

## **Questions from Senator Cruz**

### **I. Directions**

Please provide a wholly contained answer to each question. A question's answer should not cross-reference answers provided in other questions.

If a question asks for a yes or no answer, please provide a yes or no answer first and then provide subsequent explanation. If the answer to a yes or no question is sometimes yes and sometimes no, please state such first and then describe the circumstances giving rise to each answer.

If a question asks for a choice between two options, please begin by stating which option applies, or both, or neither, followed by any subsequent explanation.

If you disagree with the premise of a question, please answer the question as-written and then articulate both the premise about which you disagree and the basis for that disagreement.

If you lack a basis for knowing the answer to a question, please first describe what efforts you undertook as Public Policy Director for Facebook in order to ascertain an answer to the question and then provide your tentative answer as a consequence of its reasonable investigation. If even a tentative answer is impossible at this time, please state why such an answer is impossible and what efforts you and Facebook intend to take to provide an answer in the future. Please further give an estimate as to when the Committees shall receive that answer.

If it is impossible to answer a question without divulging confidential or privileged information, please clearly state the basis for confidentiality or privilege invoked and provide as extensive an answer as possible without breaching that confidentiality or privilege. For questions calling for answers requiring confidential information, please provide a complete answer in a sealed, confidential form. These materials will be kept confidential. For questions calling for privileged information, please describe the privileged relationship and identify the privileged documents or materials that, if disclosed, would fully answer the question.

If the answer to a question depends on one or more individuals' memory or beliefs and that individual or those individuals either do not recall relevant information or are not available to provide it, please state the names of those individuals, what efforts you undertook to obtain the unavailable information, and the names of other individuals who may have access to that information.

To the extent that an answer depends on an ambiguity in the question asked, please state the ambiguity you perceive in the question, and provide multiple answers which articulate each possible reasonable interpretation of the question in light of the ambiguity.

To the extent that a question inquires about your actions or Facebook's actions, omissions, or policies, the question also asks about any entities that Facebook owns, controls, or contracts with to provide services, including but not limited to services related to content moderation or advertising sales, including any and all subsidiaries and affiliates of Facebook or any contractor. If context suggests that a question may ask about Facebook as a service rather than as an entity,

please answer the question as applied to both Facebook as a service as well as all of Facebook's entities and platforms (e.g. Instagram, WhatsApp).

## **II. Questions**

- 1. Please state the number of users or advertisement purchasers elected to any political office or standing as a candidate for any political office in the United States (including any state, local, or municipal office) that have been banned, shadow banned, or in any other way had posts, content, or advertisements demoted, downgraded, restricted, or blocked (whether permanently or temporarily) by Facebook, any of its employees or contractors, or any algorithm designed by Facebook or any of its employees or contractors. In providing this answer, please include all incidents involving any restriction on content or advertising, even if Facebook subsequently reversed or altered its decision.**
  - a. Please provide a complete list of the above-described incidents, naming each user or advertisement purchaser affected, the post(s), content or advertisement(s) that lead to Facebook's decision, and the political affiliation of the user or advertisement purchaser elected to or standing for political office. If Facebook is unable to provide a complete list, please provide the most complete list possible after a reasonable and thorough investigation, including, without limitation, all such incidents that are already a matter of public record.**
  - b. Does Facebook take the political affiliation of any user or advertisement purchaser that is elected to political office or standing for political office in the United States into account when determining whether to take any adverse action regarding that user or advertisement purchaser? For purposes of this question, please disclose instances when any individual moderator has ever taken such factors into account in making the decision to restrict any content or advertising in any way on behalf of Facebook, even if such consideration was contrary to Facebook policy.**
  - c. Does Facebook require or provide any internal training or education to moderators or administrators of its platform regarding how to enforce Facebook's policies in a politically neutral manner? If so, please indicate whether this training is mandatory or optional, what positions at Facebook may or must attend such training, the frequency with which these positions are required or able to attend such training, and the nature, extent, and duration of the training.**
  - d. Does Facebook take the stance on any political issue—for example, abortion—that a user or advertisement purchaser that is elected to political office or standing for political office has adopted into account when determining whether to take any adverse action regarding that user or advertisement purchaser? For purposes of this question, please disclose whether any individual moderator has ever taken such factors into account**

**in making the decision to restrict any content or advertising in any way on behalf of Facebook, even if such consideration was contrary to Facebook policy.**

- e. Conversely, does Facebook take such adopted stances into account when providing advertisement rates, coverage, duration, or any other factor affecting the cost or quality of an advertisement on Facebook? Again, for purposes of this question, please disclose whether any of Facebook’s employees has ever taken such factors into account, even if such consideration was contrary to Facebook policy.**

While we do not typically comment on specific cases of content removal for privacy reasons, when we identify or learn of content that violates our policies, we remove that content regardless of who posted it. Decisions about whether to remove content are based on our Community Standards. The political affiliation of the user generating the content has no bearing on that content removal assessment. We have removed content posted by individuals and entities across the political spectrum.

On April 24, 2018, we published the detailed guidelines our reviewers use to make decisions about reported content on Facebook. These guidelines cover everything from nudity to graphic violence. Our Community Standards are global, and all reviewers use the same guidelines when making decisions.

We published these guidelines because we think it is important to provide clarity on where we draw lines on complex and continuously evolving issues, and we hope that sharing these details will prompt an open and honest dialogue about our decision-making process that will help us improve how we develop and enforce our standards. We make these guidelines public because we believe that the more companies are open about their policies, the more we can all learn from one another.

We recognize that our policies are only as good as the strength and accuracy of our enforcement—and our enforcement is not perfect. We make mistakes because our processes involve people, and people are not infallible. We are always working to improve.

With respect to training, our content reviewers undergo extensive training when they join and thereafter are regularly trained and tested with specific examples on how to uphold our Community Standards and take the correct action on a piece of content. This training occurs when policies are clarified, or as they evolve.

Our reviewers are not working in an empty room. There are quality control mechanisms as well as management onsite to help or provide guidance to reviewers if needed. When a reviewer is not clear on the action to take based on our Community Standards, they can pass the content decision to another team for review. We also audit the accuracy of reviewer decisions on an ongoing basis to coach them and follow up on improving when errors are made. And when we are made aware of incorrect content removals, we review them with our Community Operations team to prevent similar mistakes in the future.

We are also working to reduce unconscious bias. Our publicly available Managing Unconscious Bias class encourages our people to challenge and correct bias as soon as they see it—in others, and in themselves.

With respect to ads, people can run ads on Facebook, Instagram, and Audience Network on nearly any budget. The exact cost associated with an ad being shown to someone is determined in Facebook’s ad auction. Furthermore, the auction price depends on who the advertiser wants to reach and the advertiser’s objectives. It does not differ by virtue of whether the prospective purchaser is a Democrat or Republican.

- 2. Has Facebook ever conducted any investigation, whether formal, informal, or otherwise, to determine whether its content moderation policies or advertising rules have a disparate impact on users or advertisers based on partisan identity (e.g. Republican) or issue positions (e.g. pro-life)?**
  - a. If so, please provide the results of such investigation.**
  - b. If not, why not? Will Facebook conduct such an investigation and provide the results of that investigation?**
  - c. Has Facebook ever conducted any investigation, whether formal, informal, or otherwise, to determine whether its content moderation policies or advertising rules have a disparate impact on users who advocate for or against certain political or issue positions (e.g. abortion)?**
    - i. If so, please provide the results of such investigation.**
    - ii. If not, is Facebook willing to conduct such an investigation and provide its results?**

We engaged an outside advisor, former Senator Jon Kyl, to advise the company on potential bias against conservative voices. We believe this external feedback will help us improve over time and ensure we can most effectively serve our diverse community and build trust in Facebook as a platform for a broad spectrum of ideas. We also asked Laura Murphy, a highly respected civil rights and civil liberties leader, to guide a civil rights audit. After speaking with more than 90 civil rights organizations, Laura provided an important update on our progress in December 2018. The audit remains ongoing.

We continue to expand our list of outside partner organizations to ensure we receive feedback on our content policies from a diverse set of viewpoints.

We have made our detailed reviewer guidelines public to help people understand how and why we make decisions about the content that is and is not allowed on Facebook. And we have launched an appeals process so that people can contest content decisions with which they disagree. We are also instituting additional controls and oversight around the review team, including robust escalation procedures and updated reviewer training materials. These improvements and safeguards are designed to encourage free expression.

