

PROBLEMS WITH PROBLEMATIC SPEECH ON SOCIAL MEDIA

William Fleischman, Leah Rosenbloom

Villanova University (USA), The Workshop School (USA)

william.fleischman@villanova.edu; leah.rosenbloom@workshopschool.org

ABSTRACT

In this paper, we consider some of the tensions and conflicts between freedom of speech on the Internet, and other public goods and individual rights. The dimensions of the problem include: Threats of physical violence to individuals; threats directed at groups defined by ethnic, national, religious, sexual or gender identity, or political orientation; abusive, harassing, and/or hateful speech; incitement to self-harm; doxing; social exclusion; and dissemination of false information. Since initiating this study, we have also come to see an additional dimension, the importance of which we were slow to recognize. This is a pattern of misleading, self-contradictory, content-free, and deceptive speech on the part of spokespersons for one of the dominant social media platforms – Facebook. We make provisional suggestions for discouraging the actions of troll armies and for applying more vigorous measures of transparency in regard to political advertising on social media.

KEYWORDS: freedom of speech, violent and abusive speech, internet, social media.

1. INTRODUCTION

The popularization of the Internet promised a radical democratization of communication: Everyone can be a publisher, cost of entry is low, and access is available to anyone connected to the Internet. But early on, prescient individuals understood that “cheap speech,” in Eugene Volokh’s pungent phrase, carried other implications not all of which are entirely conducive to the dissemination of reliable information or the reasoned discourse of the marketplace of ideas.

As Tim Wu points out in “Is the First Amendment Obsolete?” (Wu, 2017), the assumption that the most serious threats to freedom of speech come principally from governmental actors is no longer entirely valid. Direct censorship, either in the form of government action or by content filters and human content monitors employed by social media platforms, can now be supplemented or supplanted by the actions of privately constituted troll armies or bands of individuals and/or robots programmed to drown out disfavoured speech. These means are at the disposal of powerful private interests and loosely organized partisan groups.

In this paper, we consider some of the tensions and conflicts between freedom of speech on the Internet, and other public goods and individual rights. We argue that the widest scope should be afforded individuals’ right to free expression, but believe that social media platforms should be held to certain standards of responsibility for preventing or redressing harms resulting from speech on these platforms.

2. THE TROUBLED AND VIOLENT TERRAIN...

“We’ve got a speech problem on the Internet!” is an observation that covers a lot of ground. The dimensions of the problem include: Threats of physical violence to individuals; threats directed at groups defined by ethnic, national, religious, sexual or gender identity, or political orientation; abusive, harassing, and/or hateful speech; incitement to self-harm; doxing; social exclusion; and dissemination of false information.

2.1. Troll Armies

Gamergate (Wikipedia, 2019) is a well-known example of an online hate mob. Lately, troll armies have figured prominently in polarized political discourse. Tim Wu (2017) cites two examples: David French, a writer associated with the conservative *National Review*, and Rosa Brooks, a professor of law at Georgetown University, both targets of online mobs for criticism of the current U.S. president.

The rhetoric was murderous and hateful in both instances – Nazi imagery and the face of his daughter in the gas chamber in the case of French (French, 2016), and extremely violent misogynistic language directed at Brooks. (Brooks, 2017)

2.2. Reverse Censorship and Flooding

Another technique used by governments to marginalize dissident speech involves mobilizing a volume of opposing information to drown out inconvenient speech or distort the informational environment to render the speech dubious and unimportant. An important variant of reverse censorship, used in political advertisement, floods public discourse with patently false information or “fake news.” It is widely understood that in 2016 targeted political advertisements disseminating false information were instrumental in both the U.K. Brexit referendum (Cadwalladr, 2019) and the U.S. Presidential election. (Lapowsky, 2018)

3. ... NONETHELESS THERE ARE GOOD REASONS TO PROTECT EVEN EXTREME SPEECH...

In spite of these examples, there are important reasons to favor protecting even extreme speech on the web.

