#### In the

# Supreme Court of the United States

ASHLEY MOODY, ATTORNEY GENERAL OF FLORIDA, et al..

Petitioners,

υ.

NETCHOICE, LLC, DBA NETCHOICE, et al., Respondents.

NETCHOICE, LLC, DBA NETCHOICE, et al., Petitioners,

υ.

KEN PAXTON, ATTORNEY GENERAL OF TEXAS, Respondent.

On Writs of Certiorari to the United States Courts of Appeals for the Fifth and Eleventh Circuits

# BRIEF OF AMICI CURIAE CENTER FOR SOCIAL MEDIA AND POLITICS AT NEW YORK UNIVERSITY ET AL. IN SUPPORT OF NEITHER PARTY

Jacob M. Karr
Counsel of Record
Jason M. Schultz
TECHNOLOGY LAW AND
POLICY CLINIC
NEW YORK UNIVERSITY
SCHOOL OF LAW
245 Sullivan Street
New York, NY 10012
(212) 998-6042
jacob.karr@law.nyu.edu

Counsel for Amici Curiae

## TABLE OF CONTENTS

'ABLE OF AUTHORITIESi
NTEREST OF AMICI CURIAE
SUMMARY OF ARGUMENT4
ARGUMENT
I. Independent social science research plays a critical role in helping the public and policymakers understand the effects of social media on democracy, politics, and society
II. Social media platforms impose significant limitations on access to the data needed to conduct rigorous research, improve public knowledge, and develop effective public policy.
A. Platforms unilaterally control access to their data10
B. Platforms voluntarily disclose insufficient information to the public15
C. Independent researcher access to data can be restrictive, incomplete, and subject to withdrawal at any time, for any reason, and with little recourse16
III. Mandating meaningful platform transparency is the only way to ensure that independent researchers, the public, and policymakers will be able to answer urgent questions about how platforms operate and influence contemporary society24
CONCLUSION

### TABLE OF AUTHORITIES

Cases
Packingham v. North Carolina, 582 U.S. 98 (2017)7
Other Authorities
A.W. Geiger & Leslie Davis, A Growing Number of American Teenagers – Particularly Girls – Are Facing Depression, Pew Rsch. Ctr. (Jul. 12, 2019)
About the Twitter API, X20
$Advancing\ the\ State\ of\ the\ Art,\ Google\ Research10$
Aleksandra Urman & Mykola Makhortykh, How Transparent Are Transparency Reports? Comparative Analysis of Transparency Reports Across Online Platforms, Telecomms. Pol'y, Jan. 6, 2023, at 1
Alexandra A. Siegel et al., Trumping Hate on Twitter? Online Hate Speech in the 2016 U.S. Election Campaign and its Aftermath, 16 Q.J. Pol. Sci. 71 (2021)
Alexandra A. Siegel, Online Hate Speech, in Social Media and Democracy 56 (Nathaniel Persily & Joshua A. Tucker eds., 2020)14
Andrew M. Guess et al., How Do Social Media Feed Algorithms Affect Attitudes and Behavior in an Election Campaign?, 381 Science 398 (2023)1, 8
Bobby Allyn, Montana Becomes 1st State to Approve a Full Ban of TikTok NPR (Apr. 14, 2023) 26

Bobby Allyn, States Sue Meta, Claiming Instagram, Facebook Fueled Youth Mental Health Crisis, NPR (Oct. 24, 2023)
Chris Stokel-Walker, Twitter's \$42,000-per-Month API Prices Out Nearly Everyone, Wired (Mar. 10, 2023)
Ctrs. for Disease Control & Prevention, Youth $Risk$ Behavior Survey (2023)27
Daphne Keller, Platform Transparency and the First Amendment, 4 J. Free Speech L. 1 (2023)29
David Lazer et al., Computational Social Science, 323 Science 721 (2009)9
David Lazer et al., Meaningful Measures of Human Society in the Twenty-First Century, 595 Nature 189 (2021)
Deen Freelon, Computational Research in the Post-API Age, 35 Pol. Commc'n 665 (2018)22
Exec. Bd., Twitter's New API Plans Will Devastate Public Interest Research, Coal. Indep. Tech. Rsch. (Apr. 3, 2023)
Eytan Bakshy, Solomon Messing & Lada A. Adamic, Exposure to Ideologically Diverse News and Opinion on Facebook, 38 Science 1130 (2015)
Facebook Community Standards, Meta11
Ferenc Huszár et al., Algorithmic Amplification of Politics on Twitter, Proc. Nat'l Acad. Scis. U.S., Dec. 21, 2021, at 1

iv
Georgia Wells et al., Facebook Knows Instagram Is  Toxic for Teen Girls, Company Documents  Show, Wall St. J. (Sept. 14, 2021)12
Gregory Eady et al., Exposure to the Russian Internet Research Agency Foreign Influence Campaign on Twitter in the 2016 US Election and Its Relationship to Attitudes and Voting Behavior, Nature Commc'ns, Jan. 9, 2023, at 1
Gregory Eady et al., How Many People Live in Political Bubbles on Social Media? Evidence from Linked Survey and Twitter Data, SAGE Open, JanMar. 2019, at 1
Introducing the Ad Archive API, Meta (Aug. 22, 2018)19
Irene V. Pasquetto et al., Tackling Misinformation:  What Researchers Could Do with Social Media  Data, Harvard Kennedy Sch.: Misinformation  Rev. (Dec. 9, 2020)
Jeff Horwitz, Facebook Says Its Rules Apply to All.  Company Documents Reveal a Secret Elite That's Exempt., Wall St. J. (Sept. 13, 2021)12
Jennifer Allen et al., Research Note: Examining Potential Bias in Large-Scale Censored Data, Harvard Kennedy Sch.: Misinformation Rev. (July 26, 2021)
Jeremy B. Merrill & Ariana Tobin, Facebook Moves to Block Ad Transparency Tools—Including Ours ProPublica (Jan. 28, 2019)

