HORIZON FORUM

PHILANTHROPY | EQUITY | DIALOGUE

WHAT DO THE EXPERTS SAY:

The Free Speech, Hate Speech Debate

February 1, 2021

Annotated Bibliography #1
Prepared by: Nagham El Karhili

HORIZON FORUM

Horizon Forum is a fiscally sponsored project of the Proteus Fund, a 501(c)(3). The initiative aims to strategically engage with sector leaders in philanthropy to ensure their charitable missions are insulated from extremism, bigotry, and hate activity.

Horizon Forum's primary activity consists of hosting semi-annual dialogues with stakeholders and researchers familiar with the institutional landscape of philanthropy and unique challenges of self-regulation therein. It also partners with peer institutions and university-based research centers to provide public facing thought leadership and data-based recommendations for stakeholders. Finally, the program also provides direct services for foundations, assisting them in establishing processes and policies for vetting grantees and ensuring they protect their mission to nurture the public good and "do no harm" with their philanthropy.

WHAT DO THE EXPERTS SAY

"What Do the Experts Say" is an ongoing series that reviews recent discussions affecting the philanthropic sector by thought leaders and experts from a variety of sources in the form of an annotated bibliography. It captures the state of the debate today, helping readers navigate an otherwise complex and multidimensional field.

CONTENTS

INTRODUCTION

As current levels of social and political polarization reach new heights, so do discussions about extremism, disinformation, and hate speech. But what do these terms mean for institutions such as philanthropy, technology, and media which straddle and blur the boundaries between public and private life? In many ways, current debates follow old patterns and paradigms surrounding the limits of free speech. As a constitutionally protected right, free speech protections spur discussion across the spectrum of constitutional rights, public safety, and political activity. As a result, conversations around hate speech are often fraught and highly contested, leaving the United States as one of the few democracies in the world to not have some form of legislation prohibiting its amplification and proliferation.

Although broad consensus on what constitutes hate speech is unlikely to be reached in the near future, leading voices from the law, nonprofit, and academic sectors are actively shaping the current state of the conversation.

The legal field has contributed to shaping the public understanding of free speech and hate speech in theoretical and practical ways. Legal scholars, political philosophers, and ethicists explore the definition through in-depth analyses of hate speech sub-categories ranging from racist speech, to incitement, and violence. While disagreements continue to fuel debates for clearer categories and appropriate legal repercussions, theoretical legal inquiries focus on drawing a link between the nature of the speech and ways in which it can directly lead to violence. Politically speaking, this debate often falls along predictable partisan and ideological lines.

More recently, the tech industry has played an increasingly active role in shaping the practical definition of hate speech. Operating with very little government regulation, social media platforms have unprecedented control over content that is capable of reaching broad swaths of the public. Tech companies such as Google, Apple, and Facebook, have taken aggressive action to moderate hateful content on their platforms through their terms of service policies. They also utilize content moderation tools such as profile removals, app de-platforming, and stricter user control. The use of such tools can often lead to claims of censorship or bias.

Additionally, academic voices have weighed in on the issue through philosophical interrogations and sociological studies. On the one hand, ethical questions are posed around the moral right to freedom of expression and the direct causation between speech

and violence. On the other hand, sociological approaches center the experiences of victims of hate speech. By prioritizing the human and social impact, this inter-disciplinary approach presents a more granular and practical understanding of hate speech, along with providing opportunities available to manage the problem in the medium and long terms.

Contained is a review of recent discussions by thought leaders and experts from a variety of sources that capture the state of the debate today, helping readers navigate an otherwise complex and multidimensional field.

LEGAL CONVERSATIONS

In the News:

Edsall, Thomas. "Have Trump's Lies Wrecked Free Speech?" *The New York Times.* January 6, 2021.

> In light of the Trump presidency, public debates have arisen around the limits of the constitutional protection of the First Amendment. Thomas Edsall, a political journalism professor at Columbia University, reviews the arguments of key thoughts leaders in the national debate exploring the challenge that unbridled free speech poses for American democracy. The article highlights the lack of consensus amongst scholars and specialists, while emphasizing the centrality of this conversation in our current political moment.

Stengel, Richard. "Why America Needs a Hate Speech Law." *The Washington Post.* November 7, 2019.

