Center for American Progress

Combating Hate and White Nationalism in the Digital World

By Aastha Uprety and Danyelle Solomon   August 8, 2018

While the digital world has bolstered the free exchange of ideas and revolutionized the global economy, it also provides new fertile ground for old evils. The spread of violent white nationalist ideology online, rampant algorithmic bias, and homogeneity in the technology industry threaten to undo decades of racial progress and further entrench inequality.

Historically, racial progress in the United States has almost inevitably been followed by racist pushback. The Civil War, emancipation, and Reconstruction were followed by more than 80 years of state-sanctioned violence against African American communities. The gains coming out of the Civil Rights era of the 1960s were immediately met with white backlash and resentment that seeped into the nation's politics. Policies targeted communities of color with draconian actions, not the least being the weaponization of the criminal justice system. More recently, the historic election of President Barack Obama was followed by a resurgence of violent white nationalism—the alt-right movement[1]—which was key to the election of Donald Trump.

This perplexing pendulum of racial progress followed by a rise in racism helps explain the recent resurgence of white nationalism in the United States. This maddening cycle owes itself to Americans' addiction to collective denial and selective ignorance when it comes to the nation's history.[2]

A progressive digital agenda that promotes inclusivity and diversity is needed to address this latest swing and mitigate hate and violence. Without proper oversight, regulation, and accountability to combat the ugliest online realms, society will never reap the full benefits of the digital world. To that end, this issue brief offers some concrete steps policymakers and industry leaders can take to fight racism online and within the industry.

Monitor and mitigate hate online

In 2017, hundreds of white nationalists, using the online chat and voice app Discord, gathered to organize and plan the so-called Unite the Right rally in Charlottesville, Virginia—a violent protest that used the removal of a statue of Confederate Gen. Robert E. Lee as a pretense. On August 11, the torch-wielding mob marched through the city chanting the Nazi slogan "blood and soil" and beat students on the University of Virginia campus. The following day, the mob assaulted counter protesters with mace and clubs. One member rammed his car into a crowd of civilians, killing one woman and injuring dozens more.[3]

The attacks in Charlottesville were a wake-up call revealing the digital world's potential for incubating violent ideologies and inspiring domestic terrorists. In reality, hate groups have long relied on online technology to advance their agendas; the digital world provides channels through which hate groups can operate anonymously and communicate without detection. Their ability to reach narrow audiences across a large geography allows these groups to raise money, spread racist propaganda, and lure and indoctrinate new followers.

Halting the spread of violent white nationalism online will require nonprofits and private sector companies in the media and technology industries to devote substantial resources toward research and the development of best practices to address this threat. This research should seek to fully identify and understand patterns of behavior; platforms for communication and indoctrination; and mechanisms for exchanging money and weapons.

In the meantime, media and technology companies should begin implementing clear terms-of-use policies, expand enforcement mechanisms, and put in place measures to ensure transparency and accountability.

Most platforms have terms of use, which lay out the rules users must agree to and abide by to use the service.[4] As private companies, online platforms have the right to regulate content on their platform. Companies should monitor hate speech and, if appropriate, act when content violates their terms of use policy. In order to be sufficient, terms of use should state that:

- Users may not engage in promoting hate or violence based on race, color, religion, gender, sexual orientation, gender identity, or disability and that users who do so are subject to suspension or termination.

- Content that fits the above category is subject to removal. A thorough understanding by content reviewers of social, cultural, and political norms can help determine what content warrants removal.

- Users may appeal if they believe they were wrongly affected by the policy.

