholas J. Belkin & W. Bruce Croft, *Information Filtering and Information Retrieval: Two Sides of the Same Coin?,* 35 COMM. ACM 29, 33 (1992).

[15] *See id.*

[16] *See* Resp. Br. 12 (explaining that YouTube recommends videos "primarily based on what viewers with similar YouTube browsing histories watched").

their subjective evaluation of past articles are likely to agree again in the future."[17] So if Bill has watched many videos that were watched by Jane, it might make sense to recommend that he watch other videos she has seen but that he has not.

Over the years, the set of signals that recommender systems use has grown, and the specifics of the algorithms for combining them have become more sophisticated. "Collaborative filtering," for example, can use signals about other people's reactions to content items. Location has become available more often as a contextual signal. And individuals' past behaviors, such as search histories, have become more available as part of the user signals on which recommendations can be based.

Modern recommender systems, such as the one used by YouTube, thus may be more sophisticated than early systems. Today's systems nonetheless build on prior efforts, and their fundamental structure has remained unchanged since well before the enactment of Section 230.

Recommender systems fall within the core of Section 230 because they filter, pick, choose, search, subset, and organize content. A provider of such software, including YouTube, is an access software provider. *See* 47 U.S.C. § 230(f)(4).

---

[17] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom & John Riedl, *GroupLens: An Open Architecture for Collaborative Filtering of NetNews,* 1994 PROC. ACM ON COMPUTER SUPPORTED COOPERATIVE WORK 175, 176.

II. YouTube computers that provide recommendations are servers regardless of whether users submit queries.

Petitioners seek to draw a distinction between search engine results and YouTube's recommendations, noting that search results respond to search queries, while YouTube responds to implicit signals. *See* Pet. Br. 44. The government disagrees, noting that "the salient point is that the algorithms simply direct to particular users videos that were created and developed without YouTube's involvement," not whether the recommendations are in response to "specific user queries." U.S. Br. 30.

Nothing in Section 230 supports distinguishing between implicit and explicit requests for purposes of deciding liability. Many common functions protected by Section 230 do not depend on an explicit query from the user. For example, malware protection and spam filters run automatically in the background, rather than requiring user queries. Section 230 nonetheless precludes civil liability for actions taken in good faith by an interactive computer service to "restrict access to" material that is considered to be "objectionable, whether or not such material is constitutionally protected." 47 U.S.C. § 230(c)(2). This provision immunizes good faith actions by programs that "filter adware and malware." *Zango, Inc.* v. *Kaspersky Lab, Inc.,* 568 F.3d 1169, 1173–74 (9th Cir. 2009). By its plain terms, it also immunizes an email provider's spam filters. *Holomaxx Techs.* v. *Yahoo!, Inc.,* No. 10-cv-4926-JF, 2011 WL 865794, at *4–5 (N.D. Cal. Mar. 11, 2011).

Section 230's definition of "interactive computer service" thus does not turn on whether the user first submits a query. *See* 47 U.S.C. § 230(f)(2). Instead, the definition hinges on the provision of access to a computer server. *See id.*

As Petitioners recognize, a "server" is a computer that executes software that enables multiple users to access information accessible to the server. *See* Pet. Br. 45. Yet Petitioners argue that if "a website's computer sends a user (by whatever method) material that the user has not requested, that computer is not operating as a 'server' within the meaning of section 230(f)(2)." *Id.* at 46.

From a computer scientist's perspective, Petitioners' position fails because the client-server paradigm does not depend on whether a *user*—a natural person—has submitted a query. Instead, the *client* is a program running on a computer, and the *server* responds to requests by that client computer program.[18]

When YouTube sends recommendations, its computer that does so is a server that is responding to a client program. "When a user directs her browser to the youtube.com website, or opens the YouTube app on an Internet-enabled smartphone, YouTube has provided the user with access to its server." U.S. Br. 33. Petitioners' argument that it matters whether YouTube provides recommendations in response to queries does not withstand scrutiny.[19]

---

[18] *See, e.g.,* Alok Sinha, *Client-Server Computing,* 35 COMM. ACM 77, 78 (1992).

[19] Moreover, YouTube is recommending videos, not the computer location pointers (URLs) that identify how the videos can be accessed. It should not matter for purposes of liability that the computer location pointers are YouTube URLs, *see* Pet. Br. 40,

III. Section 230 should not treat recommender systems differently from search engines, which rely on the same techniques.

When search engines return results to a user, they rely on the same techniques that recommender systems use. To respond to a search query, a search engine must make two determinations: (1) what content is responsive to the query; and (2) which content is most likely to be useful to the user.[20] For search queries with an abundance of responsive results, the latter problem predominates, because "[t]he number of pages that could reasonably be returned as relevant is far too large for a human user to digest."[21] The key problem for these queries isn't figuring out what content matches the query, but how best to rank the many responsive results.

Deciding how to rank web pages in response to the query "best pizza near me," for example, necessarily involves recommending which web pages are more or less likely to be of interest to the user. Every time a search engine ranks results, it is providing recommendations.

