ONE HUNDRED FIFTEENTH CONGRESS

Congress of the United States

House of Representatives

COMMITTEE ON ENERGY AND COMMERCE

2125 RAYBURN HOUSE OFFICE BUILDING
WASHINGTON, DC 20515–6115

Majority (202) 225–2927
Minority (202) 225–3641

September 28, 2018

Mr. Jack Dorsey
CEO
Twitter, Inc.
1355 Market Street, Suite 900
San Francisco, CA 94103

Dear Mr. Dorsey:

Thank you for appearing before the Committee on Energy and Commerce on Wednesday, September 5, 2018, to testify at the hearing entitled "Twitter: Transparency and Accountability."

Pursuant to the Rules of the Committee on Energy and Commerce, the hearing record remains open for ten business days to permit Members to submit additional questions for the record, which are attached. To facilitate the printing of the hearing record, please respond to these questions by the close of business on Monday, October 15, 2018. Your responses should be mailed to Ali Fulling, Legislative Clerk, Committee on Energy and Commerce, 2125 Rayburn House Office Building, Washington, DC 20515 and e-mailed in Word format to ali.fulling@mail.house.gov.

Thank you again for your time and effort preparing and delivering testimony before the Committee.

Sincerely,

Greg Walden
Chairman

cc: Frank Pallone, Ranking Member

Attachment

Additional Questions for the Record

The Honorable Greg Walden

1. Mr. Dorsey, you testified that Twitter uses thousands of "signals" to determine and decide what to show, downrank, or filter. Please outline the specific behavioral signals that were implicated in the search auto-suggest issue. Additionally, please outline the specific signals that Twitter currently uses and how those signals determine what content to show, downrank, or filter.

2. Mr. Dorsey, you testified that "about 600,000 accounts" were impacted by the auto-suggest search issue. With respect to these accounts, please provide the following:

   a. The number of Members of Congress' accounts that were impacted. Please provide a complete list of all the names of Congressional members, or their accounts, that were affected.

   b. The number of Republican Members of Congress' accounts that were impacted.

   c. The number of Democratic Members of Congress' accounts that were impacted.

   d. The number of other Federal, State or local government accounts that were impacted.

   e. Please provide a complete list of all the names of the remaining non-Congressional accounts.

3. Mr. Dorsey, you testified that Twitter needs to do a better job with your Terms of Service to make them more understandable and approachable. Please explain the steps Twitter is taking or intends to take to rework the Twitter Terms of Service to accomplish this?

4. Mr. Dorsey, how does Twitter determine whether the information included in Twitter's Terms of Service is understandable and approachable?

   a. How does Twitter define understandable with respect to Twitter's Terms of Service?

   b. How does twitter define approachable with respect to Twitter's Terms of Services?

5. Mr. Dorsey, does Twitter ever review, use, or consider data or information about a user unrelated to Twitter to make decisions about content posted by a Twitter user?

   a. If yes, what data or information does Twitter consider and why?

6. Mr. Dorsey, you testified that Twitter's verification program is not where you'd like it to be and that it needs a "serious reboot." When was the verification program shut down? Please explain the current problems and inadequacies with the verification program and what steps Twitter is taking or intends to take to address this.

7. Mr. Dorsey, in Twitter's July 26, 2018 blog post "Setting the record straight on shadow banning," company officials indicated:

   "For the most part, *we believe the issue had more to do with how other people were interacting with these representatives' accounts than the accounts themselves* (see bullet #3 above). There are communities that try to boost each other's presence on the platform through coordinated engagement. *We believe these types of actors engaged with the representatives' accounts-- the impact of this coordinated behavior, in combination with our implementation of search auto-suggestions, caused the representatives' accounts to not show up in auto-suggestions*" (emphasis added).

   For additional context, the bullet no. 3 referenced above pointed out: "How other accounts interact with you (e.g. who mutes you, who follows you, who retweets you, who blocks you, etc)".

   Yet, in response to a question from Rep. Walberg about what "specific signals or actions of other accounts interacting with the representatives' account would you suggest contributed to the auto-suggest issue," you asserted that the "behaviors we were seeing were actual violations of our terms of service."

   Please explain in detail whether the auto-suggest search issue was a consequence of "interaction" issue described in Twitter's blog post or violations of Terms of Service as you described in your testimony, or both? If it was a consequence of Terms of Service violations, how and why would such violations render affected accounts invisible to Twitter users?

8. Mr. Dorsey, in response to a question about the Meghan McCain incident and the inadequacies of Twitter's abuse prioritization mechanism, you indicated "[i]n this particular case, the reason why was because [the violent and physical harm element] *was captured within an image rather than the tweet text itself*" (emphasis added). Is currently Twitter without the technological tools to police harmful and abusive content embedded in either images, .gifs, links, videos, and audio clips? If yes to any, how do human reviewers police harmful and abusive content embedded in either images, .gifs, links, videos, and audio clips?