Expanding enforcement mechanisms should involve employing artificial intelligence, user flags, and well-trained human content reviewers to help to identify and remove hateful content. Some companies have responded to hateful content on their platforms with appropriate actions. For example, following the rally in Charlottesville, PayPal disabled some key right-wing groups from receiving donations on their platform.[5] In addition, Reddit deleted a few of its most hateful subreddits in 2015, which led to a nearly 90 percent decrease in hate speech by users who once frequented those forums.[6]

Along with clearly defined terms of use and robust enforcement mechanisms, companies must commit to transparency and accountability. To protect freedom of expression and ensure that content reviewers take relevant context into account, clear pathways should be in place for users who are denied services due to terms-of-use violations to appeal such decisions. Companies should disclose the methods through which reviewers determine which content to remove and if a user's actions warrant suspension or termination. In addition, companies should regularly release data on how many users have been denied services, why the users were removed, how many appealed, and how many appealed successfully. Facebook, for example, released its Community Standards Enforcement Preliminary Report, which provides some data on community standards violations and subsequent content removal.[7] YouTube released a similar report on enforcement of their community guidelines.[8] Public access to such information is helpful to researchers studying the impact of content monitoring and removal on curbing online hate.

Lack of oversight on social media platforms helped facilitate Russia's interference in the 2016 elections. Russian-purchased fake ads targeted users based on their ethnicity, interests, and prior internet histories, and attempted to show discontent and suppress voter participation among key groups such as African Americans.[9] Some ads also targeted conservative voters by using anti-Muslim and anti-immigrant language.[10] In fact, more than half of Russian ads on Facebook used race as a central theme.[11] This instance was a chilling attack on U.S. democracy and also served as an illustration of how content that fuels racial resentment can dominate social media platforms. In May, Facebook committed to undergoing a civil rights audit,[12] partly due to a letter from a group of civil rights organizations in response to both Russian ads and hate speech on the platform.[13] It has also removed fake Russian pages ahead of the midterm elections.[14]

The digital world's incubation of violent white nationalist groups has allowed them to spread their influence and organize domestic terror attacks in communities of color. The same online tools used by foreign actors to undermine the U.S. democracy further demonstrate how vulnerable these platforms can be to misuse. Nonprofits and companies that provide online services to users must demonstrate their commitment to combating extremism—while protecting free expression—through research, clearly defined terms of use, effective enforcement, and transparency.

Remove online racial bias in all its forms

Racial bias is the implicit or explicit prejudice against racial minorities that results from predominant, often harmful, societal stereotypes about certain racial groups.[15] In the digital world, racial bias often manifests itself in algorithms used to design social media platforms and search engines.[16] Algorithms, the functions used by computer programs to make automated decisions, can reflect both the biases of their human programmers as well as the biases in the data these algorithms use to make decisions.[17] When algorithms produce racially biased outputs on online platforms, companies can discriminate—intentionally and unintentionally—against users of color or promote propaganda that reinforces false narratives and that plays to negative stereotypes. This is not only wrong from an ethical perspective; algorithmic bias also damages the economy by excluding consumers and validates extremist views. Companies must do more to combat racial bias in online algorithms.

Some social media platforms drive content to a user by relying on algorithmic predictions of what the user wants to see based on prior interests, which can sometimes result in intentional or inadvertent discrimination. Targeted ads for credit cards, for example, raise questions about who could be on the receiving end of predatory marketing, especially if race is a distinguishing factor.[18] Some vendors will price discriminate, or offer different prices to different potential consumers based on an internet user's ZIP code. While common, this practice can result in racial discrimination by geography.[19]

While targeting specific audiences is part of the appeal and efficiency of online marketing, advertisers and online platforms must be mindful of practices that may cause harm or unfairly target racial minorities. For example, researchers found that Google posted ads for criminal background checks, as well as credit cards with exorbitant fees and high interest rates on an African American fraternity's website.[20] Recently, Facebook has also come under fire for allowing advertisers to target specific audiences by categories including "Ethnic Affinity," which enabled sellers to basically racially discriminate by excluding users associated with certain characteristics from seeing their ad.[21] For example, sellers could exclude user traits such as "African Americans" and "Spanish speakers" from their target audience.[22] After this practice was uncovered, the National Fair Housing Alliance sued Facebook,[23] claiming that the social media company was violating the Fair Housing Act. Online platforms such as Facebook should commit to updating the algorithms and guidelines used to detect whether an ad is discriminatory. Facebook should be able to catch and prevent not only blatant discrimination–such as preventing a user whose interests include the topic "African American" from seeing a specific ad—but also more discrete discrimination, such as housing redlining via limiting audiences to certain ZIP codes.