> Richard Stengel, a former public diplomacy official in the Obama administration, argues that regulations around speech need to be updated and modernized to fit new technological and political realities. The author contextualizes the U.S. standard of free speech as a global outlier with outdated laws. While noting the exceptional nature of America's focus on free speech as an essential component of its democracy, the author also notes the challenges facing Truth in today's marketplace of ideas.

Turley, Jonathan. "No, the U.S. Does Not Need European-Style Hate Speech Laws." *USA Today.* November 8, 2019.

> Responding to, and rejecting, calls for hate speech legislation, Jonathan Turley points to democratic thinkers, and policy makers as the main culprits in the war against free speech. By centralizing free speech as one of the most foundational freedoms in American democracy, the author presents any revisions or restrictions as a political assault led by left-wing politicians. To support his argument, Turley points to various European statistics of free speech criminalization. In the American context, the article highlights democrats' use of corporations as a tool to limit free speech.

Maccandless Farmer, Britt, "[Why the ACLU Defends White Nationalists' Free Speech](#)." *CBS News.* March 10, 2019.

In her article, investigative journalist Britt Maccandless Farmer comments on the American Civil Liberties Union's complex commitment to free speech. Overall, the article recounts the ACLU's history of defending various groups' right to free speech, including white nationalist groups. Specifically, Mccandless Farmer highlights the ACLU's controversial decision to defend Neo Nazi groups' right to march in Charlottesville in 2018. Written a year after the violent march, the article notes that the civil rights advocacy group recently added new criteria allowing the group to vet the potential for violence before accepting free speech cases.

Nonprofit and NGO Reports:

Guterres, A. "[United Nations Strategy and Plan of Action on Hate Speech.](#)" *United Nations.* 2019.

Recognizing the rise in speech and acts of intolerance across the globe, the United Nations outlines their plan forward through this Strategy and Plan of Action on Hate Speech. Led by the United Nations Secretary-General António Guterres, the document elaborates on the definition of hate speech, the UN's main objectives in addressing it, along with ways to respond to its impact on most vulnerable societies.

DiPietro, Becca. "[There's a World of Difference Between Free Speech and Hate Speech](#)." *Center for American Progress.* April 21, 2017.

This piece from the Center for American Progress highlights the complex dynamics of free speech on university campuses in the context of a 2017 event at Georgetown University. The Georgetown University College Republicans invited Nonie Darwish, a prominent anti-Islam writer, to promote her recent book. Protestors demanded that Darwish not be given a platform due to her calls for the annihilation of Islam. DiPietro's article affirms the importance of universities protecting access to free speech regardless of politics, but draws distinctions between university sponsored events and an individual's right to free expression. She urges that the roots of bias must be acknowledged. Most importantly, she calls for critics of the protestors to recognize the students position and desire for safe and welcoming environments.

American Civil Liberties Center. "[Freedom of Expression](#)." 2020.

> This is the official ACLU position on the free speech debate. This position paper establishes the ACLU as a core organization in fighting to protect free speech, even in the case of white supremacists. The ACLU maps the legal history of freedom of speech in court cases such as Abrahms v U.S. (two dissenting opinions establish the "clear and present danger" test), Brandenberg v. Ohio (speech is protected unless it means to produce "imminent lawless action"), or Texas v. Johnson and U.S. v. Eichman (the protection of symbolic speech). Free speech does not extend to the legally obscene (Miller v. California), defamation (New York Times Co. v. Sullivan), or "fighting words" (Chaplinsky v. New Hampshire). In their defense of the rights of white supremacist and other hate groups speech, the ACLU maintains that the First Amendment cannot only protect popular speech, but must be extended to distasteful speech as well.

Academic Sources:

Fino, Audrey. "[Defining Hate Speech: A Seemingly Elusive Task](#)." *Journal of International Criminal Justice.* 2020. 18-1.

> This article reviews the status of international criminal law on hate speech. As a baseline, Audrey Fino uses the most egregious form of hate speech that has been prosecuted as an international crime — that of direct and public incitement to genocide. Specifically, the author analyzes the legal parameters of hate speech as persecution (a crime against humanity) and hate speech as instigation (a mode of liability). In both instances, the human rights law specialist finds that interpretations of hate speech are consistent with the earlier ad hoc tribunals' jurisprudence and, more generally, with international human rights law. The author concludes that current laws allow criminalization only of the most extreme forms of incitement to violence.