9. Mr. Dorsey, what policies and procedures does Twitter have in place to notify accounts and users when their messages or content have been removed, suspended, banned, or otherwise rendered invisible? What steps going forward does Twitter intend to take to better notify accounts and users?

10. Mr. Dorsey, the Twitter Rules preclude users from posting graphic violence and adult content on the Twitter platform.

    a. Please provide Twitter's definition of "graphic violence" and outline the specific factors Twitter uses to determine whether content meets this definition.

    b. Please provide Twitter's definition of "adult content" and outline the specific factors Twitter uses to determine whether content meets this definition.

11. Mr. Dorsey, how does Twitter review the platform to identify content that meets Twitter's definition of graphic violence or adult content?

    a. Is content on the platform proactively scanned or reviewed by Twitter before it is flagged by a Twitter user?

        i. If yes, is the screening done by algorithms or humans?

            1. If the screening is done by algorithms, please outline the factors and signals the algorithms use to determine what is considered "graphic violence" and "adult content" as defined by Twitter.

            2. If the screening is done by humans, please provide the number of employees responsible for reviewing content on the Twitter platform to identify "graphic violence" or "adult content."

        ii. If no, why does Twitter solely rely on users to flag potentially violative content?

12. Mr. Dorsey, the Twitter Rules indicate that some content that may be violative of the Twitter Rules will be allowed but marked as "sensitive media."

    a. What factors does Twitter consider when deciding whether content should be prohibited and removed from the Twitter platform or should be allowed and remain on the platform, but marked as "sensitive media?"

    b. Is this decision made by algorithms or humans?

        i. If it is made by algorithms, what factors or signals do the algorithms consider when making this decision?

        ii. If it is made by humans, please provide the number of employees responsible for making this decision and identify what factors those employees consider.

13. Mr. Dorsey, the Twitter Rules state that "at times, [Twitter] may prevent certain content from trending" and that it may be content that either "violates Twitter Rules" or "may attempt to manipulate trends."

    a. When content does not violate the Twitter Rules or attempt to manipulate trends, does Twitter ever prevent such content from "trending?"

        i. If so, please provide an example of content that Twitter has not allowed to "trend" despite it being in compliance with the Twitter Rules.

    b. Please explain what factors Twitter considers when determining whether content "attempts to manipulate trends."

14. Mr. Dorsey, the Twitter Rules provide that users may not engage in hateful conduct. Please provide Twitter's definition of "hateful conduct" and outline the specific factors Twitter uses to determine whether content meets this definition.

    a. When content is flagged as "hateful conduct", what factors does Twitter consider when deciding whether to remove the content?

        i. Does Twitter consider data points or information about a user unrelated to the specific content to make this determination?

15. Mr. Dorsey, when a Twitter user flags content as potentially violative of the Twitter Rules, what specific steps does Twitter undertake to determine whether that content should be removed from the platform?

16. Mr. Dorsey, what specific steps does Twitter undertake to determine whether a Twitter user's account should be suspended?

17. Mr. Dorsey, what specific steps does Twitter undertake to determine whether a Twitter user's account should be banned?

18. Mr. Dorsey, does Twitter only use content on the platform to make decisions about suspending or removing Twitter accounts or does Twitter use data from outside entities to make those decisions?

    a. If yes, please identify the outside entities Twitter relies on for information about Twitter users to make these decisions.

19. Mr. Dorsey, in late August, you were interviewed by CNN and you indicated "[Twitter has] changed a lot. But [Twitter hasn't] changed the underlying fundamentals." Please identify the underlying fundamentals of Twitter.

20. Mr. Dorsey, in late August, you were interviewed by CNN and you indicated "We've seen abuse, we've seen trolling, we've seen harassment, we've seen misinformation." Did

you not anticipate this potential, either when you launched the platform or at any other point since launching?

21. Mr. Dorsey, as the company grew, were there ever discussions about implementing greater security and enhanced transparency measures?

    a. If yes, who was involved in those conversations and, if no, why not?

22. Mr. Dorsey, Twitter agreed to a "code of conduct" on Hate Speech under threat from the European Union. Has it changed any of its policies as it affects the U.S. based on this agreement, and has Twitter ever removed any U.S. based accounts on behalf of any foreign government?

23. Mr. Dorsey, with the exclusion of the "Network of Enlightened Women," which Twitter appears to have added recently and have no record opposing online censorship, none of the 48 groups and individuals on the Twitter Trust and Safety Council lean even mildly right-of-center. Does Twitter have plans to include more pro-free speech or conservative voices in their product and policy decisions?