Joe Bak-Coleman, <i>TikTok's API Guidelines Are a Minefield for Researchers</i> , Tech Policy Press (Feb. 22, 2023)
John Albert, Platforms' Promises to Researchers:  First Reports Missing the Baseline, AlgorithmWatch (Feb. 16, 2023)24
Jonathan Nagler & Joshua A. Tucker, <i>The Social Media Data We Need to Answer Key Research Questions</i> , CSMaP (May 4, 2022)24
Josh Constine, Zuck Says Ad Transparency Regulation Would Be 'Very Good If Done Well,' TechCrunch (Nov. 1, 2017)20
Joshua A. Tucker, Computational Social Science for Policy and Quality of Democracy, in Handbook of Computational Social Science for Policy 381 (Eleonora Bertoni et al. eds., 2023)9
Kevin Roose, Inside Facebook's Data Wars, N.Y Times (July 14, 2021)19
Laura Edelson & Damon McCoy, We Research Misinformation on Facebook. It Just Disabled Our Accounts, N.Y. Times (Aug. 10, 2021)22
Laura Edelson, Platform Transparency Legislation: The Whos, Whats and Hows, Lawfare (Apr. 29, 2022)
Malcolm R. Parks, Big Data in Communication Research, 64 J. Commc'n 355 (2014)9
Matt Richtel et al., Surgeon General Warns that Social Media May Harm Children and Adolescents, N.Y. Times (May 23, 2023)27

Matti Vuorre & Andrew K. Przybylski, Global Well- Being and Mental Health in the Internet Age, Clinical Psych. Sci., Nov. 27, 2023, at 128
Max Zahn, No Evidence of TikTok National Security Threat but Reason for Concern, Experts Say, ABC News (Mar. 28, 2023)
Megan A. Brown, The Problem with TikTok's  New Researcher API Is Not TikTok,  Tech Policy Press (Mar. 1, 2023)
Mia Sato, Researchers Will Get Access to TikTok  Data - Pending Company Approval, Verge  (Feb. 21, 2023)
Mike Schroepfer, Update on Our Progress on AI  and Hate Speech Detection,  Meta (Feb. 11, 2021)14
Monika Bickert, Working to Keep Facebook Safe, Meta (July 17, 2018)11
Monika Zalnieriute, "Transparency Washing" in the Digital Age, 8 Critical Analysis L. 139 (2021)
Nathaniel Persily & Joshua A. Tucker, Conclusion:  The Challenges and Opportunities for Social  Media Research, in Social Media and  Democracy 313 (Nathaniel Persily &  Joshua A. Tucker eds., 2020)7, 8, 9, 18, 25, 26
Nicolas Kayser-Bril, AlgorithmWatch Forced to Shut Down Instagram Monitoring Project After Threats from Facebook, AlgorithmWatch (Aug. 31, 2021)

(July 11, 2022)16
Reports, TikTok
Research Partnership to Understand Facebook and Instagram's Role in the U.S. 2020 Election, Meta
Research Platform Addendum, Meta (Nov. 4, 2021)16
Research, Meta10
Sandra González-Bailón et al., Asymmetric Ideological Segregation in Exposure to Political News on Facebook, 381 Science 392 (2023)22
Spandana Singh & Leila Doty, The Transparency Report Tracking Tool: How Internet Platforms Are Reporting on the Enforcement of Their Content Rules, New America: Open Tech. Inst. (Dec. 9, 2021)14, 15
Svea Windwehr & Jillian C. York, <i>Thank You For Your Transparency Report, Here's Everything That's Missing</i> , Elec. Frontier Found. (Oct. 13, 2020)
Teresa Carr, Why Researchers Want Broader Access to Social Media Data, Nieman Lab (May 4, 2022)
The Facebook Files, Wall St. J11
TikTok Research API Services Terms of Service, TikTok (Aug. 10, 2023)16, 17
Transparency Reports Meta 13

## viii

Trans	sparency Rep	orts,	X			13			
Zeve Sanderson et al., Feedback on EU Article 40, CSMaP (May 23, 2023)29									
	Sanderson		,						
C	omments on A	AI, (	CSMa	P (July	7, 2023)	28			

#### INTEREST OF AMICI CURIAE1

The Center for Social Media and Politics at New York University (CSMaP) is an academic research institute that studies the impacts of social media and other digital technologies on society.<sup>2</sup> CSMaP works to strengthen democracy in the digital age by conducting rigorous research, advancing evidencebased public policy, and training the next generation of scholars. Founded in 2019, CSMaP leverages cutting-edge data collection infrastructure and big data analysis to interrogate assumptions about social media and engage directly with policymakers, civil society groups, industry professionals. and journalists.

CSMaP experts perform computational social science research at scale to better understand how media politics, social shapes often belving conventional wisdoms shared by those across the political spectrum. See, e.g., Andrew M. Guess et al., How Do Social Media Feed Algorithms Affect Attitudes and Behavior in an Election Campaign?, 381 Science 398, 402 (2023) (finding that personalized social media feed algorithms did not significantly increase political polarization during the 2020 U.S. election); Alexandra A. Siegel et al., Trumping Hate on Twitter? Online Hate Speech in the 2016 U.S.

<sup>&</sup>lt;sup>1</sup> No counsel for any party authored this brief in whole or in part, and no person or entity other than *Amici* or their counsel made a monetary contribution to fund the preparation or submission of the brief. Sup. Ct. R. 37.6.

<sup>&</sup>lt;sup>2</sup> The views expressed herein do not necessarily represent those of New York University.

Election Campaign and its Aftermath, 16 Q.J. Pol. Sci. 71, 74 (2021) (providing systematic evidence that hate speech did not increase on Twitter during the 2016 U.S. election).

Darren Linvill and Patrick Warren study online deception and the behavior of coordinated inauthentic influence operations on social media, especially those operated by authoritarian states or state-backed actors.<sup>3</sup> They also lead the Media Forensics Hub, an interdisciplinary research project of the Watt Family Innovation Center, with a team of researchers working to study and combat online deception with the goal of building society's resilience to the dangers it poses.