Gelber, Katharine. "[Differentiating Hate Speech: A Systemic Discrimination Approach](#)." *Critical Review of International Social and Political Philosophy*. 2019. 1-22.

> Katharine Gelber, a professor of Politics and Public Policy at the University of Queensland develops a systemic discrimination approach to defining a narrowly construed category of 'hate speech', as speech that harms to a sufficient degree to warrant government regulation. In this article, she extends current literature on how hate speech can harm by identifying speakers' circumstantial capacity to harm, along

with targets' harm vulnerability. Finally, she bridges the gap between conceptual understandings of hate speech and policy designed to regulate it.

IN ONLINE SPACES

In the News:

Edelman, Gilad. "[The Parler Bans Open a New Front in the 'Free Speech' Wars.](#)" *Wired.*
January 13, 2021.

The article contextualizes Parler's removal from Google and Apple platforms within
larger conversations around tech's overwhelming power to define hate speech. Gilad
Edelman points to the companies' ability to de-platform apps, along with other tech
efforts by companies like PayPal and GoDaddy stop servicing specific client based on
their links to acts of violence. Armed with this ability to 'meta-moderate', tech
platforms have an unprecedent ability to re-shape current understandings of hate
speech through direct action.

Sulleyman, Aatif. "[Elon Musk on the Difference between Big Tech Banning Hate Speech and
Hated Speech.](#)" *Newsweek.* January 12, 2021.

This article sheds light on Tesla CEO, Elon Musk's, latest commentary on the state of
hate speech in the American online sphere. After deleting Tesla and Space X's
Facebook pages in 2018, the CEO shares that in his view, Tech companies are defining
hate speech merely as speech that they hate. Following Trump's account suspension
on various platforms, the CEO publicly stated that Big Tech is the 'de facto arbiter of
free speech'.

Ramaswamy, Vivek, and Jed Rubenfeld. "[Save the Constitution from Big Tech.](#)" *The Wall
Street Journal.* January 11, 2021.

In response to the suspension of President Trump's Twitter and Facebook accounts,
this piece argues for the prosecution of tech companies' censorship of free speech.
The authors, tech entrepreneur Vivek Ramaswamy and Yale law professor Jed
Rubenfeld, critique the companies' ever changing 'terms of service' along with blanket
claims of 'hate speech'. To make their argument, the authors point to several cases
setting precedent where private sector action was treated as state action.

Radsch, Courtney. "[Trump v. Tech: What Is Censorship and Who Gets to Do It?](#)" *Medium*. January 11, 2021.

In her article, Courtney Radch contextualizes tech companies' action against hate speech in the larger framework of infrastructural censorship. The Advocacy Director at the Committee to Protect Journalists points to the long history of various platforms' ability to mute out content they deem dangerous. Rather than framing them as private companies rejecting the use of their services to specific actors, Radch argues that the pervasiveness of these platforms effectively govern our social, economic, and digital lives. Therefore, she equates their action to censorship to be remedied with government regulation, and self-regulation by the companies.

York, Jillian C. "[The Best Way to Protect Free Speech on Social Media Is to Promote Competition.](#)" *MIT Technology Review.* January 10, 2021.

In her commentary, Jillian York tackles a central issue to the free vs hate speech debate: of content moderation and account suspension. York, director for international freedom of expression at the Electronic Frontier Foundation, points to the familiarity of Trump's account suspension by giving examples of other world politicians and political parties who have been censored on social media platforms. In terms of policy response, York opposes repealing section 230 – a law that protects companies from liability for the decisions they make about the content they host. Instead, she prompts government to steer tech companies in crafting better content moderation based on research and including civil society actors.

Nonprofit and NGO Reports:

Whitehouse, John. "[Facebook Knows that Right-Wing Pages Spread Hate Speech and has Decided to Do Nothing About It](#)." *Media Matters*. December 11, 2020.

This article documents the history of Facebook's treatment of far-right media, as well as the platform's track record on civil rights issues. In doing so, it argues that the social media company has allowed dangerous hate speech to proliferate on its platform. The article also guides the reader to other resources that provide more instances of tech companies' alleged support of far-right outlets over the years.

Academic Sources:

Al-Hassan, Areej, and Hmood Al-Dossari. "Detection of Hate Speech in Social Networks: A Survey on Multilingual Corpus." *In 6th International Conference on Computer Science and Information Technology*. 2019.