24. Mr. Dorsey, what percentage of Twitter content moderation reviews are conducted by actual human beings rather than via automated processes and artificial intelligence?

25. Mr. Dorsey, how many employees have been hired and are currently assigned to content moderation and reviews? Currently, what training and instruction do reviewers receive to ensure Twitter users were in compliance with the company's Terms of Service? What salary range does Twitter pay employees assigned to content moderation and reviews?

26. Mr. Dorsey, does Twitter have a list of specific users it surveils more closely than other Twitter users because of a history of use of fake accounts?

27. Mr. Dorsey, how do you define a "fake account" and what steps does Twitter undertake to identify fake accounts?

    a. Once a fake account is identified, what specific steps does Twitter undertake to verify that an account should be removed?

28. Mr. Dorsey, in January of this year, the New York Times published an article entitled "The Follower Factory" that revealed some Twitter users purchase followers and retweets from a company who specializes in selling automated bot accounts. Is it a violation of the Twitter User Agreement to purchase followers and retweets?

    a. Since the publication of the New York Times' article "The Follower Factory" what steps has Twitter taken to address the issue of users purchasing fake accounts and retweets?

b. What are the harms and concerns about artificially inflating the number of followers on social media by paying for it?

29. Mr. Dorsey, the Twitter Rules provide that "if voices are silenced because people are afraid to speak up" then the underlying philosophy of free speech and expression means little.

    a. Do algorithms make the determination that certain content silences voices and, if so, what factors are considered by the algorithm?

    b. Do humans make the determination that certain content silences voices and, if so, what factors are considered?

30. Mr. Dorsey, the Twitter Rules specifically state that users may not use the platform for "any unlawful purpose or in furtherance of illegal activities." If Twitter determines content is unlawful or in furtherance of illegal activities, does Twitter notify law enforcement?

31. Mr. Dorsey, has Twitter, or any third parties it engages, audited its input data and ranking system to determine exactly how much bias each contributes to output bias in resulting Twitter search results? With what frequency does Twitter or others does conduct such audits or studies?

32. Mr. Dorsey, what suggestions do you have for raising the awareness of Twitter users by signaling bias in Twitter search and timeline results?

33. Mr. Dorsey, Twitter announced August 30th a new issue ads policy and certification process in the U.S. When will the company begin to enforce its new issue ads policy? Please describe in detail what was the previous issue ads policy and certification process, if there was one?

34. Mr. Dorsey, how many advertisers – either individuals, organizations, or campaigns – have applied and received certification under Twitter's new process?

35. Mr. Dorsey, please explain in detail specifically how will campaign and issue ads be labeled or highlighted in a user's timeline? How many campaign and issue ads are currently labeled under the new policy and process? From their timeline can a user now click on the promoted ads to immediately see information about the advertiser's identity and location, like a "Learn More" button?

36. Mr. Dorsey, under Twitter's August 30th issues ad policy, will both campaign and issue ads be viewable in Twitter's Ads Transparency Center? Please explain what specific detail will be viewable to a user?

37. Until 1987, the Fairness Doctrine required broadcasters to provide a right of reply to ensure the presentation of balanced views on issues of public importance, during a time when broadcast was a dominant news source for most Americans. While the FCC repealed the Fairness Doctrine, the principle of providing balanced perspectives is still important in journalism. When amplifying certain topics on the platform, does Twitter seek to ensure that it promotes a balanced variety of viewpoints on a particular issue? If so, what steps are taken to ensure representation of a variety of viewpoints, including those that are less popular with users?

38. In 1996, 10 years before the original Twitter platform was released, Section 230 of the Communications Decency Act was enacted, distinguishing interactive online platforms from traditional publishers by setting up a safe harbor protecting them from the lawsuits publishers may face over third party generated content. This was based on the premise that interactive services are essentially neutral platforms, not exercising the full editorial judgment wielded by a publisher. Although like the internet message boards that were prevalent when Section 230 was enacted, Twitter gets much of its content from the users, given the curating power of the Twitter algorithms, coupled with other efforts on the company's part to proactively promote or deprioritize particular content or users, are you exercising editorial judgment?

39. In *Six4three, LLC v. Facebook, Inc.*, a case concerning alleged anticompetitive effects of its content management practices, Facebook recently argued that it is a publisher and that its editorial decisions are therefore protected by the First Amendment.

    a. Would you characterize Twitter in the same way?

    b. Should content moderation efforts that make value judgments about quality, veracity, or tone, similar to decisions made by publishers, affect the applicability of Section 230, which specifies that Internet platforms cannot be treated as publishers?