3.1. Free Speech Protects Protest Movements

Online speech is a cornerstone of modern social change. Organizers rely heavily on social media and online communication to disseminate information and coordinate action. The first widely-studied instance of online organizing was the Arab Spring, during which Egyptian and Tunisian dissidents used Facebook, Twitter, and blogs to discuss and promote revolutionary ideas before taking to the streets (Howard et al., 2011). More recently, University students in China relied heavily on social media to share information and encourage participation in Hong Kong’s Umbrella Movement (Lee et al., 2016). Social media was also a vital part of the Euromaidan uprising in Ukraine (Bohdanova, 2014), and the native environmental movement in Standing Rock, North Dakota (Johnson, 2017). Governments recognize the potential of social media to

amplify political discontent, which is why they censor and block posts, platforms, and sometimes even the entire Internet.

Social media is appealing to activists for its immediacy and accessibility. Allowing governments to regulate content on social media may introduce leeway for them to further silence activists and destabilize revolutionary movements.

3.2. Governments Stifle Speech to Protect Private Interests

Systemic regulation and censorship of online speech often goes hand in hand with other socially repressive tactics, for instance the incarceration and “re-education” of political dissidents in China (Human Rights Watch, 2017). While the most extreme examples are dictatorial governments that overtly censor and crush opposing voices, democratic governments also monitor and undermine dissent, especially when private financial interests are involved.

Among the most influential private sector interests is Big Oil. Global financial interest in the acquisition and distribution of oil has been a key driver of worldwide surveillance and censorship, even in countries with robust free speech protections. In the UK, counter-terrorism police labeled the non-violent environmental group Extinction Rebellion alongside neo-Nazis as an “extremist ideology” (Dodd & Grierson, 2020). Under the current policies of most social media platforms, a “terrorist” designation is grounds for immediate permanent dismissal from the platform. Any individual users found in support of “terrorism” would be similarly censored or dismissed.

During the Dakota Access Pipeline protests of 2016, independent media collective Unicorn Riot reported the disproportionate censorship and arrest of social media journalists, including Facebook’s (purportedly accidental) removal of a protest livestream for violating community standards (Unicorn Riot, 2016). Facebook has a designed “Law Enforcement Online Request System” that law enforcement can use to make requests for content forfeiture and removal. The specific nature and frequency of Facebook’s compliance with those requests is not public knowledge.

3.3. “Dangerous” Speech Is Defined by People in Power

It is natural for governments to want to minimize speech that encourages violence on its citizens, or otherwise undermines national interests. There is a delicate line, however, between censorship to maintain citizens’ health wellbeing, and censorship to maintain governments’ power and authority. Governments have been known to leverage the public’s need for security to censor and oppress opposition, and this is unlikely to change in the digital age. It is necessary to protect free speech rights not just for activists and dissidents, but for all people.

Defending free speech for everyone is not easy, but it is vital in order to maintain free speech for those who need it most. American Civil Liberties Union director Anthony Romero emphasizes the importance of defending neo-Nazis’ right to peaceably assemble: “We simply never want [the] government to be in a position to favor or disfavor particular viewpoints. And the fact is, government officials...are more apt to suppress the speech of individuals or groups who disagree with government decisions” (Romero, 2017). For over half its history, the United States government was more likely to agree with the KKK than the NAACP.

4. ... BUT NOT NECESSARILY WITHOUT ANY LIMITS

First Amendment protections are intended to restrain the power of government to interfere with the freedom of expression of individuals. However, speech online occurs through an internet service provider or a social media platform which inherit First Amendment protections. As private enterprises they can and do set rules that should apply to their users. Often these rules are vague and inconsistently applied. In particular, when speech policies come into conflict with the profit motive of the platform, these policies frequently evaporate. The result is often flagrantly abusive, obscene, threatening or deceptive speech. We propose adoption of stricter and clearer standards and procedures to limit the harms associated with three aspects of speech on social media – troll armies, incitement, and dissemination of political advertisements containing verifiably false content.

