Update on Facebook's Civil Rights Audit

I. A Note from Laura Murphy

The call for a civil rights audit at Facebook reflects the deep concerns of U.S.-based advocacy groups who have rightly observed that, as we bring much of our daily lives online through Facebook and other services, we should bring our civil rights with us. As such, when Facebook asked me to lead a civil rights audit of its platform earlier this year, I accepted.

My reasoning was simple. If I could help a platform that serves over 2 billion people improve its approach to civil rights and civil liberties, the impact would be significant and meaningful for Facebook's users and potentially the broader technology community and society.

But there is also no question that Facebook faces a number of serious civil rights challenges including allegations of allowing the propagation of hate speech and misinformation and the facilitation of voter suppression, discriminatory ad targeting, privacy lapses, unfair censorship, and algorithmic bias. To be clear, the civil right groups that are raising these concerns are pro-technology but firmly anti-bias; moreover, they are concerned about the impact of the platform on public discourse and the institutions that are the foundation of American democracy.

These concerns became particularly pronounced as the public learned more about the Russian government's attempts to use Facebook to influence the U.S. national elections in 2016, with a disproportionate focus on stirring racial and religious animus, and that Cambridge Analytica had collected the personal data of millions of Facebook users for political purposes. Recent reporting that Facebook hired a communications firm to attack the credibility of key civil rights stakeholders has only deepened these concerns.

As someone who has spent her career fighting to protect and advance civil rights – including during my 20-plus years working at the American Civil Liberties Union – I too take these concerns raised by the civil rights community very seriously.

Facebook products are of value to millions of people. Beyond promoting meaningful connections, Facebook's services can be – at their best – utilized as important and accessible tools for promoting civic engagement, movement building and organizing, exposing bias, and giving people a vehicle for expression.

This report details the process for the civil rights audit I undertook; outlines my team's work in partnership with Facebook over the last six months; and publicly shares progress that has been made, including policy changes, enforcement improvements, efforts to reduce voter suppression, and increased transparency in advertising. This progress represents a start, not a destination. While this work continues, Facebook's leadership

and I agreed that an interim update is important at this moment. Facebook has committed to issuing another public update in 2019 to maintain transparency on this ongoing work.

II.   Overview of Process and Stakeholder Engagement

The policy questions I have been asked to address are complicated. While I reject the false dichotomy of choosing free speech at the expense of other rights, I have gained appreciation for the challenge of developing standards for a diverse community of users who hold various and often competing viewpoints, while ensuring safety, equality, and non-discrimination. To be clear, many of these challenges are not unique to Facebook. But given its size, scale, and impact, Facebook must explicitly demonstrate a commitment to tackle these important questions; engage in thoughtful deliberation; and work in close collaboration with its partners in the civil rights and civil liberties community.

This last point is particularly important. The audit must be conducted in a coordinated way with meaningful input from the U.S. civil rights community. The United States is uniquely positioned with a robust civil society sector, including a number of organizations that have tremendous expertise and have decades-long commitments to protecting and advancing rights and liberties. These century-old organizations represent people of color, immigrants, religious minorities, women, senior citizens, and people with disabilities, and include key leaders from the African American, Latinx, Muslim, Jewish, Christian, Sikh, Arab American, Asian American, LGBTQ, and Native American communities and others, who represent millions of Facebook users.

Since this summer, my team and I have devoted hundreds of hours to interviewing leading advocates from these communities representing 90 different organizations to understand their concerns and identify key issues for inclusion in Facebook's civil rights audit. This direct engagement has provided valuable insight into the challenges Facebook faces, including important issues such as voter suppression, hate speech, and potential discriminatory ad targeting — issues that are complex and which merit careful deliberation. The interviews are ongoing.

I also held meetings with representatives from Facebook's Policy, Product, and Enforcement teams to gain a deeper understanding of Facebook's products and services. This work was an important complement to the legal engagement the company is undertaking with the prominent civil rights law firm, Relman, Dane & Colfax. I understand that work is still ongoing and information about it will be included in future updates.