40. Earlier this year, the Allow States and Victims to Fight Online Sex Trafficking Act of 2017 (FOSTA) was signed into law. This legislation represented the first successful attempt to amend Section 230 since it was enacted in 1996.

    a. Has Twitter changed any of its procedures with regard to detecting and removing sex trafficking content as a result of FOSTA?

    b. Has the legislation impacted Twitter's bottom line?

    c. Does the Twitter user experience differ in other markets where internet platforms are not protected with safe harbor legislation similar to Section 230?

41. In your testimony, you indicated that Wall Street did not approve of Twitter's recent actions to de-activate suspicious accounts, ultimately lowering the number of followers of certain accounts. While Twitter can afford to do this in the short term, Twitter, as a

publicly held company, has a fiduciary duty to its shareholders. How do you expect the shifted focus on the health of the conversation taking place on the platform, rather than maximizing user engagement, to impact the company?

42. For some events or topics, Twitter features custom emojis that automatically appear when a user types a certain hashtag, which is a form of promoting or prioritizing certain content. How is it decided to allow a custom emoji for a hashtag, and who designs the emojis? Can users apply for this service? Are the designs done in consultation with the event coordinators or users actively promoting these hashtags?

43. The "trending topics" list is one important method Twitter uses to promote or prioritize certain content. Is a topic chosen to be featured in this list strictly based on how many posts are made on a specific hashtag within a period of time, or do any human decisions intervene? Anecdotally, some users have observed that hashtags with fewer tweets sometimes appear on trending lists while those with more tweets do not.

   a. A "tailored trends" list is generated based on a user's location and who they follow on Twitter. How does Twitter use location in determining which trends or content to display to an individual user? What is the source of the location data being used?

44. Rep. Walberg stated in his question to you that ISPs have certain transparency requirements that require the ISPs to disclose if they are altering traffic, and asked you if a similar idea would be helpful in the tech industry. You replied, "That [transparency] is a good idea and would help earn people's trust." What should those transparency requirements look like with regard to content moderation practices?

45. Mr. Dorsey, understanding that in your testimony you stated that algorithmic bias is a relatively "new problem" and that Twitter is "early" in its learning process in terms of managing it, what research, development, and other tool deployment is Twitter currently undertaking to understand, manage, and address algorithmic bias on its platform?

46. Mr. Dorsey, in your testimony, you stated that Twitter reviewers look for indicators of "fairness" and "impartiality" when reviewing algorithmic output to try and identify bias.

   a. Please provide Twitter's definition of "fairness" as it applies to Twitter's oversight of its algorithms.

   b. Please provide Twitter's definition of "impartiality" as it applies to Twitter's oversight of its algorithms.

The Honorable Michael C. Burgess

1. Mr. Dorsey, you stated during the hearing that Twitter is building technologies so that it doesn't have to wait on reports from victims to act on violent or threatening content.

   a. When do you anticipate implementation of these new technologies?

   b. Currently, how many reports of inappropriate or threatening content does Twitter receive per day?

      i. How many of these reports result in action?

2. Mistakes do happen, and we understand that, but we would like to understand what the process is at Twitter if a mistake happens. That process is still very unclear.

   a. If a user does not believe they have violated the Twitter Rules, what recourse does that user have? What is the internal process at Twitter to review an appeal? Are employees deciding whether they agree with a users' posts?

   b. There are high profile Twitter users who seem to have a different level of recourse from, for example, a constituent of mine who uses Twitter for news and conversations with friends. Are there different channels for review in these cases?

3. Mr. Dorsey, I understand that Twitter is conducting internal investigations that will result in a transparency report.

   a. When do you anticipate being able to share this information with the Energy and Commerce Committee?

4. Mr. Dorsey, I've long been concerned about the role the Internet and online platforms, such as Google, Facebook, Pinterest and Twitter, have played in enabling access to deadly and illegal controlled substances. Given the current opioid crisis, would you please describe in detail what specific policies and procedures Twitter has in place to crack down on illegal online sales and marketing of opioids and pharmaceutical drugs.

   a. Does Twitter have policies and procedures in place to disable the ability to use the Twitter search function for the sales and marketing of controlled substances? If yes, please outline in detail those policies and procedures. If no, does Twitter plan on implementing such measures and when?

   b. Does Twitter have policies and procedures in place to report to Federal, State or international law enforcement information Twitter receives indicating that an individual or organization is engaged in the sale or marketing of controlled substances? If yes, please outline in detail those policies and procedures. If no, does Twitter plan on implementing such measures and when?

c. Has Twitter established a 24/7 point of contact with whom Federal, State or international law enforcement can communicate directly if law enforcement has information indicating that an individual or organization is engaged in the sale or marketing of controlled substances? If you have, please outline in detail what person or department is Twitter is responsible, and provide their contact information. If you have not, are you planning on implementing such measures and when?

d. Does Twitter have specific information in its "Help Center" for users to report the sale or marketing of controlled substances on the platform, similar to the help modules on "online abuse" and "self-harm and suicide"? If yes, please outline in detail what they are. If no, does Twitter plan on implementing such measures and when?