Filippo Menczer leads Indiana University's Observatory on Social Media (OSoMe).<sup>4</sup> OSoMe's mission is to transform the study of coupled media and technology networks that drive the diffusion of information; offer access to data and tools to

<sup>&</sup>lt;sup>3</sup> Darren Linvill is a Professor of Communication and Co-Director of the Watt Family Innovation Center Media Forensics Hub at Clemson University. Patrick Warren is an Associate Professor of Economics and Co-Director of the Watt Family Innovation Center Media Forensics Hub at Clemson University. Affiliations are provided for identification purposes only, and the views expressed herein do not necessarily represent those of Clemson University.

<sup>&</sup>lt;sup>4</sup> Filippo Menczer is the Luddy Distinguished Professor of Informatics and Computer Science and Director of the Observatory on Social Media at Indiana University. Affiliation is provided for identification purposes only, and the views expressed herein do not necessarily represent those of Indiana University.

investigate the diffusion of (mis)information, uncover the vulnerabilities of the media ecosystem, and develop methods for increasing the resilience of citizens and democratic systems to manipulation; and train a generation of media professionals, enabling them to employ computational skills for fulfilling the traditional watchdog function of journalism.

As social science researchers at American universities who study social media, technology, and society, *Amici* rely on access to social media platform data to perform their critical work. See, e.g., Gregory Eady et al., How Many People Live in Political Bubbles on Social Media? Evidence from Linked Survey and Twitter Data, SAGE Open, Jan.-Mar. 2019, at 1, 18–19 ("Eady et al., Political Bubbles") (analyzing roughly 1.2 billion tweets from over 640,000 Twitter accounts); Siegel et al., supra, at 74 (analyzing over 750 million tweets related to the 2016 U.S. election, in addition to almost 400 million tweets from a random sample of American Twitter users). But the platforms impose significant limitations on Amici's ability to access and analyze much of the data needed to answer urgent questions about the influence of social media on society. As social media and other digital technologies become increasingly vital to politics and public life, *Amici* have an interest in ensuring that they have access to platform data to conduct independent research that informs public discourse and policymaking. Amici respectfully submit this brief to provide the Court with the professional expertise of faculty and scholars and with broader context for understanding the issues of social media platform transparency raised by these cases.

#### SUMMARY OF ARGUMENT

Over the past two decades, social media platforms and other digital technologies have transformed society. These technologies have made it easier than ever to find information, engage with politics, and connect with people across the globe. But the public and policymakers have also raised concerns about the role that platforms play in spreading misinformation, enabling harassment, and contributing to polarization. The need to understand how platforms operate and what influence they have on politics, policy, and democracy has never been more urgent.

Independent social science research has played a critical role in helping the public and policymakers understand the wide-ranging effects of these technologies. Studies have provided insights into the recommendations of algorithmic systems, of foreign influence campaigns, patterns the relationship between social media and political behavior and beliefs, the prevalence of hate speech and harassment, and the efficacy of interventions. Yet much of this research can only be conducted by analyzing massive amounts of social media data, since assessing how these platforms shape society requires access to data on a scale commensurate with the unprecedented complexity and scope of today's online information ecosystems.

Social media platforms, however, unilaterally control and limit access to their data, erecting significant barriers to rigorous research. Voluntary public disclosures like "transparency reports" contain insufficient information. Direct researcher access is

inconsistently granted, and when it is, that access can be restrictive, incomplete, and subject to withdrawal at any time, for any reason, and with little recourse. Only some platforms have historically provided researcher access to data at all, so existing studies have skewed towards research inquiries that could make use of the data available, rather than towards the most pressing questions of public importance. Social media platforms have monopolies over information critical to the cultural, political, and social life of our democracy, and they have little incentive to help researchers paint an accurate picture of their impacts, especially when doing so may reveal them in an unflattering light.

As a result, independent researchers are limited in their efforts to study the causes, character, and scope of the various phenomena attributed to the rise of social media. There is widespread alarm over perceived problems such as a rise in hate speech across platforms, algorithmic systems that push users into ideological echo chambers or extremist rabbit holes, and the spread of inaccurate information from low-credibility news sources. Some government actors in the United States have attempted to ban the videohosting platform TikTok based on alleged national security concerns, while others seek to regulate a host of platforms out of concerns for adolescent mental health. And the rise of Generative Artificial Intelligence (AI) raises new fears about the spread of dis- and misinformation on social media.

Without researcher access to accurate, comprehensive, and timely platform data, the public and policymakers are forced to rely on guesswork

when grappling with these important cultural, social, and political issues. On the one hand, members of the public are unable to make informed decisions about social media use both as consumers in the marketplace and more fundamentally as citizens in a democratic society. On the other, policymakers are unable to develop effective social media regulation: Without an evidence-based understanding of the nature of the risks posed by platforms, they are hampered in their ability to design policies to mitigate them or evaluate those policies once implemented.

This untenable status quo points to the need for, and overriding public interest in, meaningful platform transparency mandates. Although the Court has decided not to address directly the general disclosure provisions of the Florida and Texas laws at issue in these cases, the Court's resolution of the remaining provisions—in particular the laws' individualized explanation requirements—implicates fundamental questions about the power of governments to mandate platform transparency and access to data. Amici file this brief to emphasize both the importance of research and the imperative independent mandated access to the data and information that researchers. the independent public. policymakers need to understand the ways in which social media platforms influence public discourse and democracy. Amici respectfully submit that the Court should craft rulings in these cases that leave ample room for responsible legislative and regulatory efforts aimed atmandating meaningful platform transparency and access to data. It is essential that such efforts survive constitutional scrutiny.

#### **ARGUMENT**

I. Independent social science research plays a critical role in helping the public and policymakers understand the effects of social media on democracy, politics, and society.

Social media platforms create and curate information ecosystems that have significant effects on society. Mainstream platforms are "integral to the fabric of our modern society and culture." Packingham v. North Carolina, 582 U.S. 98, 109 (2017). Today, "[t]hey are not merely places people visit online but rather central nodes of our social, economic, and political lives." Nathaniel Persily & Joshua A. Tucker, Conclusion: The Challenges and Opportunities for Social Media Research, in Social Media and Democracy 313, 322 (Nathaniel Persily & Joshua A. Tucker eds., 2020); see also David Lazer et al., Meaningful Measures of Human Society in the Twenty-First Century, 595 Nature 189, 190 (2021) ("[Platforms] have the potential to alter important patterns of human society, such as the speed of information flows, the scope of media production, and the actors responsible for defining public opinion.").