This update is not meant to cover each and every issue we examined over the past six months. Rather, it is meant to provide a high-level overview of the policy work done to date and preview our future steps. **The purpose of the audit is to ensure Facebook makes real and lasting progress in the area of civil rights and civil liberties.**

Many of the advocates and experts with whom I spoke raised the following issues – in no particular order – as a priority for Facebook to consider. This is not an all-encompassing list, but rather a highlight of some of the major concerns:

- **Voter Suppression:** Protecting against misuse of the platform to intimidate or suppress voter participation, particularly amongst minority groups and people of color. Experts and advocates specifically noted the absence of a public-facing policy dedicated to preventing voter suppression and expressed concerns that Facebook's operations team was ill-prepared to address this challenge in advance of the 2018 midterm elections.

- **Accountability Infrastructure:** Instituting sufficient protocols and a civil rights infrastructure to ensure civil rights and potential discrimination 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.

- **Advertising Targeting and Practices**: Strengthening Facebook's non-discrimination in advertising policies and addressing allegations of facilitating discriminatory ad targeting against ethnic and religious minorities, immigrants, older workers, LGBTQ individuals, women, families with children, and other protected classes.

- **Diversity and Inclusion:** Promoting greater employee diversity in all functions and levels and taking more steps to create a more inclusive workplace. There is a belief that if Facebook has more diverse voices represented at the company, it will be less likely to overlook the civil rights implications of its products and policies.

- **Fairness in Artificial Intelligence and Algorithms:** Devoting resources to ensuring that artificial intelligence tools, such as machine learning or facial recognition, do not facilitate bias and are fair to all users, especially to members of protected classes.

- **Privacy:** Developing comprehensive privacy measures that protect civil rights while preventing unlawful discrimination are top of mind for advocates in the civil rights community.

- **Transparency:** Increasing transparency in Facebook's policy-making process, enforcement and operations.

III.   2018 Audit Focus: Facebook's U.S. Voting and Election Efforts

Of the issues civil rights leaders raised this summer, there was unique urgency around the 2018 midterm elections in the United States in November, specifically protecting against voter suppression and intimidation. As such, these concerns became a primary focus of the first phase of the civil rights audit. As Facebook's Chief Operating Officer Sheryl Sandberg stated in her congressional testimony on September 5, 2018: "There is a long history in this country of trying to suppress civil rights and voting rights and that activity has no place on Facebook."

Advocates were deeply alarmed at the impact of ads run by Russia's Internet Research Agency ("IRA") that surfaced on Facebook in 2016. They expressed concern that such ads encouraged racial and religious animus, attempted to suppress minority voter turnout, and particularly targeted communities of color and racial and ethnic minorities.

It is now clear that Facebook was slow to understand the IRA's activities, a point the company has acknowledged. Eventually, Facebook did take a number of steps to address concerns that the platform had been used to spread misinformation designed to influence elections, intimidate voters, and suppress voting. These steps included significant changes related to advertisers' posting of political and issue ads along with the creation of a searchable database, expanding the third-party fact-checking program, tightening restrictions on advertising, and removing fake accounts that were used to spread misinformation and inflammatory and divisive ads.

While these were important steps in the right direction, the civil rights community continued to express concern about the dangers of election-related misinformation spreading on Facebook. In particular, civil rights groups identified the possibility of false information regarding voter registration requirements or voting logistics. Before the rise of social media, flyers containing false voting information were posted in minority neighborhoods – on street corners or in churches, schools, and public parks. Today, the same falsehoods may spread online as memes in Facebook's News Feed, with the potential to reach millions of people in a short amount of time. Containing the spread of this type of content was and remains imperative for the company to address.

Groups also observed that advertisers might attempt to appeal to racial, ethnic, or religious prejudices and stereotypes in an effort to improperly suppress, intimidate, or influence voting by racial, ethnic, and religious minorities.

These concerns and discussions with experts and advocates generated a series of recommendations and best practices that informed the changes below.

Recommendations and Facebook Actions Ahead of 2018 Mid-Terms

The civil rights experts and advocates that I engaged over the summer shared specific recommendations related to preventing voter intimidation and suppression. Key themes emerged based on these conversations. Namely, that Facebook should:

- Employ robust approaches to preventing misinformation and voter suppression;

- Develop clear and public policies prohibiting voter suppression and intimidation, including voting-related misinformation and threats of violence;

- Increase transparency regarding the technical aspects of Facebook's election integrity efforts (e.g., how its policies are operationalized); and

- Seek guidance from voting rights experts to train Facebook staff on responding to voter suppression content on the platform.