The Honorable Robert E. Latta

1. Mr. Dorsey, in your written statement you indicated Twitter conducted an internal analysis of the Members of Congress affected by the auto-suggest search issue, and that you would make that information available to the Committee, if requested. Please provide the internal analysis done by Twitter with respect to the auto-suggest search issue.

2. Mr. Dorsey, in your testimony you indicated that Twitter has made more than 30 policy and product changes since the beginning of last year to improve health on your platform. Please provide a complete accounting of all 30-plus policy and product changes made.

The Honorable Cathy McMorris Rodgers

1. Your company is an immense supporter of net neutrality regulations being applied to Internet Service Providers (ISP's) based on Title II of the Telecommunications Act written in 1934. We can all agree to the basic principles of Net Neutrality and that there shouldn't be blocking, throttling or discrimination by those companies but I also believe that Title II is the wrong approach.

    a. Do you believe ISPs and Edge Providers have a different relationship to their customers? If so, do you believe that ISPs and Edge Providers have different responsibilities regarding the throttling and blocking of content?

    b. Can you identify who your customers are?

    c. Do you believe that standard consensus rules of the road regarding ISP throttling and blocking should apply only ISPs?

    d. Would you agree to have your company adhere to the same regulations you are advocating the ISP's live by? And if not, why?

2. On Privacy, there is clearly a significant public debate about private companies' access to, storage of, and monetization of people's data.

    a. What do you believe is the role of consumers regarding control of their information?

    b. Do you believe that we need greater transparency of private companies' privacy and data usage practices?

    c. Do you believe consumers should have greater control over their information? What do you make of proposed opt-in policies for Internet Service Providers?

    d. Do you believe ISPs and Edge Providers like you should be subject to the same or different standards as it relates to consumer privacy? If so, can you elaborate?

The Honorable David McKinley

1. A recent study in the American Journal of Public Health analyzed a number of tweets over 5 months and identified nearly 2000 sites linked to illicit online drug sales on Twitter. Your website states that this is prohibited. Can you tell me how many of those 2000 sites are still up?

2. Can you explain Twitter's current process for locating and removing criminal and prohibited activity on the platform? What actions are you taking to make this process more efficient and effective?

The Honorable Gus Bilirakis

1. I have heard from my local school districts that they must consistently respond to threats of school violence. In your response you mentioned that the company conducts outreach to local entities and law enforcement officials when you see anything impacting potential physical security threats. What are the specific implementations of that process?

2. How does Twitter determine the efficacy of the company's process for responding to potential threats?

3. You said you are always looking for ways to improve the process for quickly alerting the company to potential threats and responding with outreach to the proper authorities, and would be open to an implementation that can be scaled. In what ways would you consider working directly with local school districts and institutions, which play a key role in protecting its student body?

4. You mention that Twitter takes a data-driven approach to arranging Moments, based primarily on the amount of conversation happening on a particular topic or event, and then afterwards going into "impartiality" to provide many differing perspectives. Are there editorial standards the company has set for what makes a Moment, and if so, what are they? If no standards, why not, and would you consider using an official standard?

5. Regarding algorithms used to flag malicious behavior, what is the "error" rate, meaning how often are tweets overturned upon review by an employee content reviewer?

6. How many employees do you have to review content and make decisions on policing behavior? How long does it take on average to get an "error" overturned? Is there notification to user during the steps of this process?

7. Is the content reviewing algorithm technology unique to Twitter or is this something that is developing industry wide? If the former and it shows promise, is it something they are willing to share with competitors to provide a uniform industry standard?

The Honorable Susan Brooks

1. During the hearing, I talked about a bill of mine that passed into law in 2015, the DHS Social Media Improvement Act. The bill created a working group of relevant stakeholders to put their heads together to look at the impact social media has on preparedness, response, and recovery. The working group, housed in DHS' Science and Technology Directorate, also looks at how to counteract misinformation spread via Twitter during disasters and how first responders and communities can effectively use social media during times of crisis. This working group has to-date made 3 reports, which highlight countering false information in disasters and emergencies, best practices for incorporating social media into exercises, and how to operationalize social media for public safety. You said you were not aware of this working group, but would be willing to consider these reports. Below I have provided links to the reports. If you would like more information about the group, please reach out to my Legislative Assistant, Mimi Strobel (202.225.2276 or mimi.strobel@mail.house.gov).