We believe that social media platforms should hold themselves to standards and policies that are consistently applied and promote more responsible speech. Reddit provides an example that shows this is possible. (Marantz, 2016)

4.1. The Case of Troll Armies

We propose that in cases like those cited, where troll armies coordinate hateful speech and threats against an individual, the social media platform that facilitates such an attack take action against members of the mob. The principle here is that those claiming the right to speak violently and abusively under the doctrine of freedom of expression are, in fact, acting to attempt to silence another individual, and therefore, perversely curtailing the very same right of that individual.

These situations should be relatively easy to document, once the attacked individual registers a complaint with the social media platform, since they consist of N --> 1 more or less synchronized messages (multiple sources, one target). We recognize the limitations of algorithmic detection or wholesale human moderation of every instance of hateful and threatening speech. By contrast, the cases to which we refer are not so frequent that human inspection would be impossibly difficult. Naturally, this requires nuanced consideration of the multiple messages, some of which may be reasoned arguments expressed in strong language and should be differentiated from those that simply spew hate in language and (photoshopped) images.

4.2. Incitement

In the United States, legal theory governing cases involving incitement is not entirely satisfactory. The current standard for determining whether speech constitutes illegal incitement comes from the U.S. Supreme Court opinion in the 1969 case, *Brandenburg v. Ohio*. "There, the Court held that advocacy of violence is protected unless it 'is directed to inciting or producing imminent lawless action and is likely to incite or produce such action.'" (Pew, 2015) The *Brandenburg* precedent has been cited on numerous occasions but there has been a certain difficulty. Conflicting interpretations of the word "imminent" have given rise to inconsistency in interpretation. According to Pew (2015), there is a consensus forming around the interpretation that "imminent" refers to "a matter of several days."

However, the well-known case, *Planned Parenthood of Columbia/Willamette, Inc. v. Am. Coalition of Life Activists*, illustrates another difficulty in application of *Brandenburg*. This case

stemmed from a 1995 action “in which antiabortionists uploaded approximately 200 more physicians’ names to a website, again including their photographs and addresses. Some of the physicians’ names were crossed out, others were in grey font, and the rest were in black font. The following legend accompanied the files: ‘Black font (working); Greyed-out Name (wounded); Strikethrough (fatality).’ In other words, the website recorded murders and other violent attacks against the abortion doctors. The names of the three doctors who had been murdered from 1993 to 1994 were struck through. Several physicians featured on the website, terrified for their lives, brought suit.” (Pew, 2015)

The original decision by a three-judge panel of the 9th Circuit held that the contested speech was protected in view of the lack of time frame indicating an “imminent” threat. However, the decision was reversed by the 9th Circuit sitting *en banc*. Although the court held the speech was protected under *Brandenburg*, it found another basis for declaring it unprotected. Under the “true threats doctrine,” since so many of the physicians identified on the website had already been killed or injured, the court held that no one posting the photographs, names, and addresses could believe otherwise than that those who were targeted by the website would live in fear that they might be the next target of an assailant in real life. (Pew, 2015)

The difficulty of “drawing lines” in cases bordering on incitement is apparent. On the other hand, living in a world in which the protracted state of anguish caused by the flood of hatred that engulfed David French and his family is “normal” seems deeply unsatisfactory. It seems that the only recourse in situations of this sort consists of protective reactions by the affected individual(s) such as blocking those responsible for virulently hateful attacks, and avoiding those sites where active participation results in further abuse. But how is this consistent with the idea of the web as the modern incarnation of the marketplace of ideas and reasoned discourse?

These difficulties are the result of the protections provided by the First Amendment (at least in the context of the U.S. Constitution) against government censorship of speech by individual citizens. But the situation is significantly altered when the speech occurs on a social media platform whose ownership is in private hands and whose owners have the freedom to set rules promoting a marketplace of ideas rather than a marketplace of murderous invective. Considered from this standpoint, we encounter some truly puzzling stories.