These recommendations informed Facebook's efforts to strengthen its election-related efforts and processes. Given the complexity of the issues, implementation took longer than expected, but Facebook did make important changes prior to the U.S. midterm election, including but not limited to:

- **Strengthened Voter Suppression Policy**: Facebook strengthened its policy prohibiting voter suppression under its Community Standards.[1]

  The new policy expands on Facebook's earlier policy against misrepresenting voting dates, locations, and times, making clear that content designed to misrepresent information about voting logistics, including voting methods, qualifies as voter suppression too.

  The policy expressly bans 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. Other misinformation related to voting – including false claims of polling place closures, long lines, and wait times – is proactively sent to third-party fact-checkers for review. The revised policy also expressly prohibits threats of violence related to voting or voter registration.

  Finally, and as a direct response to feedback from civil rights advocates, Facebook recognized that it needed to focus specifically on voter suppression as a distinct civil rights challenge.

- **Supporting Voter Engagement**: Facebook employed increased voter registration and Election Day reminders.

- **Voting Expert Advisors**: Facebook newly retained voting experts as outside advisers to the company's policy and operations teams. These legal experts provided training to staff prior to the election and consulted with Facebook's policy and operations teams in the weeks leading up to and on Election Day.

---

[1] Facebook's Community Standards govern what content is and is not allowed on Facebook. The policy prohibiting voter suppression is included within the "Coordinating Harm" policy and can be found here: https://www.facebook.com/communitystandards/coordinating_harm.

- **User Reporting**: Facebook added a new option for users to specifically report incorrect voting information they found on Facebook during the U.S. midterm elections.

- **State Elections Officials Reporting**: Facebook established a specific reporting channel for state election authorities to report potential voter suppression content to the relevant policy and operations teams.

- **Voting Rights Partners Reporting**: Facebook partnered with voting rights and election protection groups to report potentially suppressive content for review and action.

- **Dedicated Reviewers for Timely Action**: To ensure action was taken in a timely manner, Facebook established dedicated channels through which potentially violating content involving voter suppression would be reviewed.

- **Strengthening Election "War Room"**: In advance of the election, Facebook established a physical war room to coordinate in real time the work of election-focused employees in the lead up to the midterm election.[2] The goal of the war room effort was to assemble voting experts from across the company in one place, so they could quickly address potential problems. Feedback from civil rights groups significantly improved the war room's ability to handle voter suppression and intimidation related content. At their recommendation, subject-matter experts in voting rights law were dedicated to advising the war room on a 24/7 basis. Training materials on Facebook's new policy were distributed to war room personnel and specific points-of-contact were designated to identify and remove violating content proactively.

- **Cracking down on fake accounts.** Facebook continues to take an aggressive approach to removing fake accounts, including from foreign actors, which are one of the largest sources of misinformation.

Immediately prior to the 2018 midterm elections, I encouraged Facebook to host representatives from several civil rights organizations engaged in the civil rights audit to visit the election war room and see first-hand the company's elections operations. Facebook's executive leadership met with these voting rights advocates and experts to discuss voter suppression and election protection more broadly. During that meeting, Facebook leadership acknowledged that the work with civil rights groups to prepare for the midterm elections had a significant impact on Facebook's policies and process and represented a meaningful step toward greater collaboration.

While these were important steps, there is more work to do on voting and election-related issues. I will continue to examine these issues, including by working with Facebook to

---

[2] Fighting Election Interference in Real Time, Facebook Newsroom, Oct. 18, 2018, available at https://newsroom.fb.com/news/2018/10/war-room/.

incorporate lessons learned from the 2018 election cycle. Those changes will be discussed in future public updates.

IV. Key Changes at Facebook in 2018

In addition to the improvements discussed above, Facebook has also made a number of changes this year that address priorities that my team heard from the civil rights community. Many of these changes evolved on a separate track from the civil rights audit. These changes include:

Transparency

- **Community Standards**: In April, Facebook released a more detailed version of its Community Standards, including what had previously been internal enforcement guidelines.[3] In tandem, Facebook made available content-level appeals, which are designed to enable users to challenge content decisions at the post level, and which responds to a critical and consistent concern voiced throughout my discussions.