Suppressing content on the basis of political viewpoint or preventing people from seeing what matters most to them is directly contrary to Facebook’s mission and our business objectives.

- 3. Yes or No: Does Facebook consider itself a platform that is open to all ideas and all forms of expression that are protected by the First Amendment?**
  - a. Yes or No: Does Facebook consider itself to be a modern equivalent to the historical public square?**
  - b. Yes or no: Does Facebook consider itself to be a neutral public forum?**
  - c. When Facebook crafts its content moderation policies and advertising rules, does it seek to craft rules that are viewpoint neutral?**
  - d. In practice, does Facebook moderate content and enforce its advertising rules on a viewpoint-neutral basis?**
  - e. Has Facebook ever made any moderating decision or enforced its advertising rules in a non-viewpoint-neutral manner? Please describe all such incidents, even if they were contrary to Facebook policy.**

Facebook is first and foremost a technology company. We do not create or edit the content that our users post on our platform. While we seek to be a platform for a broad range of ideas, we do moderate content in good faith according to our published Community Standards in order to keep users on the platform safe, reduce objectionable content, and ensure users participate on the platform responsibly.

Freedom of expression is one of our core values, and we believe that the Facebook community is richer and stronger when a broad range of viewpoints is represented. We are committed to encouraging dialogue and the free flow of ideas by designing our products to give people a voice. We also know that people will not come to Facebook to share and connect with one another if they do not feel that the platform is a safe and respectful environment. In that vein, we have Community Standards that outline what is and is not allowed on Facebook.

We base our policies on principles of safety, voice, and equity. Our policy development is informed by input from our community and from experts and organizations outside Facebook so we can better understand different perspectives on safety and expression, as well as the impact of our policies on different communities globally. Based on this feedback, as well as changes in social norms and language, our standards evolve over time. Every two weeks, members of our Product Policy team, who sit in 11 offices around the world, run a meeting called the Product Policy Forum to discuss potential changes to our Community Standards, ads policies, and major News Feed ranking changes. A variety of subject matter experts participate in this meeting, including members of our safety and cybersecurity policy teams, counterterrorism specialists, Community Operations employees, product managers, public policy leads, and representatives from our legal, communications, and diversity teams. We publish the minutes from these meetings publicly (<https://newsroom.fb.com/news/2018/11/content-standards-forum-minutes/>)

and we hope that sharing these details will prompt an open and honest dialogue about our decision-making process that will help us improve in how we both develop and enforce our standards.

Decisions about whether to remove content are based on whether the content violates our Community Standards. Discussing controversial topics or espousing a debated point of view is not at odds with our Community Standards. We believe that such discussion is important in helping bridge division and promote greater understanding.

- 4. Has Facebook ever dismissed, demoted, fired, or otherwise taken adverse employment action against an employee on the basis of political speech that the employee undertook within the company, on Facebook, or elsewhere?**
  - a. If so, please list each such incident. If federal law requires Facebook to keep any of these incidents or their details confidential, please disclose as much information as federal law permits and anonymize the instances through appropriate pseudonyms and redactions. If Facebook does so, please note the legal basis for such redaction or confidentiality.**

We do not dismiss, demote, or fire employees on the basis of political speech.

- 5. Has Facebook ever dismissed, demoted, fired, or otherwise taken adverse employment action against an employee on the basis of that employee's discrimination against content or viewpoint within the company or on Facebook's platform?**
  - a. If so, please list each such incident. If federal law requires Facebook to keep any of these incidents or their details confidential, please disclose as much information as federal law permits and anonymize the instances through appropriate pseudonyms and redactions. If Facebook does so, please note the legal basis for such redaction or confidentiality.**

Content reviewers take action on content based on our Community Standards. Our Community Standards are global, and all reviewers use the same guidelines when making decisions. We seek to write actionable policies that clearly distinguish between violating and nonviolating content, and we seek to make the decision-making process for reviewers as objective as possible. We also audit the accuracy of reviewer decisions on an ongoing basis to coach them and follow up on improving when errors are made. And when we are made aware of incorrect content removals, we review them with our Community Operations team to prevent similar mistakes in the future.

Our policies are extremely granular because we want to ensure that the content review process is as objective as possible. Every week, we audit a sample of all reviewer decisions for accuracy and consistency. When a reviewer makes mistakes or misapplies our policies, we follow up with appropriate action. We also audit our auditors.

6. Does Facebook provide access to its services on a viewpoint-neutral basis? For this question and its subparts, please construe “access to its services” and similar phrases broadly, including the position or order in which content is displayed on its services, the position or order in which users or content appear in searches (or whether they appear at all), whether users or content are permitted to purchase advertisements (or be advertised), the rates charged for those advertisements, and so on.
- a. Yes or no: Has Facebook ever discriminated among users on the basis of viewpoint when determining whether to permit a user to access its services? If so, please list each instance in which Facebook has done so.
- i. If so, does Facebook continue to do so today, or when did Facebook stop doing so?
- ii. If so, what viewpoint(s) has Facebook discriminated against or in favor of? In what way(s) has Facebook done so?
- iii. If so, does Facebook consider only on viewpoints expressed on Facebook, or does it discriminate among users based on viewpoints expressed elsewhere? Has Facebook ever based its decision to permit or deny a user access to its services on viewpoints expressed off Facebook?
- iv. Yes or no: Excluding content encouraging physical self-harm, threats of physical violence, terrorism, and other content relating to the credible and imminent physical harm of specific individuals, has Facebook ever discriminated against users or their communications on the basis of viewpoint in its services? If so, please list each instance in which Facebook has done so.
- v. Yes or no: Has Facebook ever discriminated against American users or content on the basis of an affiliation with a religion or political party? If so, please list each instance in which Facebook has done so and describe the group or affiliation against which (or in favor of which) Facebook was discriminating.
- b. Yes or no: Has Facebook ever discriminated against any American users or content on its services on the basis of partisan affiliation with the Republican or Democratic parties? This question includes advocacy for or against a party or specific candidate or official. If so, please list each instance and the party affiliation discriminated against.
- c. Yes or no: Has Facebook ever discriminated against any American users or content on its services on the basis of the user’s or content’s advocacy for a political position on any issue in local, State, or national politics? This

**question includes but is not limited to advocacy for or against abortion, gun control, immigration, criminal justice reform, and net neutrality.**

- d. Yes or no: Has Facebook ever discriminated against any American users or content on its services on the basis of the user’s or content’s religion, including advocacy for one or more tenets of that religion? If so, please list each such instance in which Facebook has done so and identify the religion, religious group, or tenet against which Facebook discriminated.**

Because Facebook is a platform for a broad spectrum of ideas, we allow for discussion of controversial topics or points of view. We believe that such discussion is important in helping bridge division and promote greater understanding. We are committed to free expression and err on the side of allowing content. Our policies do not permit content to be removed because of a user’s political affiliation or religion. Decisions about whether to remove content are based on whether the content violates our Community Standards.

However, we recognize that our policies are only as good as the strength and accuracy of our enforcement—and our enforcement is not perfect. We make mistakes because our processes involve people, and people are not infallible. When we do make a mistake, we work to make it right. And we are committed to constantly improving our efforts so we make as few mistakes as possible.

We are also committed to designing our products to give people a voice and to foster the free flow of ideas and culture. But when content violates our Community Standards, that content has no place on Facebook. We work to remove it whenever we become aware of it.

- 7. Yes or no: Has Facebook ever discriminated between users in how their content is published, viewed, received, displayed in “trending” or similar lists, or otherwise in any function or feature, based on the user’s political affinity, religion, religious tenets, ideological positions, or any ideological or philosophical position asserted? This includes either the insertion of a topic or individual into the “trending” topics feature or the prohibition of a topic’s or individual’s display in the “trending” topics feature. If so, please list each such incident as well as the basis on which Facebook discriminated against that user or content.**

Suppressing content on the basis of political viewpoint or preventing people from seeing what matters most to them is directly contrary to Facebook’s mission and our business objectives.