For one particularly flagrant example, there is the following: “A journalist on Monday tweeted, without naming them directly, that “they” need to be “killed” before “they kill us.” Although interpretations of the tweet may differ from person to person, many saw it as a tweet advocating violence against members of a particular community – Muslims – and called it genocidal in its intent. Many also reported the tweet as well as the account as abusive or advocating violence. However, Twitter doesn’t find anything wrong with the tweet and replied to many saying that the tweet advocating murder of people doesn’t violate its rules. (India Today Tech, 2018)

The full text of the tweet is as follows: “They killed us in Trains, Hijacked our Planes, held us Hostage in Hotels, Forced us to flee #Kashmir, & now Killing us for holding the Tricolor on #RepublicDay.

Truth is We Live in Fear, NOT They.

NO more. Always Carry Lethal Weapons. KILL them before they KILL us.

#MondayMotivation”

“The tweet was made from a profile that is verified and it is possible that because of the popularity of the account, Twitter decided that exhortations to kill people was probably alright to tweet from this particular account.” (India Today Tech, 2018)

Apparently, incitement to genocide is permitted if your Twitter profile is verified and popular.

In this context, it seems appropriate to quote the language pertaining to incitement articulated in Articles 19 and 20 of the International Covenant on Civil and Political Rights (ICCPR):

“While the right to freedom of expression is fundamental, it is not absolute. A State may, exceptionally, limit the right under Article 19(3) of the ICCPR, provided that the limitation is:

- Provided for by law, so any law or regulation must be formulated with sufficient precision to enable individuals to regulate their conduct accordingly;
- In pursuit of a legitimate aim, listed exhaustively as: respect of the rights or reputations of others; or the protection of national security or of public order (ordre public), or of public health or morals; or
- Necessary in a democratic society, requiring the State to demonstrate in a specific and individualised fashion the precise nature of the threat, and the necessity and proportionality of the specific action taken, in particular by establishing a direct and immediate connection between the expression and the threat.

Article 20(2) of the ICCPR obliges States to prohibit by law ‘any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.’” (Article 19. 2018)

In particular, with regard to incitement:

- “Incitement. Prohibitions should only focus on the advocacy of discriminatory hatred that constitutes incitement to hostility, discrimination, or violence, rather than the advocacy of hatred without regard to its tendency to incite action by the audience against a protected group.
- Six-part threshold test. To assist in judicial assessments of whether a speaker intends and is capable of having the effect of inciting their audience to violent or discriminatory action through the advocacy of discriminatory hatred, six factors should be considered:
- Context: the expression should be considered within the political, economic, and social context prevalent at the time it was communicated, for example the existence or history of conflict, existence or history of institutionalised discrimination, the legal framework, and the media landscape;
- Identity of the speaker: the position of the speaker as it relates to their authority or influence over their audience, in particular if they are a politician, public official, religious or community leader;

- Intent of the speaker to engage in advocacy to hatred; intent to target a protected group on the basis of a protected characteristic, and knowledge that their conduct will likely incite the audience to discrimination, hostility, or violence;
- Content of the expression: what was said, including the form and the style of the expression, and what the audience understood by this;
- Extent and magnitude of the expression: the public nature of the expression, the means of the expression, and the intensity or magnitude of the expression in terms of its frequency or volume; and
- Likelihood of harm occurring, including its imminence: there must be a reasonable probability of discrimination, hostility, or violence occurring as a direct consequence of the incitement.” (Article 19. 2018)

Our contention is that, with regard to incitement, social media platforms should have policies of self-regulation that at least meet the standard articulated in the ICCPR as described above. There is no rationalization, other than a shameless addiction to value-insensitive economic gain, that can justify labeling a message that calls for mass murder “in compliance with the rules of permissible expression” on a social media platform.

There is something profoundly discordant about the fact that, deliberating a case involving speech that was clearly intended to incite violence against identified individuals without specifying a time frame that would trigger application of the Brandenburg precedent, the 9th Circuit Court was able to find a basis for ruling the speech unprotected, whereas social media platforms which, as private entities, are under no obligation to adjudicate the nicety of determining the threat “imminent,” cannot bring themselves to act in a conservative fashion when faced by a speech act of similarly explosive and violent intent.