---

any more than it should matter whether someone recommends reading a newspaper on their desktop rather than a copy at a newsstand. *See* U.S. Br. 33 ("A website does not act as an information content provider by taking the technical steps necessary to render user-generated online content visible to others.").

[20] *See* Jon M. Kleinberg, *Authoritative Sources in a Hyperlinked Environment,* 46 J. ACM 604, 605–06 (1999).

[21] *Id.* at 606 (formatting modified); *see also* Resp. Br. 32.

Moreover, the "best" page might depend on context—including context that isn't part of the search query.[22] Someone who searches for "best pizza near me" likely isn't interested in pizzerias in another state. In response, even though the query does not identify where the user is located, a search engine may rely on implicit indicators of a likely location. If the search is being made on a modern smartphone, the app might have access to the user's geographic coordinates as determined by a GPS chip in the phone. For a query from a desktop computer, a search engine might try to estimate a location based on the IP address of the client web browser.

In some cases, a search engine might attempt to personalize results based on location even without a hint like "near me" in the query. For example, it is easy to verify that the one-word query "pizza" on a modern search engine will also return results for pizza near you; in effect the "near me" part of the query is assumed even when it is not provided.

Nor is this type of targeting based on location specific to a narrow set of transactional searches. For a user on the University of Michigan network on a fall Saturday, the best page for a query about football might be about the Michigan Wolverines; on a Sunday in Massachusetts, the best page might be about the New England Patriots; in Europe, a page about Man-

---

[22] *See, e.g.,* Yumao Lu, Fuchun Peng, Xing Wei & Benoit Dumoulin, *Personalize Web Search Results with User's Location,* 2010 PROC. ACM SPECIAL INT. GRP. ON INFO. RETRIEVAL 763, 763 (discussing "implicit local intent" queries).

chester United or another Premier League team might make more sense.[23]

A savvy search engine likewise could consider a user's past searches to improve search results.[24] For example, a search for "python" is ambiguous, and could refer to a snake or a programming language.[25] But someone who has previously often searched for information about programming or clicked through to the website for the Python language is more likely to be interested in results about the language than those about snakes.[26]

More generally, a modern search engine is designed to produce results that do not necessarily match the specific terms in a user's query if its algorithms infer that these results are likely to be relevant to the user's interests based on the context of the search, user characteristics such as location, and the user's history of actions. Algorithms for matching search results based on the user's actions[27] or other

---

[23] *Cf.* Resp. Br. 12 (noting that YouTube's recommendations may be based in part on a viewer's "location" or the "time of day").

[24] *See, e.g.,* Bin Tan, Xuehua Shen & ChengXiang Zhai, *Mining Long-Term Search History to Improve Search Accuracy,* 2006 PROC. ACM SPECIAL INT. GRP. ON KNOWLEDGE DISCOVERY & DATA MINING 718, 718.

[25] *Id.*

[26] *Id.*

[27] *See, e.g.,* J.J. Rocchio, *Relevance Feedback in Information Retrieval, in* INFO. STORAGE & RETRIEVAL: SCI. REPORT NO. ISR-9 TO THE NAT'L SCI. FOUND. § XXIII (Gerard Salton ed., 1965).

users' actions[28] date back to the foundations of the field of information retrieval in the 1960s. Neither the text of Section 230 nor sound policy supports distinguishing this type of functionality from the task that a recommendation system is performing, which is also to infer relevant results from context and users' actions.

Moreover, the notion that certain user actions constitute "queries" and others do not is essentially impossible to specify in any consistent way. For example, when a user performs an image search to locate similar images to a specific picture they find on the internet, we might think of the image as the query. But this is essentially no different from the functionality that YouTube's video recommendations offer, which is to locate similar videos based on the videos a user has found so far.

Similarly, speech interfaces to mobile or in-home devices are an increasingly common modality for issuing requests for information, and the types of voice commands that are often issued to these devices further blur the definition of what it means to be responsive to a query. Saying "play music" to a voice assistant or smartphone could be viewed as issuing a query, but it will result in a playlist of music determined by a recommendation system, just like telling your browser to take you to youtube.com will result in

[28] *See, e.g.,* M.C. Davis, M.D. Linsky & M.V. Zelkowitz, *A Relevance Feedback System Employing a Dynamically Evolving Document Space, in* INFO. STORAGE & RETRIEVAL: SCI. REPORT NO. ISR-14 TO THE NAT'L SCI. FOUND. § X (Gerard Salton ed., 1968).

a playlist of videos determined by a recommendation system.

There seemingly is no dispute that search engines are protected by Section 230. *See* Pet. Br. 47; Resp. Br. 22; U.S. Br. 13; *see also* 47 U.S.C. § 230(f)(4)(C) (including "search" within the definition of an access software provider). Nothing in the design of these systems or the text of Section 230 suggests that recommendations by search engines should be treated differently from recommendations by YouTube.

CONCLUSION

For the foregoing reasons, *amici curiae* urge the Court to affirm the judgment of the court of appeals.

Respectfully submitted,

Michael S. Kwun
Counsel of Record
KWUN BHANSALI LAZARUS LLP
555 Montgomery Street
Suite 750
San Francisco, CA 94111
(415) 630-2350
mkwun@kblfirm.com

January 19, 2023