Independent social science research has played a crucial role in helping the public and policymakers understand the wide-ranging effects of these platforms. Among many other issues, research has provided insights into the recommendations of algorithmic systems, the patterns of foreign influence campaigns, the relationship between social media and political behavior and beliefs, the prevalence of hate

speech and harassment, and the efficacy interventions. See, e.g., Gregory Eady et al., Exposure to the Russian Internet Research Agency Foreign Influence Campaign on Twitter in the 2016 US Election and Its Relationship to Attitudes and Voting Behavior, Nature Commc'ns, Jan. 9, 2023, at 1, 2 (finding no evidence of a meaningful relationship between exposure to the alleged Russian foreign influence campaign on Twitter during the 2016 U.S. election and changes in political polarization, or voting behavior); Guess et al., supra, at 402 (finding that personalized social media feed algorithms did not significantly increase political polarization during the 2020 U.S. election). Ultimately, "real gains—economic, political, and social . . . result from the public sharing of insights from analyzing social media data." Persily & Tucker, supra, at 321.

But this research is only possible with access to data that is commensurate with the complexity and scale of social media platforms. See id. at 313–14. For example, experts from Amicus CSMaP analyzed over 750 million tweets related to the 2016 U.S. election, in addition to almost 400 million tweets from a random sample of American Twitter users, to provide systematic evidence that hate speech did not increase on Twitter over the course of the 2016 presidential election campaign and its immediate aftermath. Siegel et al., supra, at 74. In another study, experts analyzed roughly 1.2 billion tweets from over 640,000 Twitter accounts and concluded that most users do not inhabit strict ideological "echo chambers," with conservatives more likely to follow accounts classified

as left-leaning than the reverse. Eady et al., *Political Bubbles*, *supra*, at 18–19.

Independent researchers working in the field of computational social science are able to analyze these massive datasets with cutting-edge tools that allow for an understanding of the actual nature of the activity on social media platforms and its impacts. See Joshua A. Tucker, Computational Social Science for Policy and Quality of Democracy, in Handbook of Computational Social Science for Policy 381, 382 (Eleonora Bertoni et al. eds., 2023); David Lazer et al., Computational Social Science, 323 Science 721, 722 (2009). Indeed, "[t]he more important questions, such as the relative prevalence of a phenomenon, trends over time, or assessments of causal relationships, ... [require] complex research designs, sustained research efforts. (often) sophisticated and methodological tools." Persily & Tucker, supra, at 324; Malcolm R. Parks, BigCommunication Research, 64 J. Commc'n 355, 356 (2014) (explaining that computational methods and tools "often provide the only means of managing and analyzing digital datasets of increasing size and complexity").

For the public and their representatives to make informed decisions about whether to use platforms or how to regulate them, they need insights gleaned from independent computational social science research based on access to and analysis of platform data at scale. Without such research, they will lack "answers to the many crucial questions concerning the relationship between social media and democracy." Persily & Tucker, *supra*, at 313.

II. Social media platforms impose significant limitations on access to the data needed to conduct rigorous research, improve public knowledge, and develop effective public policy.

Social media platforms unilaterally control and limit access to their data, erecting significant barriers to rigorous research, public understanding, and effective policymaking. Voluntary public disclosures like "transparency reports" contain insufficient information. Direct researcher access is inconsistently granted, and when it is, that access can be restrictive, incomplete, and subject to withdrawal at any time, for any reason, and with little recourse.

# A. Platforms unilaterally control access to their data.

The information needed for a full understanding of how social media platforms operate is owned and controlled by private companies, and the employees of companies are the only ones comprehensive and reliable access to it. Companies such as Google and Meta employ large research teams devoted to investigating a host of topics on their platforms, ranging from social to technical. See Advancing the State of the Art, Google Research, https://research.google; Research. Meta. https://research.facebook.com; see also, e.g., Eytan Bakshy, Solomon Messing & Lada A. Adamic, Exposure to Ideologically Diverse News and Opinion on Facebook, 38 Science 1130, 1130 (2015) (research by Facebook employees with access to a dataset of millions of active U.S. users); Ferenc Huszár et al.,

Algorithmic Amplification of Politics on Twitter, Proc. Nat'l Acad. Scis. U.S., Dec. 21, 2021, at 1, 5 (research by Twitter employees with access to an experiment involving millions of active users).

To date, however, the public has had only scant and indirect knowledge about the kinds of research inquiries platforms that pursue internally, or the results of those inquiries. Notably, when former employee Frances Haugen leaked internal documents from Facebook in 2021, the public got a peek into the kinds of research projects platforms undertake, and the kinds of internal policy deliberations that result. The leaks confirmed the scope of inquiries pursued by Facebook's research teams, from whether content moderation policies are applied consistently, to the impact of Instagram on the well-being of adolescents. See generally The Facebook Files, Wall St. J., https://www.wsj.com/articles/the-facebook-files-11631713039.

Importantly, the Haugen leaks demonstrated that a significant consequence of the platforms' opacity is that the public cannot trust what the companies themselves reveal about their operations. Many of the findings in the leaked internal reports directly contradicted public statements made by Meta. For example, the company has stated that content apply equally moderation policies to everyone. FacebookCommunity See Standards. Meta, https://transparency.fb.com/policies/communit y-standards ("Our Community Standards apply to everyone, all around the world, and to all types of content."); Monika Bickert, Working to Keep Facebook Safe, Meta (July 17, 2018), https://about.fb.com/news/

2018/07/working-to-keep-facebook-safe ("We want to make clear that we remove content from Facebook, no matter who posts it, when it violates our standards."). In fact, the company has previously exempted certain high-profile users from many, if not all, enforcement actions for violations of content policies. A leaked 2019 internal review concluded that this "favoritism" was "both widespread and not publicly defensible": "We are not actually doing what we say we do publicly." Jeff Horwitz, Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt., Wall St. J. (Sept. 13, 2021), https://www.wsj .com/articles/facebook-files-xcheck-zuckerberg-eliterules-11631541353 (quotation marks omitted). Similarly, despite internal research suggesting that Instagram use may lead to negative mental health impacts on a sizable percentage of adolescent girls, Meta has minimized this risk repeatedly in its public statements. See Georgia Wells et al., Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show, Wall St. J. (Sept. 14, 2021), https://www.wsj.com/articles/facebook-knowsinstagram-is-toxic-for-teen-girls-companydocuments-show-11631620739. Were it not for the leaks of internal information, the public would have been unaware not only of the content of these findings, but also of the fact that Meta knew about them.