- **Content Enforcement Report[4]**: In May, Facebook published a report on how much content it had removed for violating its policies against adult nudity and sexual activity, fake accounts, hate speech, spam, terrorist propaganda, and violence and graphic content. In November, the company published its second enforcement report, which also included two new categories of data: bullying and harassment, and child nudity and sexual exploitation of children. Facebook will continue to publish these reports on a biannual basis.

Ads

- **Ads Transparency for Users**: In June, Facebook released a feature that allows users to view an advertiser's currently running ad campaigns. This will further the company's 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. Nonetheless, civil rights groups have urged for greater transparency measures to be incorporated into this tool and I look forward to working with Facebook to improve its design, accessibility, and functionality.

- **Removal of Thousands of Targeting Terms**: Since last year, Facebook has removed thousands of ad targeting options for ads related to housing, employment, credit, insurance, and public accommodations, if the targeting options could be misunderstood as describing groups on the basis of race, creed, color, national origin, veteran or military status, sexual orientation, and disability status. After further review of targeting options under these standards, in August,

---

[3] https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/
[4] https://newsroom.fb.com/news/2018/05/enforcement-numbers/

Facebook removed 5,000 additional targeting options to minimize potential for advertiser abuse. This is an important area of work where much more progress needs to be made.

- **New, Mandatory Non-Discrimination Certification**: This year, Facebook 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 U.S. advertisers placing any ad on Facebook.

Commerce Tools for Users

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

Artificial Intelligence and Algorithmic Fairness

- **Resources and Outreach on Fairness for Artificial Intelligence Tools**: While algorithms and artificial intelligence offer an immense opportunity to generate value for people and communities, the technology industry is facing significant questions about ensuring fairness in these tools. The exploration of these complex, industry-wide questions is at the very beginning with no simple solutions. At this early stage, it is important that industry and external stakeholders engage with each other to articulate the right questions and learn together.

  This year, Facebook has put more resources behind exploring algorithmic and AI fairness issues by convening a multi-disciplinary team to study this issue and engaging with a range of external stakeholders, including those representing ethnic and religious minorities and communities of color, and experts who work on algorithms and AI. Facebook has also participated in a number of conferences, panels, roundtables, and workshops that have been organized by nonprofit, academic and research institutions in the United States and abroad.

  This is an area of rapid change and new study, where ongoing input from the civil rights community is essential. I will continue to work with Facebook to expand such engagement and collaboration.

V.   Looking to 2019: Road Ahead

While the civil rights audit has achieved some important results, more work remains. I look forward to continuing this effort next year with Facebook and the civil rights community. The first two areas of focus for 2019 are Content Moderation and developing a Civil Rights Accountability Infrastructure, to build upon the work we have begun.

- **Content Moderation**: Facebook has stated repeatedly that its policies are iterative and subject to continuous improvement. CEO Mark Zuckerberg recently confirmed in a public statement that content moderation and enforcement is, and will continue to be, a significant area of focus for the company.[5] Taking in feedback from civil rights experts, I will explore various approaches to content moderation as well as measures for addressing censorship and harmful and potentially discriminating content on the platform.

- **Creating a "Civil Rights Accountability Infrastructure"**: A chief concern expressed from the civil rights community is ensuring that Facebook builds an appropriate infrastructure to ensure that the changes it is making are systemic – and that civil rights issues are considered "on the front end" as Facebook develops new products, features, and policies. I plan to discuss the contours and characteristics of such an infrastructure with both civil rights experts, advocates and company officials in the coming months.

For the past several years, civil rights groups have consistently expressed, both publicly and privately, their deeply held concerns about Facebook's products, policies, and practices and their implications on civil and human rights. The work that has been done over the last six months is an attempt to capture and consolidate their concerns to produce meaningful results. Given Facebook's scope and scale, this continues to be a challenge. That being said, in the first six months of this audit, we have witnessed some progress and tangible results, including policy changes, improvements to enforcement, and greater transparency in certain areas. Importantly, Facebook has sought to deepen its engagement with the civil rights community through this process, and I want to thank my colleagues in the community for their advice, insight, and time, which directly led to these improvements. I fully realize that these leaders expect much more, and I am committed to working with them to help achieve the ambitious goals we have set, and which have the potential to positively impact Facebook's global community. I look forward to issuing another update in 2019.

---

[5] Mark Zuckerberg, "A Blueprint for Content Governance and Enforcement," Nov. 15, 2018, *blogpost available at* https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/.