When allegations of political bias surfaced in relation to Facebook’s Trending Topics feature, we immediately launched an investigation to determine if anyone violated the integrity of the feature or acted in ways that are inconsistent with Facebook’s policies and mission. We spoke with current reviewers and their supervisors, as well as a cross-section of former reviewers; spoke with our contractor; reviewed our guidelines, training, and practices; examined the effectiveness of operational oversight designed to identify and correct mistakes and abuse; and analyzed data on reviewers’ implementation of our guidelines.

Our investigation revealed no evidence of systematic political bias in the selection or prominence of stories included in the Trending Topics feature. In fact, our analysis indicated that the rates of approval of conservative and liberal topics were virtually identical in Trending Topics. Moreover, we were unable to substantiate any of the specific allegations of politically motivated suppression of subjects or sources, as reported in the media. To the contrary, we confirmed that most of those subjects were in fact included as trending topics on multiple occasions, on dates and at intervals that would be expected given the volume of discussion around those topics on those dates.

In 2016, Facebook met with Senator John Thune on this topic. We released our letter to Senator Thune and detailed our findings in a Newsroom Post. For more information, please see: <https://newsroom.fb.com/news/2016/05/response-to-chairman-john-thunes-letter-on-trending-topics/> and <https://www.commerce.senate.gov/public/cache/files/93a14e98-2443-4d27-bf04-1fc59b8cf2b4/22796A1389F52BE16D225F9A03FB53F8.facebook-letter.pdf>.

Moreover, in 2018, we removed the Trending Topics feature from Facebook because we found that users no longer found it useful.

- 8. How does Facebook moderate, prohibit, ban, or in any way otherwise restrict content or advertising that it considers to be “hate speech?”**
  - a. How does Facebook define the term “hate speech?”**
  - b. What objective metrics, if any, does Facebook use to determine whether a statement constitutes “hate speech?”**
  - c. To what extent does whether a statement constitutes “hate speech” depend on the subjective judgment of the moderator reviewing the content?**
  - d. What training, if any, does Facebook provide moderators in restricting content or users’ access to the platform on the basis of “hate speech” in a way that does not otherwise discriminate on the basis of viewpoint or partisan affiliation?**
  - e. Has Facebook ever changed its definition of “hate speech” or how it applies its hate speech policies? If so, please describe those changes.**
  - f. Does Facebook moderate, prohibit, ban, or in any way otherwise restrict content or advertising now on the basis of that ad or content being hate speech that it would have permitted at some previous time?**

We define hate speech as a direct attack on people based on what we call protected characteristics—race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disease or disability. We also provide some protections for immigrant status. For more information, please see: [https://www.facebook.com/communitystandards/hate\\_speech](https://www.facebook.com/communitystandards/hate_speech).

We recognize how important it is for Facebook to be a place where people feel empowered to communicate, and we take our role in keeping abuse off our platform seriously. That is why we have developed a set of Community Standards that outline what is and is not allowed on Facebook. Our Community Standards are designed to be comprehensive—for example, content that might not be considered hate speech may still be removed for violating our bullying policies. When we find things that violate our Standards, we remove them.

Our Community Standards are global and all reviewers use the same guidelines when making decisions. They undergo extensive training when they join and, thereafter, are regularly trained and tested with specific examples on how to uphold our Community Standards and take the correct action on a piece of content. This training occurs when policies are clarified, or as they evolve.

We seek to write actionable policies that clearly distinguish between violating and nonviolating content and we seek to make the decision-making process for reviewers as objective as possible.

Our reviewers are not working in an empty room. There are quality control mechanisms as well as management onsite to help or provide guidance if needed. When a reviewer is not clear on the action to take based on our Community Standards, they can pass the content decision to another team for review. We also audit the accuracy of reviewer decisions on an ongoing basis to coach them and follow up on improving when errors are made. And when we are made aware of incorrect content removals, we review them with our Community Operations team to prevent similar mistakes in the future.

We also introduced the right to appeal our decisions on individual posts, allowing users to ask for a second opinion when they think we have made a mistake. We believe giving people a voice in the process is another essential component of building a fair system.

We are constantly evaluating—and, where necessary, changing—our content policies to account for shifts in cultural and social norms around the world. For example, in August 2017, we expanded protections under our hate speech policies such that we now remove violent speech directed at groups of people defined by protected characteristics, even if the basis for the attack may be ambiguous. Under the previous hate speech policy, a direct attack targeting women solely on the basis of gender, for example, would have been removed from Facebook, but the same content directed at a sub-group, like “female drivers,” would have remained on the platform. We recognize that the distinction was overly narrow. As such, we no longer differentiate between the two forms of attack when it comes to violent hate speech.

**9. Did or does Facebook collaborate with or defer to any outside individuals or organizations in determining whether to classify a particular statement as “hate speech?” If so, please list the individuals and organizations.**

Hate speech has no place on our platform. Our Community Standards prohibit attacks based on characteristics including but not limited to race, ethnicity, religion, and national origin.

We speak with numerous organizations across the political spectrum to inform our policies, including our hate speech policy. But at the end of the day, we write and enforce our policies on our own, including our policies against hate speech. These policies are clearly laid out in our public Community Standards. For more information, please see: <https://www.facebook.com/communitystandards>.

As a matter of policy, we do not share the names of the groups we consult with for a number of reasons, including safety and security concerns—concerns which are especially acute in places around the world where the government may exercise censorship or control—and the fact that groups may not want to be named. That said, we typically engage with civil society organizations, activist groups, and thought leaders in areas including digital and civil rights, anti-discrimination, free speech, and human rights, as well as with academics who have relevant expertise.

**10. Did or does Facebook collaborate with or defer to any outside individuals or organizations in determining whether a given speaker has committed acts of “hate speech” in the past? If so, please list the individuals and organizations.**

**a. Does Facebook review these groups’ internal procedures in determining whether an entity is a “hate group” or committing acts of “hate speech” to determine that these determinations are not made on a partisan basis?**

In developing and iterating on our policies, including our hate speech policy, we consult with outside academics and experts from across the political spectrum and around the world, many of whom study organized hate groups and hate speech. While we do not share individual pieces of content from users with these organizations out of concerns for user privacy, we do provide in-depth examples and explanations of what the policy changes would entail. And we do not defer to these individuals or organizations when making decisions about content on our platform. Content that violates our Community Standards is removed when we are made aware of it, and content that does not violate our Community Standards is left on the platform.

As a matter of policy, we do not share the names of the groups we consult with for a number of reasons, including safety and security concerns—concerns which are especially acute in places around the world where the government may exercise censorship or control—and the fact that groups may not want to be named.

Regarding banning “hate groups,” we ban individuals or organizations that proclaim a violent or hateful mission or are engaged in acts of hate or violence. This is true regardless of ideology or motivation. We go through an extensive process to determine which people or groups we designate as dangerous, and consider a number of signals including:

- Whether they have called for or directly carried out acts of violence against people based on factors like race, ethnicity, or national origin.
- Whether they are a self-described or identified follower of a hateful ideology.

- Whether they use hate speech or slurs in their “About” section on Facebook or Instagram.
- Whether they have had Pages or Groups removed from Facebook, or accounts removed from Instagram, for posting content that goes against our hate speech policies.

**11. Under what circumstances does Facebook ban or otherwise limit the content of individuals or organizations who have spoken “hate speech” on its platform aside from the offending content?**

We believe in giving people a voice, but we also want everyone using Facebook to feel safe. That is why we have Community Standards and remove content that violates them, including hate speech.

But sometimes simply removing content that violates our Standards is not enough to deter repeat offenders. That is why every time we remove something, it counts as a strike against the person who posted it. And when it comes to Pages, we hold both the entire Page and the person who posted the content accountable. More specifically:

- If a Page posts content that violates our Community Standards, the Page and the Page admin responsible for posting the content receive a strike.
- When a Page surpasses a certain threshold of strikes, the whole Page is unpublished.
- For people, including Page admins, the effects of a strike vary depending on the severity of the violation and a person’s history on Facebook. For example, some content is so bad that posting it just once means we would remove the account immediately. In the case of other violations, we may warn someone the first time they break our Community Standards. If they continue, we may temporarily block their account, which restricts their ability to post on Facebook, or remove it all together.

Because we do not want people to game the system, we do not share the specific number of strikes that leads to a temporary or permanent suspension.