Nevertheless, the Haugen leaks raised more empirical questions than they answered. Unsurprisingly, the leaked internal information was itself incomplete, lacking much of the underlying data and research methods that would be necessary to independently evaluate the rigor of Meta's own

internal findings. Thus, the leaks not only served to highlight the information asymmetry between platforms and the public, but also reinforced the need for robust access for independent researchers to help answer key questions about the effects of social media.

The platforms' informational monopolies, along with the mismatch between their public statements and internal analyses, produce distrust among the public and policymakers about these services. As long as social media companies control access to information so tightly, their communications about the impact of their platforms are likely to be greeted with skepticism.

# B. Platforms voluntarily disclose insufficient information to the public.

Platforms have responded to critics of this information asymmetry by voluntarily providing public disclosures like "transparency reports." See, e.g., Transparency Reports, Meta, https://transparency.fb.com; Reports, TikTok, https://www.tiktok.com/transparency/en/reports; Transparency Reports, X, https://transparency.twitter.com/en/reports.html. The platforms tout these reports as evidence that they already publish adequate information. However, transparency reports fail to provide all of the information necessary for the public and policymakers to make informed decisions about the platforms.

There are numerous deficiencies in the platforms' transparency reporting practices. First, the platforms can frame information in ways that give misleading impressions. To take just one example, Facebook's transparency report for the fourth quarter of 2020

promoted the success of its automated moderation tools in removing large quantities of hate speech from the platform. See Mike Schroepfer, Update on Our **Progress** on AIand Hate Speech Detection. Meta (Feb. 11, 2021), https://about.fb.com/news/2021/ 02/update-on-our-progress-on-ai-and-hate-speechdetection. Internal numbers revealed by the Haugen however, suggested that Facebook only leaks. removed between 3 and 5 percent of what the company considered hate speech on the platform. See Spandana Singh & Leila Doty, The Transparency Report Tracking Tool: How Internet Platforms Are Reporting on the Enforcement of Their Content Rules, New America: Open Tech. Inst. (Dec. 9, 2021), https://www.newamerica.org/oti/reports/transparency -report-tracking-tool.

Defining and measuring hate speech on platforms are complex, contested inquiries. "[T]here is no clear consensus on the definition of hate speech" and "there is no consensus with regard to the most effective way to detect it across diverse platforms." Alexandra A. Siegel, Online Hate Speech, in Social Media and Democracy, supra, at 56, 56–61. Independent and rigorous research, backed by data at scale, is required to develop an accurate picture of hate speech on a platform and the effectiveness of the platform's actions to combat it. The only way to get a comprehensive picture of hate speech on the platform is to open up data to independent researchers who are able to apply different approaches to the question.

Second, because platforms themselves choose what information to release and how to categorize it, there are often large gaps in the information provided.

Key content moderation metrics may be missing See Aleksandra Urman & entirely. Mvkola Makhortykh, How Transparent Are Transparency Reports? Comparative Analysis of Transparency Reports Across Online Platforms, Telecomms. Pol'y, Jan. 6, 2023, at 1, 13 (revealing "gaps . . . particularly pronounced in the case of reporting on companies' internal moderation practices" and observing that "in many cases the so-called 'Big Tech' companies' reports tend to be more opaque and rigid than those of the smaller companies"). Facebook includes metrics for the prevalence of some categories of content, but not others. See Singh & Doty, supra. And YouTube has grouped its moderation of spam and misleading information together as one number. *Id.* 

Platforms fail to explain in their transparency reports why information is presented as it is, why certain actions were taken, the relative roles of automated tools and human oversight, and a host of other questions necessary to understand the true meaning of the reported numbers. See, e.g., Svea Windwehr & Jillian C. York, Thank You For Your Transparency Report, Here's Everything That's Missing, Elec. Frontier Found. (Oct. 13, 2020), https://www.eff.org/deeplinks/2020/10/thank-you-your-transparency-report-heres-everything-thats-missing. Consequently, the public and policymakers have little insight into the patterns of platforms' content moderation practices or potential biases in the application and enforcement of platform guidelines.

C. Independent researcher access to data can be restrictive, incomplete, and subject to withdrawal at any time, for any reason, and with little recourse.

The status quo for researcher access to data is untenable. Platforms impose severe restrictions on access that are irreconcilable with central tenets of academic research. Datasets are often incomplete and incompatible with computational social science research methods, and platforms can arbitrarily revoke access, chilling long-term investigations.

1. Platforms routinely place restrictions on independent research that are irreconcilable with academic research integrity and norms. Researchers must apply for access, providing details ranging from their qualifications to the research questions and types of analysis they plan to undertake. See, e.g., Research Platform Addendum, Meta (Nov. 4, 2021), https://developers.facebook.com/terms/facebook\_rese arch\_platform\_terms\_addendum; Program Terms & Conditions, YouTube (July 11, 2022), https://research .youtube/policies/terms; TikTokResearch Services Terms of Service, TikTok (Aug. 10, 2023), https://www.tiktok.com/legal/page/global/terms-ofservice-research-api/en.