Of course, we understand that if it is a speech act with the potential to generate numerous “clicks” and contribute significantly to the platform’s bottom line, all bets are off.

4.3. The Case of False Political Advertising

Our other proposal has to do with political speech - specifically political ads that circulate false or discredited information. Facebook is currently involved in such a dispute. We don't want to say that these things should be outlawed - there's plenty of history, going back to the election of 1800 in our country, of scurrilous political speech (McCullough, 2001). But it is troubling that ads on Facebook and other platforms appear and then disappear without any trace so there is no possibility of auditing them or providing public scrutiny. They are particularly pernicious because they are targeted to people identified as susceptible through analysis of their Facebook profiles.

Carol Cadwalladr (2019) has documented how this occurred in the Brexit referendum and how Facebook has stonewalled any serious attempt to investigate the sources of funding and means of targeting these false and vanishing ads. Facebook executives have been notably oblivious and evasive about such advertising. (Lee, 2018)

Our proposal is to force the social media platform to keep publicly accessible, auditable records of political ads so that they can be scrutinized and rebutted in the same way that's possible with

ads on other media. We are not alone in thinking that this is a reasonable measure for curbing the most egregious excesses of dishonesty in political advertising.

U.S. Senator Amy Klobuchar of Minnesota introduced a bill in 2017 bill – S. 1989, The Honest Ads Act – which appears to have stimulated pre-emptive action on the part of several social media platforms. At the present moment, Twitter has announced a complete ban on political ads.

Although Facebook claims to have set up an archive of the description we favor there are reasons to think it will fall short of the promise of promoting greater transparency in political advertising on social media. For one thing, Facebook continues to resist any voluntary action to remove ads that contain verifiably false content. In addition, we have the discouraging episode related in a recent article “On Dec. 10, [2019], just two days before the United Kingdom went to the polls, some 74,000 political advertisements vanished from Facebook’s Ad Library, a website that serves as an archive of political and issue ads run on the platform. For a while, what the company described as a “bug” wiped 40% of all political Facebook ads in the UK from the public record.” In fact, this was just one of a litany of disturbing failures that undercut the usefulness of the archive. (Smith, 2020)

5. FIRST YOU SAY YOU DO, AND THEN YOU DON’T, THEN YOU SAY YOU WILL...

Long ago, James Moor foresaw that computers would offer new capabilities and choices for action; these possibilities would, in turn, require new policies or call into question the adequacy of existing policies for ethical conduct in deployment of these new choices. He predicted further that the attempt to remedy an existing policy vacuum might bring us face to face with an underlying conceptual vacuum. (Moor, 1985)

This is an apt description of the problem of characterizing social media platforms in regard to the dissemination of news. Should these platforms be seen as neutral technology companies that have simply built an ingenious set of tools for passing along content including entertainment, artistic creation and news, or have they, in fact, evolved to play a role in the sphere of public information comparable to traditional media companies without having assumed any of the traditional responsibilities and ethical norms that define the legal and social expectations of journalism? What should we call them? What do they call themselves?

In the words of a song made famous by Ella Fitzgerald and Louis Armstrong,

“First, you say, you do
And then you don't
And then you say, you will
And then you won't
You're undecided now
So what are you gonna do?” (Genius, 2020)

If we are permitted to paraphrase somewhat facetiously: “First you say you are, and then you’re not. And then you say you’re not and then you are.” This seems to be the stance of Facebook, Google, Twitter, and all the major social media platforms when confronted with the question as to whether they are tech companies or publishers.

Except that it has nothing to do with being undecided. It is, rather, a matter of exploiting what Shoshana Zuboff, echoing James Moor, has characterized as “a lag in social evolution” in the face

of the rapid build-out of social media platform capabilities that “outrun public understanding and the eventual development of law and regulation that it produces.” (Zuboff, 2015) Many observers representing very different perspectives on the spectrum of political affiliation and belief (Levin, 2018; Dougherty, 2019; Shaw, 2019) have noted the disparity between the public posture of Facebook as a self-identified tech platform and the contrary representations it makes in court filings where it seeks the protection accorded to publishers concerning decisions made about “what not to publish.” (Levin, 2018)

We believe that regulating social media, based on the substantial advertising revenues realized from activity properly described as that of a publisher of news would, on the one hand, cut through the conceptual ambiguity cynically and opportunistically exploited by platforms like Facebook and, on the other hand, provide a basis for requiring that such media platforms put “their houses in order” by means of consistently applied journalistic oversight in regard to hate speech and incitement, the publication of which they permit.