**12. Facebook is not subject to the First Amendment’s limitations against government censorship, and is free to moderate content as it sees fit in the same way that the New York Times or Wall Street Journal do.**

- a. As Facebook defines “hate speech,” does Facebook believe that its hate speech policy affects content that would be protected from government censorship by the First Amendment?**

**b. If so, please describe what content would be subject to Facebook’s policy that is nonetheless protected from government censorship by the First Amendment.**

The goal of our Community Standards is to encourage expression and create a safe community for our 2 billion users, more than 87% of whom are located outside the United States.

We err on the side of allowing content, even when some find it objectionable, unless removing that content prevents a specific harm.

We do not allow hate speech on Facebook because it creates an environment of intimidation and exclusion and, in some cases, may promote real-world violence.

Our current definition of “hate speech” is anything that directly attacks people based on what are known as their “protected characteristics”—race, ethnicity, national origin, religious affiliation, caste, sexual orientation, sex, gender, gender identity, and serious disability or disease. We also provide some protections for immigration status. However, our definition does allow for discussion around these characteristics as concepts in an effort to allow for and encourage expression and dialogue by our users.

**13. Yes or no: Has Facebook ever removed content for “hate speech” that did not directly attack or threaten a person on the basis of his or her race, ethnicity, national origin, religious affiliation, sexual orientation, sex, gender, or gender identity, or serious disabilities or diseases? If so, what criteria did Facebook use to determine that the content violated Facebook’s policy?**

We define hate speech as a direct attack on people based on what we call protected characteristics—race, ethnicity, national origin, religious affiliation, caste, sexual orientation, sex, gender, gender identity, and serious disability or disease. We define “attack” as violent or dehumanizing speech, statements of inferiority, and calls for exclusion or segregation. Under our policy, such content does not need to attack or threaten a person directly on the basis of a protected characteristic for it to constitute hate speech.

Sometimes it is obvious that something is hate speech and should be removed—because it includes the direct incitement of violence against people possessing protected characteristics, or degrades or dehumanizes people. Sometimes, however, there is not a clear consensus—because the words themselves are ambiguous, the intent behind them is unknown, or the context around them is unclear. Language also continues to evolve, and a word that was not a slur yesterday may become one today.

Here are some of the things we take into consideration when deciding what to leave on the site and what to remove:

- **Context:** Regional and linguistic context is often critical in deciding whether content constitutes hate speech, as is the need to take geopolitical events into account.

- **Intent:** There are times someone might share something that would otherwise be considered hate speech but for non-hateful reasons, such as making a self-deprecating joke or quoting lyrics from a song. People often use satire and comedy to make a point about hate speech. In other cases, people may speak out against hatred by condemning someone else’s use of offensive language, which requires repeating the original offense. This is something we allow, even though it means some users may encounter material disturbing to them, because it gives our community the chance to speak out against hateful ideas. We revised our Community Standards to encourage people to make it clear when they are sharing something to condemn it, but sometimes when their intent is not clear, anti-hatred posts get removed in error.

**14. Can expressing a controversial opinion itself—when not transmitted to a particular user or indicated as directed at a particular individual, given the circumstances—count as a “direct attack or threat” that violates Facebook’s “hate speech” policy?**

Discussing controversial topics or espousing a debated point of view is not at odds with our Community Standards, which include our hate speech policies. We believe that such discussion is important in helping bridge division and promote greater understanding. We are committed to free expression and err on the side of allowing content.

But when something crosses the line into hate speech, it has no place on Facebook, and we are committed to removing it from our platform any time we become aware of it. We define hate speech as a direct attack on people based on what we call protected characteristics—race, ethnicity, national origin, religious affiliation, caste, sexual orientation, sex, gender, gender identity, and serious disability or disease. We also provide some protections for immigration status. We define “attack” as violent or dehumanizing speech, statements of inferiority, and calls for exclusion or segregation. Our Community Standards, the detailed guidelines our reviewers use to assess whether content violates our hate speech policies, are available here:

[https://www.facebook.com/communitystandards/objectionable\\_content/hate\\_speech](https://www.facebook.com/communitystandards/objectionable_content/hate_speech).

**15. You acknowledged that a given quote by Mother Theresa was not “hate speech.” Would any of the statements below, standing alone, violate Facebook’s “hate speech” policy?**

- There are only two sexes or two genders, male and female.**
- Sex reassignment surgery is a form of bodily mutilation.**
- The abortion of an unborn child is murder.**
- Same-sex marriage is wrong.**
- No person of faith should be required to assist a same-sex wedding by providing goods or services to a same-sex marrying couple.**
- Islam is a religion of war.**

- g. All white people are inherently racist.**
- h. Donating to the NRA funds the murder of children, such as those slain in Parkland, Florida.**
- i. The U.S. should build a wall at its southern border.**
- j. Illegal aliens need to be sent back to their home countries.**

As they are stated here, none of these statements violates our Content Policies. We allow broad discussion and criticism of ideas and institutions, like same-sex marriage, structural racism, immigration policy, and the religion of Islam. We believe that such discussions are an important way of bridging division and promoting greater understanding. It is when those statements rise to the level of attacks on people—violent or dehumanizing speech, statements of inferiority, and calls for exclusion or segregation—that they violate our policies and are removed accordingly.

Context matters in making what can be difficult determinations in some cases. Sometimes it is obvious that something is hate speech and should be removed—because it includes the direct incitement of violence against people possessing protected characteristics, or degrades or dehumanizes people. Sometimes, however, there is not a clear consensus—because the words themselves are ambiguous, the intent behind them is unknown, or the context around them is unclear. Language also continues to evolve, and a word that was not a slur yesterday may become one today.

The statement “burn flags not fags” offers a poignant example. While the statement is certainly provocative on its face, should it be considered hate speech? Is it an attack on gay people, or an attempt to “reclaim” the slur? Is it an incitement of political protest through flag burning? Or, if the speaker or audience is British, is it an effort to discourage people from smoking cigarettes (fag being a common British term for cigarette)? To know whether the statement is hate speech, more context might be needed.

- 16. Has Facebook ever studied or examined, whether formally, informally, or otherwise, the political beliefs or affiliations of its users? If so, please disclose the results of those studies or examinations.**
- a. How has Facebook used this information? Please explain each use of this information. If any uses contain information that would be protected by law as proprietary or trade secrets, please inform us so that we may arrange for procedures to keep this information appropriately confidential.**
  - b. Has Facebook ever reviewed or made use of third-party studies or examinations of the political affiliations of its users? If so, please explain when and how, including in what ways these conclusions affected Facebook’s policies or how Facebook enforces its policies.**

Facebook is able to view any information a user adds to “Political Views” in the “About” section of Timeline. Users can download their own Political Views information, as well as other information associated with their Facebook accounts, through our Download Your Information tool. We also introduced Access Your Information—a secure way for people to access and manage their information, such as posts, reactions, comments, and things they have searched for. Users can go here to delete anything from their Timelines or profiles that they no longer want on Facebook.

If someone adds this information to their profile, they can later choose to delete it. If they do so, we will remove it from our site and delete it in accordance with our Data Policy.

We prompt people on Facebook who have added a political affiliation to their profiles to review this information and decide whether they want to keep it on their profiles. More information about these prompts is available at <https://newsroom.fb.com/news/2018/05/pardon-the-interruption/>.

**17. Has Facebook ever conducted any study or investigation, whether formal, informal, or otherwise, the level of engagement that Facebook users have with accounts held by individuals who are elected to or standing for any political office in the United States? If so, please provide the results of such investigation.**

Every day, people use Facebook to engage with their elected officials and make their voices heard on issues they care about. For example, in a single month shortly before the 2018 midterm elections, over 4 million people in the United States commented on, reacted to, or shared a post by one of their elected officials, including the 1.5 million people who interacted with a state or local elected official. And over 25 million people in the US now follow at least one of their elected officials on Facebook.

As part of our effort to foster civically engaged communities on Facebook, we have developed tools to help people learn about different candidates and get information on when and where to vote ahead of Election Day. That includes Town Hall, which allows people to easily find and contact their elected officials, and Candidate Info, which lets people hear directly from their federal, state and local candidates on why they are running for office, what policy issues they care about, and what they hope to accomplish if elected. In developing these products, we learned that what people value most is hearing directly from candidates in their own words. This feedback informed our Candidate Info tool, which shows both information about the candidates as well as videos created by the candidates themselves.