These requirements effectively grant platforms the power to screen researchers and research projects. In February 2023, for example, TikTok launched a new research interface, offering social scientists a look at TikTok data that had previously been inaccessible. But TikTok has placed strict stipulations on that access, some of which are incompatible with basic

tenets of the research process. Significantly, TikTok's researcher Terms of Service bar researchers from sharing data. TikTok Research API Services Terms of Service, supra, § IV. That prohibition "flies directly in the face of academic research norms, require...open data for replication, and the independent sharing of results in peer-reviewed journals." Megan A. Brown, The Problem with TikTok's New Researcher API Is Not TikTok. Tech Policy Press (Mar. 1, 2023), https://techpolicy.pr ess/the-problem-with-tiktoks-new-researcher-api-isnot-tiktok. One research scientist who examined TikTok's requirements concluded that "there appears to be no plausible pathway to published research without violating the [Terms of Service]." Joe Bak-Coleman, TikTok's API Guidelines Are a Minefield for Researchers, Tech Policy Press (Feb. 22, 2023), https://techpolicy.press/tiktoks-api-guidelines-are-aminefield-for-researchers.

2. The data available to independent researchers is also inadequate in at least two crucial ways. First, only select platforms have historically allowed access to data at all, so research projects have necessarily focused on those platforms, while others have escaped similar scrutiny. Second, when researchers do gain access to platform data, it is frequently insufficient, abridged, or fragmentary, and the amount of data that platforms release and the speed at which researchers can access it are incompatible with most computational social science research.

Because only some platforms offer access to researchers, many independent studies are designed to make use of the data available, which is not necessarily the data that is most useful for answering the questions that are most important to the public or policymakers. Until recently, a disproportionate number of research studies analyzed Twitter data. See Persily & Tucker, supra, at 314. Yet "this [was] not because there [was] a consensus that Twitter [was] the most politically consequential social media platform. Although Twitter is certainly important for politics . . . , this imbalance in research occur[red] because Twitter data ha[d] historically been among the most easily accessible for outside research, especially compared to Facebook data." Id.

When Facebook has offered access to independent researchers, the data that it provides can be incomplete and selective. In 2018, Facebook launched a data-sharing initiative with the research consortium Social Science One to make Facebook and other industry data available to the larger scientific community. It took almost twenty months for Facebook to produce a dataset similar to the one promised in its original proposal. *Id.* at 315. Further, the company limited that dataset, which delineated the numbers and types of people who saw and engaged with specific URLs posted to the website, to URLs that were publicly shared 100 times or more. Id. This restriction hampered researchers from developing a complete picture of precisely the types of activity that the dataset was intended to disclose. The discretion that platforms maintain in voluntarily releasing data has potentially significant effects on the substantive findings relevant to the public and policymakers. In the case of Social Science One, for instance, the data overestimated the prevalence of both true and false

news by as much as 400 percent. Jennifer Allen et al., Research Note: Examining Potential Bias in Large-Scale Censored Data, Harvard Kennedy Sch.: Misinformation Rev. (July 26, 2021), https://misinforeview.hks.harvard.edu/article/researc h-note-examining-potential-bias-in-large-scale-censored-data.

In another notable example, for over four years Facebook made available engagement metrics for posts from public groups and pages through its data analytics tool known as CrowdTangle. See Kevin Roose, Inside Facebook's Data Wars, N.Y Times (July 14, 2021), https://www.nytimes.com/2021/07/14/ technology/facebook-data.html. In response to a journalist's use of CrowdTangle to rank the pages that garnered the most engagement, however, company publicly disparaged the usefulness of that data, and the inferences that could be drawn from it, as compared to more representative data that was only available internally. *Id.* Put simply, the company criticized the adequacy of the limited data that it made publicly available by citing richer data that it did not.

Platforms may also place other barriers in the way of independent researchers. In 2018, Facebook announced the launch of a searchable advertisement database, available to researchers. Rob Leathern, Introducing the Ad Archive API, Meta (Aug. 22, 2018), https://about.fb.com/news/2018/08/introducing-the-ad-archive-api. Prior to the launch, CEO Mark Zuckerberg had told investors that "even without legislation, we're already moving forward on our own to bring advertising on Facebook to an even higher

standard of transparency than ads on TV or media." Josh Constine, ZuckSaysTransparency Regulation Would Be Very Good If Done Well,' TechCrunch (Nov. 2017), https://techcrunch.com/2017/11/01/zuck-may-supportad-regulation. Researchers soon discovered, however, that the company had not provided data about the level of engagement with ads or about the criteria used for its ad targeting algorithm. Moreover, because researchers could only explore the database via keyword search, researchers could not verify the size and scope of the dataset. See Monika Zalnieriute, "Transparency Washing" in the Digital Age, 8 Critical Analysis L. 139, 148 (2021). As long as social media companies are free to decide what data to release and in what form, social scientists will be frustrated in their efforts to inform the public and policymakers.

3. Even where platforms voluntarily make some data available, they can revoke that permission at any time, for any reason, and with little recourse. This creates a level of uncertainty that discourages researchers from pursuing long-term studies and makes the investment of resources into those studies risky.

Of the major social media platforms, only Twitter previously offered routine access to data. Until its recent acquisition, Twitter had been known for its relative transparency, and for providing developers and researchers free access to its Application Programming Interface (API), which allows "to programmatically researchers retrieve and analyze" data. See About the Twitter API, X, https://developer.twitter.com/en/docs/twitterapi/getting-started/about-twitter-api. This access proved invaluable for researchers to "systematically and reliably collect public tweets posted by public figures, gather information about network dynamics, investigate bots and other inauthentic activity, or analyze conversations around specific Exec. Bd., Twitter's New API Plans Will Devastate Public Interest Research, Coal. Indep. Tech. Rsch. (Apr. 3, 2023), https://independenttechresearch.org/le tter-twitters-new-api-plans-will-devastate-publicinterest-research ("Rsch. Coal. Letter"). In the past three years alone, researchers produced "more than 17,500 academic papers based on the platform's data." Chris Stokel-Walker, Twitter's \$42,000-per-Month Nearly APIPricesOut Everyone, Wired (Mar. 10, 2023), https://www.wired.com/story/twitterdata-api-prices-out-nearly-everyone.