   - April 2016 - **"From Concept to Reality: Operationalizing Social Media for Preparedness, Response and Recovery;"** (https://www.dhs.gov/sites/default/files/publications/SMWG_From-Concept-to-Reality-Operationalizing-Social-Media-508.pdf);

   - March 2017 - **"Best Practices for Incorporating Social Media into Exercises;"** (https://www.dhs.gov/publication/best-practices-incorporating-social-media-exercises); and,

   - March 2018 - **"Countering False Information on Social Media in Disasters and Emergencies"** (https://www.dhs.gov/sites/default/files/publications/SMWG_Countering-False-Info-Social-Media-Disasters-Emergencies_Mar2018-508.pdf).

The Honorable Earl "Buddy" Carter

1. Mr. Dorsey, you mentioned during my questioning that you wanted to reboot the verification process. Can you please describe the steps that Twitter is taking to overhaul that process? What will the new process look like? What is the timeline?

2. Mr. Dorsey, you mentioned that there is a new behavioral trend regarding algorithms and the illegal opioid sales present on your platform. You also mentioned that you need to look at how your algorithms are determining when they see this activity and can take action. Please describe how you will be reviewing your algorithms for their behavior and what changes you will be implementing.

3. Mr. Dorsey, intellectual property violations not only hurt those who directly develop this content, but also those industries that have developed around them. Across Georgia, thousands of well-paying jobs have been developed as a result of this growth. How is Twitter taking action to address intellectual property theft or rebroadcasting to keep up with changing practices? A quick search found numerous accounts that have been up for years with thousands of followers that continuously tweet out these links. How has Twitter not addressed these accounts and what steps will you take moving forward?

The Honorable Frank Pallone, Jr.

1. At the hearing, I asked you about Twitter's training and resources for content moderation. You could not provide specific responses at the time. Please provide specific responses to the following questions:

   a. How many human content moderators does Twitter employ in the U.S. and how much do they get paid?

   b. How many hours of training is given to them to ensure consistency in their decisions?

2. What steps is Twitter taking to improve the consistency of its enforcement and the metrics that demonstrate improvement?

The Honorable Debbie Dingell

1. What is the minimum number of individuals you allow to be targeted for advertisements?

2. Do you have any existing partnerships with device makers?

3. When you look at the performance of your AI, what is an acceptable error rate?

4. Will you be deploying any of these AI programs to combat election propaganda?

5. Because influence operations are always changing tactics, how often do you retrain your models to adapt to these changes?

6. Does Twitter have the capacity to monitor IP addresses that access multiple accounts to look for networks of malicious activity?

7. Like you, I believe understanding how AI comes to certain decisions is vital to building public trust. Although the study of this is in its infancy, what work is Twitter doing to move towards explainable AI?"

The Honorable Jerry McNerney

1. As I noted at the hearing, some social media platforms have recently been accused of facilitating discriminatory advertising—such as housing and employment ads. I want to know what Twitter is doing to identify and remedy any potential similar practices. During our exchange at the hearing, you stated "[w]e do regular audits of how our ads are targeted and how they're delivered and we work to make sure that we have fairness within them."

    a. Are these audits performed by Twitter employees or are they conducted by third parties?

    b. What role do algorithms have in carrying out these audits?

    c. Does each audit examine an individual ad? If not, how many ads are made part of each audit?

    d. How often are these audits conducted?

    e. When did Twitter begin conducting these audits?

    f. When you said "fairness within them" what did you mean by this?

    g. Do these audits specifically examine whether an ad is being targeted in a discriminatory way?

    h. Do these audits specifically examine whether any discriminatory effects may have resulted from Twitter's own ad optimization process?

    i. Are there any other steps that Twitter takes to determine whether ads are being targeted by advertisers in a discriminatory way? Please explain in detail what these steps entail.

j. Are there any other steps that Twitter takes to determine whether discriminatory effects have occurred as a result of its own ad optimization process? Please explain in detail what these steps entail.

k. What happens once an instance is identified in which an ad has been targeted in a discriminatory way or ultimately delivered in a way that resulted in discriminatory effects?

2. Does Twitter have any protocols in place for what actions should be taken when it is discovered that ads are being targeted in a discriminatory way or being delivered in a way that ultimately results in discriminatory effects? If so, please provide a copy of the documents that describe these protocols.