**18. Under what circumstances does Facebook either ban content criticizing Facebook’s decision to restrict content or users, or otherwise require users to remove such content critical of Facebook as a condition of using the platform?**

We are an open platform for a broad spectrum of ideas, a place where we want to encourage self-expression, connection, and sharing. At the same time, when people come to Facebook, we always want them to feel welcome and safe. That is why we have rules against bullying, harassing, and threatening someone.

Our Community Standards and Ads Policies outline the content that is not allowed on the platform, such as hate speech; fake accounts; and praise, support, or representation of terrorism/terrorists. When we find content that violate these Standards, we remove it. There are other types of problematic content that, although they do not violate our policies, are still misleading or harmful and that our community has told us they do not want to see on Facebook—things like clickbait or sensationalism. When we find examples of this kind of content, we reduce its spread in News Feed using ranking and, increasingly, we inform users with additional context so they can decide whether to read, trust or share it.

The goal of our Community Standards is to encourage expression and create a safe environment. We base our policies on input from our community and from experts in fields such as technology and public safety. Our policies are also rooted in the following principles:

- **(1) Safety:** People need to feel safe in order to build community. We are committed to removing content that encourages real-world harm, including (but not limited to) physical, financial, and emotional injury;
- **(2) Voice:** Our mission is all about embracing diverse views. We err on the side of allowing content, even when some find it objectionable, unless removing that content can prevent a specific harm. Moreover, at times we will allow content that might otherwise violate our standards if we feel that it is newsworthy, significant or important to the public interest. We do this only after weighing the public-interest value of the content against the risk of real-world harm; and
- **(3) Equity:** Our community is global and diverse. Our policies may seem broad, but that is because we apply them consistently and fairly to a community that transcends regions, cultures, and languages. As a result, our Community Standards can sometimes appear less nuanced than we would like, leading to an outcome that is at odds with their underlying purpose. For that reason, in some cases when we are provided with additional context, we make a decision based on the spirit, rather than the letter, of the policy.

**19. How many individuals at Facebook have the ability to moderate, remove, downgrade, conceal, or otherwise censor content, ban, suspend, warn, or otherwise discipline users, or approve, price, review, or refuse advertisements on the platform? (For this question only, we refer to these individuals as moderators.) This question includes individuals with the power to alter search results and similar mechanisms that suggest additional content to users in order to promote or demote content, whether individually or routinely through an algorithm or by altering any of the platform’s search functions. (Please include all employees, independent contractors, or others with such ability at Facebook.)**

- a. How many moderators work for Facebook? This includes individuals who serve in moderating functions part-time or as independent contractors. This question includes individuals with the power to alter search results and similar mechanisms that suggest additional content to users in order to**

**promote or demote content, whether individually or routinely through an algorithm or by altering any of the platform’s search functions.**

- b. Who are the individuals responsible for supervising these moderators as their conduct relates to American citizens, nationals, businesses, and groups?**
- c. On average, how many pieces of content does a moderator remove a day?**
- d. On average, how many users does a moderator discipline a day?**
- e. On average, how many advertisements does a moderator approve, disapprove, price, consult on, review, or refuse a day?**

We employ more than 30,000 people at Facebook who work on safety and security—about half of whom are content reviewers. Our content review teams work around the world, 24 hours a day and in dozens of languages, to review content. Those reviewers respond to more than two million pieces of content every day from people all over the world. We issue a transparency report with a more detailed breakdown of the content we take down.

Each day can be slightly different, with shifts lasting no more than 8 hours, and much less than 8 hours being spent reviewing content. A typical day would include elements such as reviewing content, receiving coaching, taking mandatory and wellness breaks, having lunch, participating in team huddles or meetings, or training.

With respect to the review of ads, we are also committed to getting better at enforcing our advertising policies. We review many ads proactively using automated and manual tools, and reactively when people hide, block, or mark ads as offensive. We are taking aggressive steps to strengthen both our automated and our manual review. We are also expanding our global ads review teams and investing more in machine learning to better understand when to flag and take down ads, such as ads that offer employment or credit opportunity, while including or excluding multicultural advertising segments. Enforcement is never perfect, but we will get better at finding and removing improper ads.

- 20. As Facebook has previously acknowledged, Silicon Valley is predominantly politically liberal, and Facebook’s employees are likewise predominantly liberal. Republicans and conservatives are concerned that such a political monoculture leads to disproportionate sanctions against conservatives and conservative views, such as those researchers find prevail in academia. To Facebook’s credit, it has devoted significant resources to hearing Republicans and conservatives out regarding our concerns about this potential basis of bias.**

**To that end, please answer the following questions based on any information Facebook has, whether formal or informal, as to the political beliefs or political involvement of Facebook’s personnel. If Facebook requires more time to gather this information, please let us know when we can expect a response.**

- a. **What percentage of Facebook’s Board of Directors self-identify as “liberal” or Democrats versus “conservative” or Republicans?**
- b. **How many of Facebook’s Board of Directors have donated or raised money for Democrats, the Democratic National Committee, or political action committees primarily supporting Democrats? For Republicans and their counterparts?**
- c. **What percentage of Facebook’s senior management have worked in Democratic administrations? In Republican administrations?**
- d. **What percentage of Facebook’s senior management self-identify as “liberal” or Democrats versus “conservative” or Republicans?**
- e. **How many of Facebook’s senior management have donated or raised money for Democrats, the Democratic National Committee, or political action committees primarily supporting Democrats? For Republicans and their counterparts?**
- f. **What percentage of Facebook’s senior management have worked in Democratic administrations? In Republican administrations?**

We do not maintain statistics on these data points.

- 21. Does Facebook conduct any voter outreach, for example encouraging users to vote in an election or register to vote in elections?**
- a. **If so, does Facebook consider the political party of those reached by its voter outreach efforts when designing or engaging in those efforts?**
  - b. **If so, do Facebook’s voter outreach efforts disparately reach registered Democrats or Republicans?**
  - c. **Please list each such voter outreach effort that Facebook has conducted, including the year, the election, and the candidates in that election, and the means and extent to which Facebook engaged in voter outreach.**
  - d. **Has Facebook or any employees, contractors, or subsidiaries ever engaged in any voter outreach in order to influence the outcome of any election?**
  - e. **Has Facebook ever conducted any investigation, whether formal, informal, or otherwise, to determine the political leanings or party affiliation of the users it reaches with voter outreach efforts?**
  - f. **Has Facebook ever conducted any investigation, whether formal, informal, or otherwise, to determine the political leanings or party affiliation of the users that respond to or interact with voter outreach efforts?**

- g. Has Facebook ever conducted specific voter outreach efforts to reach any identifiable demographic, including by race, sex, nationality, sexual orientation or gender identity, marital status, geography, language, or age?**
- h. Has Facebook ever used its platform to influence public debate or the outcome of an election, either through direct communications or through the enforcement of its policies?**

We want all candidates, groups and voters to use our platform to engage in elections. We want it to be easy for people to find, follow, and contact their elected representatives and those running to represent them. That is why, for candidates across the political spectrum, Facebook offers the same levels of support in key moments to help campaigns understand how best to use the platform.

As part of our effort to foster civically engaged communities on Facebook, we have also developed non-partisan tools to help people learn about different candidates and get information on when and where to vote ahead of Election Day. In October 2018, we unveiled a new tool, Candidate Info, which lets people hear directly from their federal, state, and local candidates about why they are running for office, what policy issues they care about, and what they hope to accomplish if elected.

As part of our ongoing efforts to prevent people from misusing Facebook during elections, we are broadening our policies against voter suppression—action that is designed to deter or prevent people from voting. These updates were designed to address new types of abuse that we are seeing online.

We prohibit offers to buy or sell votes as well as misrepresentations about the dates, locations, times, and qualifications for casting a ballot. And we expressly ban misrepresentations about how to vote, such as claims that you can vote using an online app and statements about whether a vote will be counted (e.g. “If you voted in the primary, your vote in the general election won’t count.”). We have also recently introduced a new reporting option on Facebook so people can let us know if they see voting information that may be incorrect. And we have set up dedicated reporting channels for state election authorities so that they can do the same.

In addition to working to prevent voter suppression, we are also building on our non-partisan efforts to encourage voter registration and engagement. When people turn 18, and ahead of elections, we remind them to register to vote. We help them find their polling places and remind them to vote on Election Day. Last year, we also added a feature that lets people ask their friends to join them in registering to vote. As a result of these efforts, Facebook and Instagram helped register an estimated 2 million people in 2018, according to our nonpartisan partner TurboVote.