Under its new leadership, however, X has put its platform data behind a paywall that effectively renders large-scale social science research all but impossible. Access to the millions of tweets that researchers previously enjoyed for free now costs tens if not hundreds of thousands of dollars per month prohibitively expensive for academic projects. Id. Many academics who had relied on this access for years have suddenly found their research derailed by private decision-making over which they have little recourse. See Rsch. Coal. Letter, supra (surveying public-interest researchers and identifying "over 250 projects that would be jeopardized ..., including research into the spread of harmful content, (dis)information flows, crisis informatics,

news consumption, public health, elections, and political behavior").

Twitter's recent policy shift is dramatic, but it is not the only example of a platform changing its mind about granting access to data. In 2018, Facebook shut down its Pages API, which had allowed researchers access to posts, comments, and associated metadata from public Facebook pages. Deen Freelon, Computational Research in the Post-API Age, 35 Pol. Commc'n 665, 665 (2018). This move essentially closed off all independent access to a key source of Facebook content, unless granted via research collaboration with the company. Researchers were given no warning, and virtually overnight, many of their methods and projects were rendered obsolete. Id. Subsequently, in a more targeted move, Facebook disabled the accounts of researchers affiliated with New York University's nonpartisan research group Cybersecurity for Democracy, which was studying how disinformation spreads online through its Ad Observer browser plugin. See Laura Edelson & Damon McCov, We Research Misinformation on Facebook. JustDisabledOur Accounts. N.Y. Times (Aug. 10, 2021), https://www.nytimes.com

<sup>&</sup>lt;sup>5</sup> Experts from Amicus CSMaP have participated in such research collaborations. See, e.g., Research Partnership to Understand Facebook and Instagram's Role in the U.S. 2020 Election, Meta, https://research.facebook.com/2020-election-research; Sandra González-Bailón et al., Asymmetric Ideological Segregation in Exposure to Political News on Facebook, 381 Science 392, 393 (2023) (presenting research results from the "US 2020 Facebook and Instagram Election Study, a collaborative effort between Meta and a team of external researchers").

/2021/08/10/opinion/facebook-misinformation.html. The unpredictability of access makes it difficult for researchers to invest in studies that require long-term access, such as those that seek to track how trends change over time, or even merely those that might require extensive planning and preparation.<sup>6</sup>

These policy changes are notable because they affected research projects already underway, but it has also been common for platforms to deny researchers any access to data at all. TikTok initiated a program ostensibly offering limited data access to researchers in 2023; before then, it offered nothing. Mia Sato, Researchers Will Get Access to TikTok Data – Pending Company Approval, Verge (Feb. 21, 2023), https://www.theverge.com/2023/2/21/23604737/tiktok-research-api-expansion-public-user-data-transparency. YouTube, a platform that has existed for nearly two decades, launched its first research program only last year, and has so far granted access

<sup>&</sup>lt;sup>6</sup> Investigative journalists and civil society groups have also had their access to information revoked with little warning. In 2019, Facebook blocked the nonprofit newsroom ProPublica's ad tracking tools, which allowed users to see how they were being targeted by ads. Jeremy B. Merrill & Ariana Tobin, Facebook Moves to Block Ad Transparency Tools—Including Ours, ProPublica (Jan. 28, 2019), https://www.propublica.org/article/f acebook-blocks-ad-transparency-tools. In 2021, the nonprofit AlgorithmWatch was forced to shut down its ongoing research into Instagram's newsfeed algorithm after Meta accused it of violating itsterms ofservice. Nicolas Kayser-Bril, AlgorithmWatch Forced to Shut Down Instagram Monitoring Project After Threats from Facebook, AlgorithmWatch (Aug. 31, 2021), https://algorithmwatch.org/en/instagramresearch-shut-down-by-facebook.

to a mere handful of researchers. John Albert, *Platforms' Promises to Researchers: First Reports Missing the Baseline*, AlgorithmWatch (Feb. 16, 2023), https://algorithmwatch.org/en/platfor ms-promises-to-researchers. The arbitrary decisions to revoke access by some platforms take place against a backdrop of other platforms that provide virtually no access whatsoever.

III. Mandating meaningful platform transparency is the only way to ensure that independent researchers, the public, and policymakers will be able to answer urgent questions about how platforms operate and influence contemporary society.

Asimpactful as previous and ongoing independent research efforts have been despite the limitations that social media platforms have placed on them, they have only scratched the surface of understanding basic facts about contemporary society. Social science researchers like Amici are limited in their efforts to get a handle on the causes. character, and scope of the various phenomena attributed to the rise of social media. Amici have been stymied by platforms' unwillingness to provide access to their data, and Amici's research has often raised further questions that necessitate access to data across multiple platforms that they do not have. There are also research questions of critical importance to the public and policymakers that social scientists have simply not been able to address at all because researchers have not had access to the necessary data. See Jonathan Nagler & Joshua A. Tucker, The Social Media Data We Need to Answer Key Research

Questions, CSMaP (May 4, 2022), https://csmapnyu.org/news-views/news/the-social-media-data-we-need-to-answer-key-research-questions; see also, e.g., Irene V. Pasquetto et al., Tackling Misinformation: What Researchers Could Do with Social Media Data, Harvard Kennedy Sch.: Misinformation Rev. (Dec. 9, 2020), https://misinforeview.hks.harvard.edu/article/tackling-misinformation-what-researchers-could-do-with-social-media-data (describing in detail misinformation "research that could hypothetically be conducted if social media data were more readily available").

In turn, the public and policymakers are left in the dark at a time when "[t]he need for real-time production of rigorous, policy-relevant scientific research on the effects of new technology on political communication has never been more urgent." Persily & Tucker, supra, at 313. "In the absence of comprehensive data, all of us—citizens, journalists, pundits, and policy makers—are crafting narratives about the impact of social media that may be based on incomplete, sometimes erroneous information . . . . " Teresa Carr, Why Researchers Want Broader Access toSocialMedia Data. Nieman Lab (May 4, 2022), https://www.niemanlab.org/2022/05/w hy-researchers-want-broader-access-to-social-mediadata. Legislators and regulators advancing public policy to tackle real problems are at a loss in trying to understand the nature of the risks posed by platforms, designing policies to mitigate them, and evaluating those policies once implemented. Proposed policies could unduly limit free speech online and harm innovation with little ability for researchers to

measure their intended effects. See Persily & Tucker, supra, at 321–22 (explaining the "dangers" of making public policy "without the advantage of the insights that can be gained through analysis of social media data").