3. During the hearing, you acknowledged that it is possible for discrimination to result from how advertisers are able to target ads on your platform (specifically, that they are able to establish criteria that includes and excludes categories of users). Yet, as the company's CEO, you were unable to answer my question if Twitter has ever taken down an ad because of potential discriminatory effects and instead told me that you would have to follow up to get the information. It was incredibly troubling that you did not know the answer to this question after acknowledging that discrimination is possible.

a. Has Twitter ever taken down any ads because they were being targeted in a discriminatory way? If so, how many?

b. Has Twitter ever taken down any ads because of discriminatory effects that resulted from Twitter's own ad optimization process? If so, how many?

c. Has Twitter ever taken down any ads because the content of the ad was discriminatory? If so, how many?

4. One reason that it is difficult for us to know if ads on Twitter's platform are having discriminatory effects is because there is no real way for watchdog groups to help identify potential bias. When I asked you whether there is a way for watchdog groups to examine how non-political ads are being targeted, you stated "[y]es, our Ads Transparency Center is comprehensive of all ads." However, upon reviewing the information made available in Twitter's Ads Transparency Center, it appears that no information is available about how non-political ads are being targeted.

a. On what date do you plan to make available information about how non-political ads are being targeted and who is seeing the ads?

b. Why does the Ads Transparency Center only include information about ads during the last seven days? Do you plan to extend this time period?

5. When I asked you if Twitter is running any educational campaigns to inform users about how their data is being used, you stated "[n]ot at the moment, but we should be looking at that ..."

   a. What specific steps has Twitter taken since September 5, 2018 towards running educational campaigns that let users know what information it collects about them and how that information is used?

   b. Has Twitter taken any other steps since September 5, 2018 to improve how it informs users about what information is being collected and how that information is used?

   c. Will you commit to running educational campaigns so that users are able to better understand what information is being collected about them and how it is used?

   d. If so, on what date can we expect Twitter to launch these educational campaigns?

6. Many consumers do not realize that Twitter makes inferences about them. Buried at the bottom of one of its webpages titled "Your Twitter Data" is a link to "inferred interests from Twitter."

   a. Aside from how ads are targeted, what are other ways in which "inferred interests from Twitter" are used to personalize a user's experience?

   b. Is each user able to see all of the "inferred interests from Twitter" that have been made about them?

   c. If users are unable to see all of the "inferred interests from Twitter," how does Twitter determine which interests users are able to see?

   d. When making these inferences, does Twitter take into account the user's online activity on Twitter's platform as well online activity off of its platform?

   e. Will you commit to making available to users all of the "inferred interests from Twitter" that have been made about them?

7. Many consumers do not realize that Twitter collects "inferred interests from partners" about them. This information is similarly buried at the very bottom of the webpage titled "Your Twitter Data."

   a. Is each user able to view all of the "inferred interests from partners" that are used to reach them on Twitter?

   b. If users are unable to view all of the "inferred interests from partners," how does Twitter determine which ones users are able to see?

c. Will you commit to making available to users all of the "inferred interests from partners"?

8. Many consumers do not realize that Twitter can track their activity across the web.

   a. Is there a place on Twitter where users can go and see in plain text information about the websites that the user has visited?

   b. Will you commit to making this information available to users in plain text?

9. During the hearing, I asked whether Twitter stores previously collected data about users after the user chooses to no longer have his or her activity tracked. You said "I believe it's erased, but we'll have to follow up with the details."

   a. If a user disables "Track where you see Twitter content across the web," does Twitter still store any previously collected information?

   I also asked you if you would commit to erasing data when users opt out of the data being collected. You said "yes, but let just make sure I understand the constraints and ramifications of that." I want to make sure that you still stand by your commitment.

   b. Will you commit to giving users the option to have this information permanently deleted?

10. It is my understanding that Twitter also collects information both on and off its platform about non-users.

    a. How are non-users able to see all of the information that Twitter collects about them?

    b. How are non-users able to delete all of the information that Twitter collects about them?

The Honorable John Sarbanes

Media reports have indicated Twitter has become increasingly sensitive to allegations it suffers from anti-conservative basis – an allegation that continues to be advanced by many conservatives.[1] Some, including myself, have raised concern that Republican elected officials and conservative leaders are intentionally drumming up accusations of political bias at social

---

[1] Rogers, J. (2018, August 02). Twitter CEO gets why conservatives are suspicious of Big Tech. Retrieved September 19, 2018, from http://www.foxnews.com/tech/2018/08/02/twitter-ceo-gets-why-conservatives-are-suspicious-big-tech.html

media companies like Twitter in an effort to "work the refs," securing favorable treatment as recompense for non-existent bias.[2]

As our Committee works to understand better the policies in place at your company to prevent political bias, I am committed to better understanding how exactly those policies are being developed. To that end, I greatly appreciate Twitter's thorough and candid responses to the following questions.