## **Questions from Senator Hirono**

### **1. With regard to Facebook's content moderation practices:**

- a. How many content moderators does Facebook employ worldwide? Please provide the total number content moderators along with a breakdown by country of residence, by state of residence (if country of residence is the United States), and by employment status (i.e., how many content moderators are Facebook employees v. contractors).**

We have invested significantly in safety and security and now have over 30,000 people working in this area, about half of whom review content. The majority of our content reviewers are people who work full-time for our partners and work at sites managed by these partners. We have a global network of partner companies so that we can quickly adjust the focus of our workforce as needed. This approach gives us the ability to, for example, make sure we have the right language or regional expertise—and allows us to quickly hire in different time zones. Our partners have a core competency in this type of work and are able to help us adjust as new needs arise or when a situation around the world warrants it.

We have just over 20 content review sites around the world in countries including Germany, Ireland, Latvia, Spain, Portugal, the Philippines, and the United States. Our reviewers come from many backgrounds; reflect the diversity of our community; bring a wide array of professional experiences, from military veterans to former public sector workers; and are native language speakers.

Although our content review operation is global, we do not have content review sites in all locations for the languages we support. We focus on having centralized locations, which allows for increased infrastructure and support, including onsite leadership to answer questions, market specialists to address uncertain review decisions, training, and especially support and resiliency programs. These offices look and feel like Facebook offices and have many of the same amenities.

- b. Please describe the training provided to content moderators.**

Our content reviewers undergo extensive training when they join, with over 80 hours of instructor-led, hands-on learning and shadowing of veteran reviewers. They are trained and tested with specific examples on how to uphold the Community Standards and take the correct action on a piece of content. There is also ongoing training when policies are clarified, or as they evolve.

We are always working to improve our operations and the training and support that are provided to each person that reviews content on behalf of Facebook. Some of these initiatives include improving training materials to include more multimedia to support all learning types; providing additional training resources for well-being, resiliency, and unconscious biases; and providing additional marketized examples for our global network of content reviewers.

**c. What is the average salary of a content moderator?**

In 2015, we introduced a new set of standards for people who do contract work in the US, including: a \$15 minimum wage; a minimum 15 paid days off for holidays, sick time, and vacation; and, for new parents that do not receive paid leave, a \$4,000 new child benefit that gives them the flexibility to take paid parental leave. Since 2016, we have also required vendors in the US to provide comprehensive healthcare to all of their employees assigned to Facebook.

In the years since, it has become clear that \$15 per hour does not meet the cost of living in some of the places where we operate. After reviewing a number of factors including third-party guidelines, we are committing to a higher standard that better reflects local costs of living. This means a raise to a minimum of \$20 per hour in the San Francisco Bay Area, New York City, and Washington, D.C., and \$18 per hour in Seattle. We will be implementing these changes by mid-next year, and we are working to develop similar standards for other countries.

For workers in the US that review content on Facebook, we are raising wages even more. Their work is critical to keeping our community safe, and it is often difficult. That is why we have paid content reviewers more than minimum wage standards, and why we will surpass this new living wage standard as well. We will pay at least \$22 per hour to all employees of our vendor partners based in the Bay Area, New York City, and Washington, D.C.; \$20 per hour to those living in Seattle; and \$18 per hour in all other metro areas in the US. As with all people who do contract work, we are working to develop similar international standards. This work is ongoing, and we will continue to review wages over time.

**d. On average, how many hours per week does a content moderator work?**

Each day can be slightly different, with shifts lasting no more than 8 hours, and much less than 8 hours being spent reviewing content. A typical day would include elements such as reviewing content, receiving coaching, taking mandatory and wellness breaks, having lunch, participating in team huddles or meetings, or training. Counseling is also offered onsite during the day.

**e. On average, how many posts, likes, status updates, etc. does a content moderator review per week?**

In total, content reviewers review more than two million pieces of content every day. We issue a transparency report with a more detailed breakdown of the content we take down.

The latest transparency report was just released on May 23, 2019 and can be found at <https://transparency.facebook.com/community-standards-enforcement>.

**f. On average, how much time does a content moderator have to determine if a post, like, status update, etc. violates Facebook's Community Standards?**

Reports are reviewed 24 hours a day, 7 days a week, and the vast majority of reports are reviewed within 24 hours. Content reviewers are not required to evaluate any set number of

posts—for example, nudity is typically very easy to establish and can be reviewed within seconds, whereas something like impersonation could take much longer to confirm. We provide general guidelines for how long we think it might take to review different types of content to make sure that we have the staffing we need, but we encourage reviewers to take the time they need.

We are continually working to find the right balance between content reviewer well-being and resiliency, quality, and productivity to ensure that we are getting to reports as quickly as possible for our community.

**g. What percentage of content moderators have reported a diagnosis of or symptoms of post-traumatic stress disorder (PTSD), drug abuse, anxiety, and/or another psychological disorder as a result of their work?**

The safety and well-being of all our content reviewers is the highest priority.

All content reviewers—whether full-time employees or those employed by partner companies—have access to well-being and resiliency resources. This includes access to trained professionals for individual and group counseling. And, as with all people doing contract work, content reviewers also have comprehensive healthcare benefits.

We have a team of four clinical psychologists across three regions who are tasked with designing, delivering, and evaluating resiliency programs for everyone who works with objectionable content. This group works closely with our vendor partners and each of their dedicated resiliency professionals to help build resiliency programming standards for their teams and share best practices. We collaborate with our partners to ensure they are providing the necessary levels of support, including psychological support, to anyone reviewing Facebook content. In addition, Facebook actively requests and funds an environment that ensures this support is in place for the reviewers employed by our partners. This includes the environment they work in, with contractual expectations around space for resiliency and wellness, wellness support and benefits including health care, paid time off, and bonuses.

Generally, our partners must provide a resiliency plan that is reviewed and approved by Facebook. This includes a holistic approach to well-being and resiliency that puts the needs of their employees first. If someone is affected by the content that they are reviewing, they can get up and take an immediate break, go to a space that is dedicated for well-being or request from their manager to review another content type, if that opportunity exists. Counselors are also onsite for reviewers to talk to.

In addition, we are also employing technical solutions to limit reviewers' exposure to graphic material as much as possible. For the first time, we are adding preferences that let reviewers customize how they view certain content. For example, they can now choose to temporarily blur graphic images by default before reviewing them. We made these changes after hearing feedback that reviewers want more control over how they see content that can be challenging. Content review at our size can be challenging and we know we have more work to do. We are committed to supporting our content reviewers in a way that puts their well-being first.

## **Questions from Senator Blumenthal**

1. **Last May, Facebook committed to conducting an independent audit of its civil rights issues on its platform. While it has labeled the audit a top priority for this year, it has not announced changes aside from those required by lawsuit settlements or Congressional and public pressure.**
  - a. **What changes has Facebook made specifically as a result of the civil rights audit?**

In May 2018, we accepted the call to undertake a civil rights audit. We asked Laura Murphy, a highly respected civil rights and civil liberties leader, to guide the audit. After speaking with more than 90 civil rights organizations, Laura provided an important update on our progress in December 2018. The audit remains ongoing.

Laura's work has helped us build upon crucial election-related efforts, such as expanding our policy prohibiting voter suppression. We updated our policy to expressly ban misrepresentations about how to vote, such as claims that you can vote using an online app and statements about whether a vote will be counted. The revised policy also prohibits threats of violence related to voting or voter registration. As a direct response to feedback from civil rights advocates, we are focusing on voter suppression as a distinct civil rights challenge.

In addition to working to prevent voter suppression, we are also building on our efforts to encourage voter registration and engagement. We remind people on Facebook to register to vote when they turn 18 and ahead of elections. We also help them find their polling places. And last year, we added a feature that lets people ask their friends to join them in registering to vote. As a result of these efforts, Facebook and Instagram helped register an estimated 2 million people in 2018, according to our nonpartisan partner TurboVote.

We have also made changes to our policies against hate and discrimination as a result of the civil rights audit. In March 2019, for example, we announced a ban on praise, support, and representation of white nationalism and white separatism on Facebook and Instagram. The decision came after more than 20 conversations with members of civil society and academics who are experts in race relations around the world. They confirmed what our own data showed—that white nationalism and white separatism cannot be meaningfully separated from white supremacy and organized hate groups.