The standard should be “You may publish anything the law allows as long as you refrain from monetizing it. If, however, you wish to derive streams of revenue from advertising associated with provocative speech, then apply the standards of good journalistic practice and responsibility to the publication thereof.”

REFERENCES

- Article 19. (2018) Responding to ‘hate speech’: Comparative overview of six EU countries. *Article 19*. Retrieved from <http://europeanjournalists.org/mediaagainsthate/wp-content/uploads/2018/02/Final-compilation-off-regional-research-digital.pdf>
- Bohdanova, T. (2014). Unexpected revolution: the role of social media in Ukraine’s Maidan uprising. *European View*, vol. 13, pp. 133-142. <http://doi.org/10.1007/s12290-014-0296-4>
- Brooks, R. (2017). And then the Breitbart lynch mob came for me, *Foreign Policy*. Retrieved from <https://foreignpolicy.com/2017/02/06/and-then-the-breitbart-lynch-mob-came-for-me-bannon-trolls-trump/>
- Cadwalladr, C. (2019). Facebook’s role in Brexit and the threat to democracy. Retrieved from https://www.ted.com/talks/carole_cadwalladr_facebook_s_role_in_brexit_and_the_threat_to_democracy?language=en
- Dodd, V. & Grierson, J. (2020, 10 January). Terrorism police list Extinction Rebellion as terrorist ideology. *The Guardian*. Retrieved from <https://www.theguardian.com/uk-news/2020/jan/10/xr-extinction-rebellion-listed-extremist-ideology-police-prevent-scheme-guidance>
- Dougherty, M. (2019, 25 June). What is Facebook? *National Review*. Retrieved from <https://www.nationalreview.com/2019/06/facebook-legal-status-platform-publisher/>
- Feiner, L. (2020, 14 February). Facebook won’t count influencer posts like Bloomberg’s memes as political ads. Retrieved from <https://www.cnn.com/2020/02/14/facebook-wont-archive-sponsored-political-ads.html>
- French, D. (2016). The price I’ve paid for opposing Trump, *National Review*. Retrieved from <https://www.nationalreview.com/2016/10/donald-trump-alt-right-internet-abuse-never-trump-movement/>