In their study, Areej Al-Hassan and Hmood Al-Dossari present a background on hate speech and its related detection approaches. Based on the demand for automatic detection of hate speech, the authors highlight the scarcity of complex language content detection (e.g. Arabic). Specifically, the authors point to the difficulties faced by Natural Language Processing tools when dealing with languages like Arabic. Finally, Al-Hassan and Al-Dossari point to Machine Learning Models and Deep Leaning architectures as the way forward for content detection in complex languages.

Chetty, Naganna, and Sreejith Alathur. "Hate Speech Review In the Context of Online Social Networks." *Aggression and Violent Behavior*. 2018. 40.

Naganna Chetty and Sreejith Alathur define hate speech as "an offensive kind of communication mechanism that expresses an ideology of hate using stereotypes." Although their work outlines different categories of speech from terrorism to extremist ideological speech, the authors focus on is its prevalence on online social networks. Based on those definitions, the authors also point to the discrepancies in hate speech laws globally and point to Internet Service Providers (ISPs) and online social networks, to effectively counter both hate speech and terrorism.

ACADEMIA

Nonprofit and NGO Reports:

The Knight Foundation. "The First Amendment on Campus 2020 Report." 2020.

> This collaborative effort by the Knight Foundation and Gallup offers an overview of current student views on campus free speech. The report is the result of online surveys of over 3,000 U.S. full-time undergraduate college students, including an oversample of students at historically black colleges and universities (HBCUs) between 2016-2019. The report highlights an overall increase in students' need for targeted speech restrictions, along with nuances dependent on racial and gender gaps when it comes to the extent to which students see the First Amendment as a safeguard and on whether colleges should protect students from certain speech.

Academic Sources:

Barendt, Eric. "What Is the Harm of Hate Speech?." *Ethical Theory and Moral Practice.* 2019. 22-3.

> Eric Barendt, a Law professor at University College in London, argues against weak legal definitions of hate speech. He focuses his critique on the longstanding causal link between hate speed and the harm caused by it. In Barendt's view, reconsiderations of hate speech can only be based on constitutive arguments proving the harm in the speech itself rather than its potential causes.

Brown, Alexander. "What is Hate Speech? Part 1: The Myth of Hate." *Law and Philosophy*. 2017. 36-4.

> Political and law theorist Alexander Brown provides a conceptual analysis of the ordinary term "hate speech". His analysis favors dynamic definitions of hate speech, while rejecting that emotions, feelings, or attitudes of hate must be part of the essential nature of hate speech. Part of larger theoretical debates on the issue, his work elaborates on the multiple meanings of hate speech without the need to involve the literal meaning of "hate.

Brown, Alexander. "What is Hate Speech? Part 2: Family Resemblances." *Law and Philosophy*. 2017. 36-5.

Digging deeper on his previous argument rejecting a singular understanding of hate speech, Alexander Brown argues that hate speech is better analyzed as a family resemblance concept. This means that remedies of hate speech should be sensitive to the heterogeneous collection of expressive phenomena that fall under the category of hate speech.

Howard, Jeffrey W. "Free Speech and Hate Speech." *Annual Review of Political Science.* 2019. 22.

Should hate speech be banned? In Jeffrey Howard's view, current participants in the debate are speaking past each other. To remedy this misalignment, the political philosopher proposes a disaggregation of the concept into smaller analytical stages. First, he tackles the scope of the moral right to freedom of expression, and whether hate speech falls within the right's protective range. If it does, hate speech bans are necessarily unjust. If not, he turns to the second stage, by assessing whether speakers have moral duties to refrain from hate speech. The article canvasses several possible duties, including the duty not to threaten, harass, offend, defame, or incite. This multi-stage approach presents an in-depth definition of hate speech with corresponding legal enforcements.

Bangstad, Sindre. "Can There Be An Anthropology of Hate Speech?." *Anthropology News*. 2017. 58-3.

While recognizing the lack of direct causation between hate speech and hate crime, Sindre Bangstad sheds some light on the benefits of an anthropological perspective on the issue. First, he underlines anthropology's disciplinary focus on the lived experiences of hate speech victims rather than abstract political philosophical theory. Here, anthropology would then lead to practical solutions rather than a continuation of previous abstract conceptualizations of hate speech. Most importantly, anthropology is a contextualized framework with particular focus to power relations which are of key importance in issues of hate speech.