Other recent changes we made that address priorities raised by the civil rights community include:

- **Commerce Discrimination Policy:** We implemented a new policy which prohibits discriminatory language in users' commerce-related posts on Facebook's Marketplace and Buy-Sell Groups. The policy complements the existing prohibitions we maintain against discriminatory conduct in our Advertising Policies.

- **Community Standards:** In April 2018, Facebook released a more detailed version of its Community Standards, including the internal review guidelines we use to enforce our standards. In tandem, Facebook made available content-level appeals, which are designed to enable users to challenge content decisions at the post level.
  - **Community Standards Enforcement Report:** In May 2018, we published our first ever Community Standards Enforcement Report, which included details on our enforcement efforts across graphic violence, adult nudity and sexual activity, terrorist propaganda, hate speech, spam, and fake accounts. Since then, we have built out the Community Standards Enforcement Report such that the most recent version, published on May 23 of this year, includes metrics across nine policy areas, as well as data on the number of appeals we received and the proportion of those appeals that prompted us to reverse our initial decision.
  - **Ads Transparency for Users:** In June 2018, we released a feature that allows users to view all ads that an advertiser is running. This furthers our efforts to combat discrimination by allowing people to see ads regardless of whether they were included in the target audience selected by the advertiser. Users can also report ads to the company, further curtailing advertisers' potential misuse of Facebook's tools.
  - **New, Mandatory Non-Discrimination Certification:** In 2018, we expanded the non-discrimination certification process requirement for advertisers. Previously, this certificate only applied to advertisers placing housing, employment, and credit ads. The new policy covers all US advertisers placing any ad on Facebook.
  - **Ads Settlement Agreement:** In March 2019, we announced changes to the way we manage housing, employment, and credit ads on our platform as part of historic settlement agreements with leading civil rights organizations and based on ongoing input from civil rights experts. Anyone who wants to run housing, employment, or credit ads are no longer allowed to target by age, gender or zip code. Additionally, any detailed targeting option describing or appearing to relate to protected classes are also unavailable. And we are building a tool so you can search for and view all current housing ads in the US targeted to different places across the country, regardless of whether the ads are shown to you. We are committed to doing more to protect against discrimination in ads, and we look forward to engaging in serious consultation and work with key civil rights groups, experts, and policymakers to help us find the right path forward.
- b. **What issues have been identified by the audit that Facebook plans to address?**

Laura's report from December includes several areas of focus, and we are working hard to address each. The specific points Laura's update identified were as follows:

- **Voter Suppression:** Protecting against misuse of the platform to intimidate or suppress voter participation, particularly among minority groups and people of color. As discussed in Response to Question 1.a above, we have taken steps to expand on our policies against voter suppression, but recognize that we must continue to focus on voter suppression as a distinct civil rights challenge.
  - **Accountability Infrastructure:** Instituting sufficient protocols and a civil rights infrastructure to ensure civil rights are considered in the development of products, services or policies before they are rolled out.
  - **Content Moderation and Enforcement:** Protecting users in minority groups from hateful and/or racist expression and ensuring that content policies are equitably enforced, and that voices of activists and civil rights advocates are not unfairly censored. As part of this effort, Laura has called for increased transparency in our policy-making process, enforcement, and operations.
  - **Diversity and Inclusion:** Promoting greater employee diversity in all functions and levels and taking more steps to create a more inclusive workplace.
  - **Privacy:** Developing comprehensive privacy measures that protect civil rights while preventing unlawful discrimination are top-of-mind for advocates in the civil rights community.
  - **Fairness in Artificial Intelligence and Algorithms:** Devoting resources to help ensure that artificial intelligence tools, such as machine learning or facial recognition, do not facilitate bias and are fair to all users.
- c. **Will Facebook commit to making the outcomes of the civil rights audit, issues identified by the audit, and changes made as a result of the audit available to the public?**

As Laura Murphy noted in her December 2018 update, she will release another update on the civil rights audit this summer, as well as a final report upon completion of the audit.

2. **In your written testimony, you mentioned that Facebook is working to create an independent body so people could appeal decisions by content reviewers. You note that individuals could appeal not just for content that was taken down, but also for content that was reported and nonetheless left up.**

a. **What opportunities would Facebook’s proposed appeals process provide to individuals to voice their concerns about decisions by content reviewers?**

As our CEO Mark Zuckerberg has said, we do not think we should be alone in making decisions about online speech. That is why we seek input from external experts, academics, and representative groups when writing decisions about our content policies. We are also in the process of creating an external Board to review challenging content decisions. The Board will be a body of independent experts who will review content decisions, focusing on important and

disputed cases. It will provide details on each of its decisions and will be able to reverse our decisions about whether to allow or remove certain posts on the platform. Facebook will accept and implement the Board's decisions. The Board's decisions may also reveal policies that we need to reconsider or revise.

In January 2019, we released a draft charter that provides a suggested approach for the Board's structure, scope, and authority. The draft is a starting point for discussions that we have been having all over the world on how the Board should be structured and designed. The draft does not answer every proposed question and instead offers a suggested approach to how requests for review to the Board could be surfaced. For example:

- Questions will be referred to the Board by Facebook users who disagree with a decision, as well as by Facebook itself. Facebook will also refer content decisions to the Board for consideration when it considers specific cases that are especially difficult to resolve, when it finds that recurring issues have occasioned significant public debate and discussion or when existing policy and enforcement practices seem to lead to many decisions inconsistent with Facebook's values.
- Cases will be heard by panels formed from a rotating set of an odd number of members. Panels that have convened to decide cases could, at the conclusion of their session, choose a slate of eligible cases for subsequent panels to decide. A majority of that panel must agree to select a case.

Since Mark first proposed this idea in November, we have held one-on-one conversations, round tables, and workshops all around the world to engage experts and seek input on the draft charter. This outreach also included a public consultation process from anyone interested in providing their thoughts. This feedback will be released in a report at the end of June and will help to answer the questions that will be presented in a Final Charter.

- b. Would Facebook's proposed appeal process allow individuals to join together and file an appeal collectively, when a piece of content affects more than one user?**

To start, the Board will be hearing individual cases. We want to ensure that we are setting the Board up for success; it could be expanded in its scope as it becomes more mature.

- 3. In your written testimony, you mention that Facebook has created Community Standards outlining what content is permissible. You mention that your systems and human reviewers work in concert to identify and remove published content that violates these Community Standards.**

- a. How does Facebook decide which of the millions of reports it receives each week to address first?**

The people who use Facebook help us by reporting accounts or content that may violate our policies. Our content review teams around the world help review these reports 24 hours a day and in more than 50 languages. Over the course of 2018, we have more than doubled the number

of people working on safety and security issues to a total of 30,000, about half of whom are content reviewers.

We prioritize safety-related reports—for example, content posted by someone who may be in distress and is calling out for help.

**4. In your written testimony, you mentioned that Facebook has solicited external feedback on your content moderation policies from civil rights groups.**

**a. Which civil rights groups did Facebook consult with on its content moderation policies?**

We do not share the names of groups we consult with for a number of reasons—among them safety and security concerns, which are especially acute in places where the government may exercise censorship or control, and the fact that groups may not want to be named. That said, we typically engage with civil society organizations, activist groups, and thought leaders in areas including digital and civil rights, anti-discrimination, free speech, and human rights. We also engage with academics who have relevant expertise.

**b. What recommendations did these civil rights groups provide to Facebook?**

See Response to Question 1.

**c. What recommendations did these civil rights groups provide that Facebook did not implement, and why?**

See Response to Question 1.

**5. The Southern Poverty Law Center has raised concern about a specific list of designated hate groups that are active on Facebook. Last year, only 58 of over 200 of these designated hate groups had had their accounts suspended by Facebook. Recently, Facebook has announced a new policy banning not just white supremacist groups from your platform, but white nationalist groups as well, a broader category.**

**a. Will Facebook commit to monitoring these identified hate groups that are still allowed to operate on your platform for infringing content?**

Our policies against extremist content and organized hate groups are longstanding. Our Community Standards are clear that we do not allow hate groups to maintain a presence on Facebook. This is true regardless of the ideology espoused; we do not want Facebook to be a platform for hate. We have an extensive process that we follow in determining which organizations are designated as hate organizations and have worked with a number of different academics and organizations around the world to refine this process. We consider a number of different signals, among them organizations and their leaders that have called for or directly carried out violence against people based on things like race, ethnicity, and national origin.