- Genius (2020). Undecided Lyrics, *Genius.com*. Retrieved from <https://genius.com/Ella-fitzgerald-and-louis-armstrong-undecided-lyrics>
- Howard, P., Duffy, A., Freelon, D., Hussain, M., Mari, W. & Mazaid, M. (2011). Opening closed regimes. *Project on Information Technology and Political Islam*. Working paper 2011.1. Retrieved from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/117568/2011_Howard-DuffyFreelon-Hussain-Mari-Mazaid_PITPI.pdf?sequence=1
- Human Rights Watch (2018, 26 February). China: Big Data fuels crackdown in minority region. Retrieved from <https://www.hrw.org/news/2018/02/26/china-big-data-fuels-crackdown-minority-region#>
- India Today Tech (2018, 30 January). Tweet calling for killing of Muslims in India doesn't violate rules, says Twitter. Retrieved from <https://www.indiatoday.in/technology/news/story/tweet-calling-for-killing-of-muslims-in-india-doesn-t-violate-rules-says-twitter-1156743-2018-01-29>
- Johnson, H. (2017). #NoDAPL: social media, empowerment, and civic participation at Standing Rock. *Library Trends*, vol. 66 no. 2, pp. 155-175. Retrieved from <https://muse.jhu.edu/article/686889/pdf>
- Kelly, J., Blood, D. & O Murchu, C. (2019, 10 December). Facebook under fire as political ads vanish from archive. *Financial Times*. Retrieved from <https://www.ft.com/content/e6fb805e-1b78-11ea-97df-cc63de1d73f4>
- Lapowsky, I. (2018). Mark Zuckerberg speaks out on Cambridge Analytica Scandal, *Wired*. Retrieved from <https://www.wired.com/story/mark-zuckerberg-statement-cambridge-analytica/>
- Lee, D. (2018). Mark Zuckerberg, missing in inaction, *BBC News*. Retrieved from <https://www.bbc.com/news/technology-46231284>
- Lee, F. Chen, H-T. & Chan, M. (2016). Social media use and students' participation in a large-scale protest campaign: the case of Hong Kong's Umbrella Movement. *Telematics and Informatics*, vol. 34, pp. 457-469. Retrieved from https://www.researchgate.net/profile/Hsuan_Ting_Chen/publication/306075497_Social_Media_Use_and_University_Students'_Participation_in_a_Large-scale_Protest_Campaign_The_Case_of_Hong_Kong's_Umbrella_Movement/links/59e0313045851537160163a2/Social-Media-Use-and-University-Students'-Participation-in-a-Large-scale-Protest-Campaign-The-Case-of-Hong-Kongs-Umbrella-Movement.pdf
- Levin, S. (2018, 3 July). Is Facebook a publisher? In public it says no, but in court it says yes, *The Guardian*. Retrieved from <https://www.theguardian.com/technology/2018/jul/02/facebook-mark-zuckerberg-platform-publisher-lawsuit>
- Marantz, A. (2018). Reddit and the struggle to detoxify the internet, *New Yorker*. Retrieved from <https://www.newyorker.com/magazine/2018/03/19/reddit-and-the-struggle-to-detoxify-the-internet>
- Moor, J. (1985). What Is Computer Ethics?, *Metaphilosophy*, vol. 16, no. 4, pp. 266-275.
- McCullough, D. (2001, at 545ff). *John Adams*, Simon & Schuster, New York
- Nelson, J. (2017). Should Google and Facebook be considered media companies? Retrieved from <https://rossdawson.com/blog/google-facebook-considered-media-companies/>

- Patterson, B. (2017). Police Spied on New York Black Lives Matter Group, Internal Police Documents Show, retrieved from <https://www.motherjones.com/crime-justice/2017/10/police-spied-on-new-york-black-lives-matter-group-internal-police-documents-show/>
- Pew, B. (2015). How to Incite Crime with Words: Clarifying Brandenburg’s Incitement Test with Speech Act Theory, *BYU Law Review*, vol. 2015, issue 4, no. 8. Retrieved from <https://digitalcommons.law.byu.edu/cgi/viewcontent.cgi?article=2996&context=lawreview>
- Romero, A. (2017, 15 August). Equality, justice and the First Amendment. Retrieved from <https://www.aclu.org/blog/free-speech/equality-justice-and-first-amendment>
- Shaw, C. (2019, 20 September). Facebook admits in court filing that it is a publisher, opening itself up to libel suits. *The New American*. Retrieved from <https://www.thenewamerican.com/tech/computers/item/33469-facebook-admits-in-court-filing-that-it-is-a-publisher-opening-itself-up-to-libel-suits>
- Smith, R. (2020, 14 January). The UK election showed just how unreliable Facebook’s security system for elections really is. *BuzzFeed News*. Retrieved from <https://www.buzzfeednews.com/article/rorysmith/the-uk-election-showed-just-how-unreliable-facebooks>
- Unicorn Riot (2016, 20 September). Statement on recent censorship of #NoDAPL coverage. Retrieved from <https://unicornriot.ninja/2016/statement-recent-censorship-nodapl-coverage/>
- Wikipedia (2019). Gamergate controversy. *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Gamergate_controversy
- Wu, T. (2017) Is the First Amendment obsolete? Knight Foundation First Amendment Institute. Retrieved from <https://knightcolumbia.org/content/tim-wu-first-amendment-obsolete>
- Zuboff, S. (2015), Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, vol. 30, pp. 75-89.