Search engine algorithms can also reflect racial bias. One study found that entering black-identifying names into Google displayed ads suggestive of the person having an arrest record, a phenomenon that did not occur as frequently for white-identifying

names.[24] For example, a search for the first names "Latanya" and "Latisha" displayed ads for a background check service, but a search of the names "Kristen" and "Jill" rendered neutral results.[25] Unfortunately, platforms have attempted to justify these instances as objective depictions of online content, but platforms need to take seriously the influence they have over public opinion.[26] Hate groups can take advantage of search engine algorithms to target individuals and elevate their websites in search results, which is only exacerbated when platforms fail to monitor and minimize hate content online. For example, Dylann Roof, who murdered nine African Americans at a church in Charleston, South Carolina, pointed to online searches as the beginning of his descent into the white supremacist online world of hate.[27] Roof said that when he typed "black on White crime" into Google, the search results were populated with propaganda declaring the prevalence of black people assaulting and killing white people, which in part fueled his racial animus.[28]

Technology companies must take steps to prevent racial bias on their platforms by ensuring algorithms do not inadvertently exclude marginalized groups from the digital marketplace or promote violent extremist propaganda.

Implement effective diversity and inclusion policies

Another manifestation of structural racism in the digital world is the lack of diversity within the most powerful technology and social media companies in the world. According to data from 2016, the technology workforce in Silicon Valley was only 2.2 percent black and 4.7 percent Hispanic.[29] This persistent lack of inclusion is a problem that costs companies billions, reduces product quality, and limits the development of innovative solutions to combat online hate and prevent algorithmic bias. To address this issue, the industry must devote significant resources toward recruiting candidates with diverse backgrounds and building safe and inclusive workplace environments.

One effective method of recruiting and hiring more diverse candidates is to target multicultural professional associations at colleges and universities. Another option is the diverse-slate hiring approach, which short lists more than one candidate from underrepresented minority groups and can increase the likelihood of hiring a minority.[30] Diverse hiring should be a priority at all levels of employment, from engineers to executives. Recently, Facebook committed to increasing the diversity of its board members,[31] and Google has pledged to focus their efforts on hiring black and Hispanic women.[32]

However, hiring is not the only concern: Many high-tech companies have trouble retaining employees from underrepresented minority groups.[33] A 2017 study found that 37 percent of people who left their jobs in a technology industry or function listed unfair treatment as a major factor in their decision to leave, and this trend was most prevalent

among underrepresented people of color. This kind of turnover costs companies an estimated $16 billion per year in employee replacement.[34] Additionally, women of color who work in technology roles face unique challenges, including being less likely to be in managerial positions[35] and being paid less than their male counterparts.[36]

To solve the problem of retention, companies must commit to fostering inclusive workplace cultures through methods that go beyond diversity training, such as establishing a diversity executive or council to continually lead inclusion efforts[37] and creating employee resource groups for different minority employees.[38] Companies must also commit to comprehensive and transparent methods for reducing the pay inequity between white workers and underrepresented populations, especially women of color.

The rise of hate online should be another reminder that technology companies must open their doors to underrepresented minorities and leverage those varied perspectives to find new solutions for the greater good. A focus on diversity in recruitment and workplace retention is a good place to start.

Conclusion

The internet is not immune to America's sordid legacy and, like most of America's institutions, is equally as porous to racism. While most Americans use the internet to communicate with family and friends or to conduct business and pay bills, there are some that use these platforms for harm. As this nation continues to increase its reliance on technology, it is imperative that technology industries and policymakers work together to develop and implement strategies and policies that mitigate racism, hate, and extremism online.

Aastha Uprety is a fellow for Race and Ethnicity at the Center for American Progress. Danyelle Solomon is the senior director of Race and Ethnicity at the Center.