Under this policy, we have banned more than 200 white supremacist groups from using our services. Last fall, we started using technology to more proactively identify hate groups globally, including white supremacists.

- 6. Tech companies occasionally remove voices that spread hatred, lies, and bigotry. Mr. Parker testified that people used the internet to “regurgitate demonstrably and undeniably false information about [the Sandy Hook shooting] while simultaneously attacking victims’ families for profit.” Those lies had real consequences.**
  - a. What criteria is used by Facebook to assess content that provokes or facilitates the online and offline harassment of crime victims?**
  - b. What staff and resources has Facebook made available to proactively monitor for and address the online and offline harassment of crime victims?**
  - c. What are the specific steps that crime victims such as Mr. Parker should take in order to elevate their cases to Facebook to receive specialized assistance?**

We do not tolerate harassment on Facebook because we want people to feel safe to engage and connect with their community. Our bullying and harassment policy prohibits targeting anyone maliciously by “posting content about a violent tragedy, or victims of violent tragedies that include claims that a violent tragedy did not occur.” It also prohibits posting content with claims that the victim or survivors are “acting [or] pretending to be a victim [or] otherwise paid or employed to mislead people about their role in the event.”

In the wake of violent tragedies, we have attempted to make ourselves available to the families of victims and survivors so that we can be responsive to their needs in a way that is easiest or best for them. In some cases, we set up direct channels for families to reach out to us; in others, we reach out to organizations and individuals who may want to report potentially violating content on the platform or use Facebook tools to fundraise or promote a cause.

- 7. I commend Facebook for providing increased information about enforcement of its content moderation policies. It is vital that Facebook continue to provide substantive information about hateful or abusive activities against individuals, including the demographics of people targeted, and rates of appeals. These data are important to researchers and civil society organizations trying to study the problems and target resources to affected communities including crime victims.**
  - a. Please provide specific information on the following:**
    - i. Of the content reported by individuals, what percentage of these reports are found to violate your rules?**

We are committed to making Facebook a place that is open and authentic, while safeguarding people’s private data and keeping our platform safe for everyone. We publish regular reports to give our community visibility into how we enforce policies, respond to data

requests, and protect intellectual property, while monitoring dynamics that limit access to Facebook products.

On May 23, 2019, we published our third Community Standards Enforcement Report. This report shows our enforcement efforts on our policies against adult nudity and sexual activity, fake accounts, hate speech, spam, terrorist propaganda, bullying and harassment, child nudity and sexual exploitation of children, regulated goods, and violence and graphic content, for the six months from October 2018 to March 2019. The report includes data on (1) the prevalence of Community Standards violations on Facebook; (2) how much content we took action on and (3) how much violating content we found before users reported it. And for the first time, we are also sharing data on our process for appealing and restoring content to correct mistakes in our enforcement decisions. For more information, please see: <https://transparency.facebook.com/community-standards-enforcement>.

**ii. What are the demographics of the users deemed engaging in hateful activities?**

See Response to Question 7.a.i.

**iii. What are the demographics of the users receiving hateful content?**

See Response to Question 7.a.i.

**iv. What are the demographics of the users reporting hateful activities?**

See Response to Question 7.a.i.

**v. What percentage of hateful activities are conducted by repeat offenders?**

See Response to Question 7.a.i.

**vi. What is the average review time of reported hateful activities?**

Reports are reviewed 24 hours a day, 7 days a week, and the vast majority of reports are reviewed within 24 hours. When there are credible threats of violence, we aim to respond much faster. To support these efforts, we are investing in people, technology, and programs. Specifically, we are building new tools so that we can more quickly and effectively detect abusive, hateful or false content. We have, for example, designated several hate figures and organizations for repeatedly violating our hate speech policies, which has led to the removal of accounts and content that support, praise or represent these individuals or organizations. We are also investing in artificial intelligence that will help us improve our understanding of dangerous content.

We have also more than tripled the number of people who work on safety and security at Facebook to 30,000, about half of whom are content reviewers.

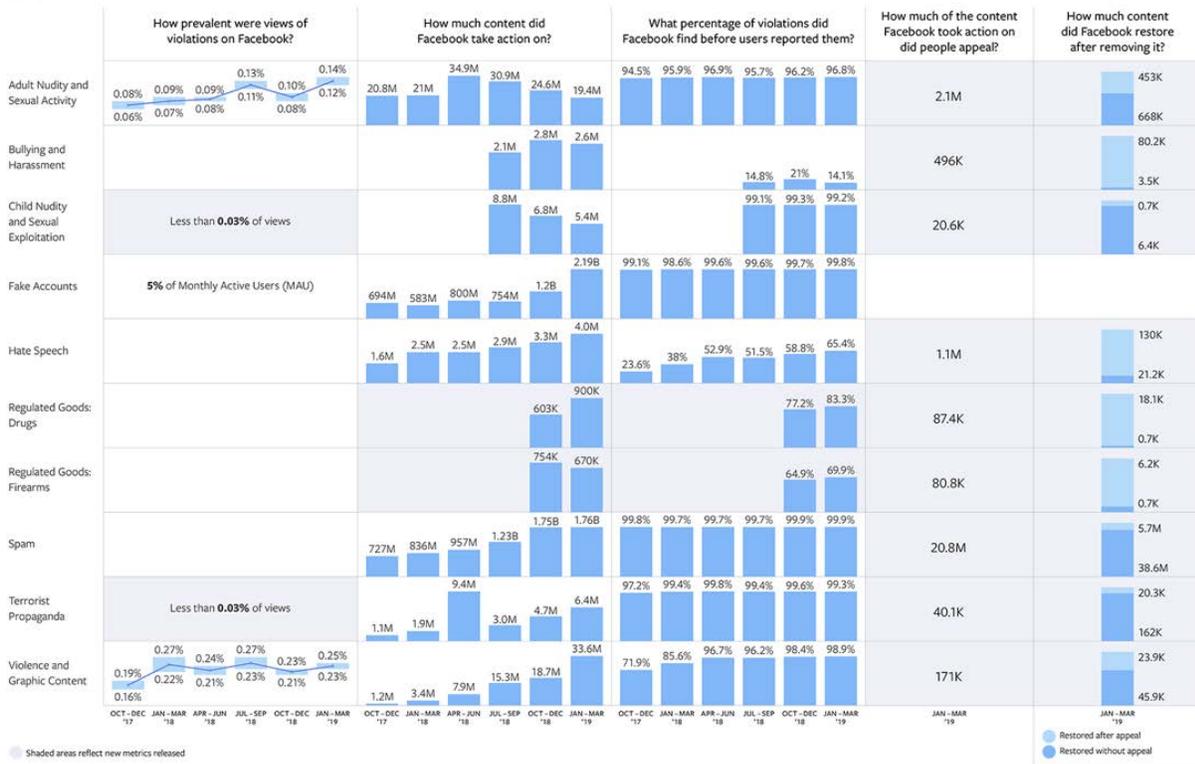
**vii. How many appeals of rules violations do you receive? What percentage of appeals are granted?**

In April 2018, we announced the launch of content-level appeals, and we have since made appeals available for all types of content that is removed from Facebook. We recognize that we make enforcement errors on both sides of the equation—what to allow and what to remove—and that our mistakes cause a great deal of concern for people, which is why we need to allow the option to request review of the decision and provide additional context that will help our team see the fuller picture as they review the post again. This type of feedback will allow us to continue improving our systems and processes so we can prevent similar mistakes in the future.

On May 23, 2019, we published our third Community Standards Enforcement Report, which for the first time includes data on our process for appealing and restoring content to correct mistakes in our enforcement decisions. We restore content when we know we have made a mistake in enforcement and do so in cases even without an appeal. For more information, please see: <https://transparency.facebook.com/community-standards-enforcement>.

**Data Snapshot: Facebook’s Community Standards Enforcement Report**

OCTOBER 2017 – MARCH 2019



Facebook is developing the metrics not shown here, and will share them as soon as meaningful and accurate measures and related data are available.

Source: Facebook's Community Standards Enforcement Report, May 2019

© 2019 Facebook, Inc.