1. It was reported in June that, in an effort to respond to complaints of bias against conservatives on Twitter, you met with conservative leaders and Republican officials, including Senator Ted Cruz and Trump Administration official Mercedes Schlapp.[3]

    a. Did these reported meetings take place? If so, what was the impetus for scheduling these meetings? Were they requested by you or by the persons with whom you met?

    b. What topics were discussed at these meetings? What specific requests were made of you at these meetings? Please provide any documents, meeting notes, or memoranda relating to the deliberation.

    c. To what extent was potential future regulation of social media platforms generally or Twitter specifically discussed with current government officials? Please provide any documents, meeting notes, or memoranda relating to the deliberation.

    d. To what extent have internal deliberations at Twitter regarding these accusations of bias focused on potential future regulation of social media platforms generally or Twitter specifically? Please provide any documents, meeting notes, or memoranda relating to the deliberation.

    e. Have any similar meetings with progressive or Democratic leaders taken place, been offered, or been requested? Please provide any documents, meeting notes, or memoranda relating to the deliberation.

    f. Have any similar meetings with victims' rights, anti-harassment, civil rights, or anti-hate group advocates, or any other individuals or groups concerned with rampant racist and misogynist harassment on Twitter taken place, been offered, or been requested? Please provide any documents, meeting notes, or memoranda relating to the deliberation.

[2] Usher, N. (2018, August 01) How Republicans trick Facebook and Twitter with claims of bias. Retrieved September 19, 2018, from https://www.washingtonpost.com/news/posteverything/wp/2018/08/01/how-republicans-trick-facebook-and-twitter-with-claims-of-bias/?utm_term=.a3de578781de

[3] Ghosh, S. (2018, June 28) Jack Dorsey secretly met Ted Cruz and other Republicans about Twitter's supposed liberal bias. Retrieved September 19, 2018, from https://www.businessinsider.com/jack-dorsey-twitter-secretly-republicans-liberal-bias-2018-6

g. Have any similar meetings with election security, campaign finance, or good government advocates, or any other individuals or groups concerned with Twitter's role in the functioning of our democratic institutions taken place, been offered, or been requested? Please provide any documents, meeting notes, or memoranda relating to the deliberation.

2. After reported "anti-conservative" bias, the accusations of which are questionable, at Facebook, specifically regarding its "Trending" section, Facebook removed human editors from the process and used algorithms to moderate the Trending section.[4] Almost immediately, the Trending section was overrun with false information and hoaxes masquerading as news stories.[5]

   a. To what extent has Facebook's experience with charges of anti-conservative bias, its overcorrection, and its resulting issues with "fake news" informed Twitter's decision-making regarding similar issues?

   b. What value do you, Twitter as a company, and Twitter's algorithms place on veracity? How does Twitter balance competing concerns over equal treatment of political viewpoints and legitimate concerns about misinformation and harassment?

   c. Is the deliberate spreading of misinformation a cause for suspending or banning a Twitter account? Please provide any documents, meeting notes, or memoranda relating to Twitter's formal policies on the subject.

   d. Does Twitter have rules regarding the veracity of promoted content? If so, what guidelines and processes are in place to evaluate the veracity of promoted content? If not, have there been any internal discussions or deliberations about moderating promoted content to prevent the spread of misinformation?

3. Relatedly, after Twitter took steps to remove "trolls" and fake accounts, conservatives accused Twitter of executing a purge of conservative accounts. In reality, Twitter was responding to concerns about widespread use of bots and fake accounts during the 2016 election.[6]

---

[4] Gershgorn, D., Murphy, M., Wong, J. (2016, August 26). Facebook is trying to get rid of bias in Trending news by getting rid of humans. Retrieved September 19, 2018, from https://qz.com/768122/facebook-fires-human-editors-moves-to-algorithm-for-trending-topics/

[5] Ohlheiser, A. (2016, August 29). Three days after removing human editors, Facebook is already trending fake news. Retrieved September 19, 2018, from https://www.washingtonpost.com/news/the-intersect/wp/2016/08/29/a-fake-headline-about-megyn-kelly-was-trending-on-facebook/?utm_term=.ac991f8ee14b

[6] Scola, N. (2018, February 21). Twitter purges accounts, and conservatives cry foul. Retrieved September 19, 2018, from https://www.politico.com/story/2018/02/21/twitter-purges-accounts-conservatives-357028

a. What do you and Twitter as a company consider Twitter's responsibilities and obligations to our democratic institutions in light of your platform's growing importance to campaigning and governing?

b. To what extent are you and Twitter as a company concerned that allegations of anti-conservative bias are an attempt to prevent Twitter from taking necessary steps to prevent the abuse of the platform that occurred during the 2016 election?

Please provide responses to these questions and any relevant documents, meeting notes, or memoranda relating to these matters.