Users and policymakers alike have long expressed alarm over perceived phenomena such as a rise in hate speech across platforms, algorithmic systems that push users into ideological echo chambers or extremist rabbit holes, and the spread of inaccurate information from low-credibility news sources. Although *Amici*'s research into these areas has revealed what can be achieved when researchers have access to large-scale data, important questions remain unexamined and unanswered. *See supra* Section I.

Emerging areas of concern across the social sciences also face challenges to understanding due to a lack of available data from social media companies. For example, governments in the United States have attempted to ban the video-hosting platform TikTok based on alleged national security concerns. See Bobby Allyn, Montana Becomes 1st State to Approve a NPR (Apr.  $Full \quad Ban \quad of \quad TikTok,$ 14, https://www.npr.org/2023/04/14/1170204627/montan a-becomes-1st-state-to-approve-a-full-ban-of-tiktok. But the public and policymakers lack information that would help guide policy decisions and public debate, as well as access to the requisite data for measuring the effects of any state or federal policies that may be implemented. See Max Zahn, No Evidence of TikTok National Security Threat but Reason for Concern, Say, ABC News (Mar. 28, Experts https://abcnews.go.com/Technology/evidence-tiktoknational-security-threat-reason-concernexperts/story?id=98149650. In this highly sensitive area of foreign policy, the public and policymakers should not have to rely on guesswork or supposition.

Governments are also attempting to regulate a host of platforms out of concerns for adolescent mental health. Numerous studies and surveys show that rates of depression and suicide among teenagers have sharply increased in recent years. See, e.g., A.W. Geiger & Leslie Davis, A Growing Number of American Teenagers – Particularly Girls – Are Facing Rsch. Pew Depression, Ctr. (Jul. 12, 2019), https://www.pewresearch.org/shortreads/2019/07/12/a-growing-number-of-americanteenagers-particularly-girls-are-facing-depression; Ctrs. for Disease Control & Prevention, Youth Risk BehaviorSurvey 2 - 3(2023),https://www.cdc.gov/healthyyouth/data/yrbs/pdf/YRB S\_Data-Summary-Trends\_Report2023\_508.pdf (tracking increased rates of poor mental health and suicidal thoughts and behaviors among young people from 2011 to 2021). Government officials at all levels have pointed to social media as a possible cause. See, e.g., Matt Richtel et al., Surgeon General Warns that Social Media May Harm Children and Adolescents, N.Y. Times (May 23, 2023), https://www.nytimes.com /2023/05/23/health/surgeon-general-social-mediamental-health.html; Bobby Allyn, States Sue Meta, Claiming Instagram, Facebook Fueled Youth Mental HealthCrisis. 24, NPR (Oct. 2023), https://www.npr.org/2023/10/24/1208219216/statessue-meta-claiming-instagram-facebook-fueled-youthmental-health-crisis.

But researchers have found it difficult to assess whether social media contributes to rising rates of mood disorders, such as depression and anxiety, or related behaviors, particularly incidents of self-harm or suicide. See Matti Vuorre & Andrew K. Przybylski, Global Well-Being and Mental Health in the Internet Age, Clinical Psych. Sci., Nov. 27, 2023, at 1, 15 (explaining that "[r]esearch on the effects of Internet technologies is stalled because the data most urgently needed are collected and held behind closed doors"). The public and policymakers do not have the data necessary to determine what role social media has or has not played in this apparent mental health crisis. "Until these data can be transparently analyzed for the public good, the potential harmful effects of the Internet and other digital environments will remain unknown." Id. at 15. Young people and their families will be unable to make informed decisions about social media use, and public health officials and legislators will be unable to craft effective regulations.

Finally, the rise of Generative AI raises new fears about the spread of dis- and misinformation on social media. See Zeve Sanderson et al., White House OSTP Comments onAI, **CSMaP** (July 7, 2023), https://csmapnyu.org/news-views/news/white-house-The development ostp-comments-on-ai. foundational models—which enable the generation of text, image, and video at scale—will undoubtedly impact the networked information ecosystem. These developments have the potential to facilitate the widescale production of spam, harassment, and false information, much of which will be both cheaper to produce and more difficult to detect. Access to

platform data is needed to understand the impact of the burgeoning AI landscape on the information ecosystem. Zeve Sanderson et al., *Feedback on EU Article 40*, CSMaP (May 23, 2023), https://csmapnyu.org/news-views/news/csmapfeedback-on-eu-article-40.

When researchers do not have access to accurate, comprehensive, and timely data, the public and policymakers are forced to rely on guesswork when grappling with these critical cultural, social, and political issues. Instead, they need better information "to shape their understanding of and participation in public discourse, elections, and other building blocks of democracy." Daphne Keller, *Platform Transparency and the First Amendment*, 4 J. Free Speech L. 1 (2023).

Lawmakers have begun to recognize mandating greater platform transparency, including consistent and comprehensive access to data, is the only way to ensure that independent researchers will be able to overcome current barriers and enable evidence-based debate about the role and impact of social media platforms. Beyond Florida and Texas, legislative proposals abound. See, e.g., Laura Edelson, Platform Transparency Legislation: The Whos, Whats and Hows. Lawfare (Apr. 29, 2022), https://www.lawfaremedia.org/article/platformtransparency-legislation-whos-whats-and-hows. It is essential that these efforts survive constitutional scrutiny.

#### **CONCLUSION**

In analyzing the specific provisions at issue in the Court should cases, recognize the democratic imperative for greater platform transparency and craft rulings that leave ample room for responsible legislative and regulatory efforts meaningful aimed mandating platform transparency and access to platform data.

Respectfully submitted,

Jacob M. Karr
Counsel of Record
Jason M. Schultz
TECHNOLOGY LAW AND
POLICY CLINIC
NEW YORK UNIVERSITY
SCHOOL OF LAW
245 Sullivan Street
New York, NY 10012
(212) 998-6042
jacob.karr@law.nyu.edu

Counsel for Amici Curiae

December 7, 2023