Endnotes

1 Hope Not Hate, "The International Alternative Right," available at https://alternativeright.hopenothate.com/?intro=0 (last accessed August 2018).

2 Danyelle Solomon, "Suppression: A Common Thread in American Democracy," June 16, 2017, available at https://www.americanprogress.org/issues/race/news/2017/06/16/434561/suppression-common-thread-american-democracy/.

3 German Lopez, "Charlottesville protests: a quick guide to the violent clashes this weekend," Vox, August 14, 2017, available at https://www.vox.com/identities/2017/8/14/16143168/charlottesville-va-protests.

4 Twitter, "Hateful conduct policy," available at https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy (last accessed July 2018); Facebook, "Community Standards," available at https://www.facebook.com/communitystandards/introduction/ (last accessed July 2018).

5 Jonathan Berr, "PayPal cuts off payments to right-wing extremists," CBS News, August 16, 2017, available at https://www.cbsnews.com/news/paypal-suspends-dozens-of-racist-groups-sites-altright-com/.

6 Devin Coldewey, "Study finds Reddit's controversial ban of its most toxic subreddits actually worked," TechCrunch, September 11, 2017, available at https://techcrunch.com/2017/09/11/study-finds-reddits-controversial-ban-of-its-most-toxic-subreddits-actually-worked/.

7 Facebook, "Community Standards Enforcement Preliminary Report" (2018), available at https://transparency.facebook.com/community-standards-enforcement.

8 Google, "YouTube Community Guidelines enforcement," available at https://transparencyreport.google.com/youtube-policy/overview (last accessed July 2018).

9 The Moscow Project, "Suppressing the Vote with Facebook Ads," September 28, 2017, available at https://themoscowproject.org/dispatch/suppressing-vote-facebook-ads/.

10 Craig Timberg and others, "Russian ads, now publicly released, show sophistication of influence campaign," The Washington Post, November 1, 2017, available at https://www.washingtonpost.com/business/technology/russian-ads-now-publicly-released-show-sophistication-of-influence-campaign/2017/11/01/d26aead2-bf1b-11e7-8444-a0d4f04b89eb_story.html.

11 Chas Danner, "More Than Half of Russian Ads Focused on Race," New York Magazine, May 12, 2018, available at http://nymag.com/daily/intelligencer/2018/05/more-than-half-of-russian-facebook-ads-focused-on-race.html.

12 Sameer Rao, "Advocacy Organizations Are Cautiously Optimistic About Facebook's Civil Rights Audit," Colorlines, May 3, 2018, available at https://www.colorlines.com/articles/advocacy-organizations-are-cautiously-optimistic-about-facebooks-civil-rights-audit.

13 Scott Simpson, "Civil Rights Groups Urge Facebook to Address Longstanding Issues with Hate Speech and Bigotry," Press release, Muslim Advocates, October 31, 2017, available at https://www.muslimadvocates.org/19civilrightsgroupslettertofacebook/.

14 Donie O'Sullivan and Jeremy Herb, "Facebook takes down suspected Russian network of pages," CNN, July 31, 2018, available at https://money.cnn.com/2018/07/31/technology/facebook-removes-pages/index.html.

15 American Psychological Association, "Speaking of Psychology: Understanding your racial biases," available at http://www.apa.org/research/action/speaking-of-psychology/understanding-biases.aspx (last accessed July 2018).

16 Nicol Turner-Lee, "Addressing racial bias in the online economy" Brookings Institution, December 1, 2016, available at https://www.brookings.edu/blog/techtank/2016/12/01/addressing-racial-bias-in-the-online-economy/.

17 Laura Hudson, "Technology Is Biased Too. How Do We Fix It?", FiveThirtyEight, July 20, 2017, available at https://fivethirtyeight.com/features/technology-is-biased-too-how-do-we-fix-it/.

18 Latanya Sweeney, "Online ads roll the dice," U.S. Federal Trade Commission, September 25, 2014, available at https://www.ftc.gov/news-events/blogs/techftc/2014/09/online-ads-roll-dice.

19 Keyon Vafa and others, "Price Discrimination in The Princeton Review's Online SAT Tutoring Service," Technology Science (2015), available at https://techscience.org/a/2015090102/.

20 Sweeney, "Online ads roll the dice."

21 Julia Angwin, Ariana Tobin, and Madeleine Varner, "Facebook (Still) Letting Housing Advertisers Exclude Users by Race," ProPublica, November 21, 2017, available at https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin.

22 Ibid.

23 Tanvi Misra, "Facebook Is Being Sued for Housing Discrimination, Too," CityLab, March 27, 2018, available at https://www.citylab.com/equity/2018/03/facebook-is-being-sued-for-housing-discrimination-too/556580/.

24 Latanya Sweeney, "Discrimination in Online Ad Delivery" (Cambridge, MA: Harvard University, 2013), available at https://dataprivacylab.org/projects/onlineads/1071-1.pdf.

25 Ibid.

26 Luke Dormehl, "Can an algorithm be racist? Spotting systemic oppression in the age of Google," Digital Trends, March 3, 2018, available at https://www.digitaltrends.com/cool-tech/algorithms-of-oppression-racist/.

27 James McWilliams, "Dylann Roof's Fateful Google Search," Pacific Standard, July 2, 2018, available at https://psmag.com/news/dylann-roof-google-algorithms.

28 Ibid.

29 Maya Beasley, "There Is a Supply of Diverse Workers in Tech, So Why Is Silicon Valley So Lacking in Diversity?" (Washington: Center for American Progress, 2017), available at https://www.americanprogress.org/issues/race/reports/2017/03/29/429424/supply-diverse-workers-tech-silicon-valley-lacking-diversity/.

30 Stefanie K. Johnson, David R. Hekman, and Elsa T. Chan, "If There's Only One Woman in Your Candidate Pool, There's Statistically No Chance She'll Be Hired," Harvard Business Review, April 26, 2016, available at https://hbr.org/2016/04/if-theres-only-one-woman-in-your-candidate-pool-theres-statistically-no-chance-shell-be-hired.

31 Sara Ashley O'Brien, "Facebook commits to seeking more minority directors," CNN, May 31, 2018, available at http://money.cnn.com/2018/05/31/technology/facebook-board-diversity/index.html.

32 Jessica Guynn, "Google says it will focus diversity efforts on black, Hispanic women," USA Today, June 14, 2018, available at https://www.usatoday.com/story/tech/2018/06/14/google-says-focus-diversity-efforts-black-hispanic-women/703003002/.

33 Hamza Shaban, "Google diversity report: Black women make up only 1.2 percent of its U.S. workforce," The Washington Post, June 15, 2018, available at https://www.washingtonpost.com/news/the-switch/wp/2018/06/15/google-diversity-report-black-women-make-up-only-1-2-percent-of-its-u-s-workforce/?utm_term=.6398f9032786.

34 Allison Scott, Freada Kapor Klein, and Uriridiakoghene Onovakpuri, "Tech Leavers Study: A first-of-its-kind analysis of why people voluntarily left jobs in tech" (Oakland, CA: Kapor Center for Social Impact and New York: Ford Foundation, 2017), available at https://www.kaporcenter.org/wp-content/uploads/2017/08/TechLeavers2017.pdf.

35 Center for Employment Equity, "Is Silicon Valley Tech Diversity Possible Now?" (2018), available at https://static1. squarespace.com/static/5b205de896e76fd6f1b8fdd3/t/5b2 fb2a9758d46fa66f52473/1529852589305/2018-Diversity+in +Silicon+Valley+Tech+June+25+release+version.pdf.

36 Hired, "The State of Wage Inequality in the Workplace" (2018) https://hired.com/wage-inequality-report.

37 Beasley, "There Is a Supply of Diverse Workers in Tech, So Why Is Silicon Valley So Lacking in Diversity?"

38 Scott, Klein, and Onovakpuri, "Tech Leavers Study."