AMITAI ETZIONI

Should We Privatize Censorship?

Over the past few decades, the United States has increasingly relied on the private sector to carry out missions that were once considered largely the mission of the state. There are now more private cops (security guards and the like) than public police; there have been about the same number of private contractors carrying out the wars in Afghanistan and Iraq as soldiers. There is a sharp rise in for-profit prisons, and even collecting intelligence and torture have been privatized to some extent. Most recently, major technology companies such as Facebook, Google, and Twitter have been, in effect, deputized as censors. They have been pressured to remove a great variety of messages deemed offensive by large parts of the public and its elected representatives. The result is a legal, ethical, and administrative mess, but it is not easy to develop a viable alternative.

Some argue that because tech corporations are private companies they cannot censor; only the government can. The First Amendment states that Congress shall make no law abridging freedom of the press, not that private companies cannot control messages they host or circulate. Moreover, one company might have no objection to posting a message that another company has banned. Only the government can prevent access to all mediums and thus truly censor.

However, given that these companies control a very large amount of the communication space—50% or more in some categories—if they restrict someone's access, that person's power of self-expression is greatly limited. Anyone banned from Google, Facebook, and/or Twitter will find it very difficult to reach the masses through social media.

For many years after their inception, tech companies largely avoided responsibility for the content posted on their sites. The tech giants argued that they are platforms, not publishers. The Communications Decency Act of 1996 sought to regulate indecent material, chiefly pornography, on the internet. Free speech advocates, fearful that such regulation would have a chilling effect on all speech, convinced Congress to include this language in the bill: "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider." In other words, online service providers are not liable for content posted by third parties.

Over the years that followed, various critics argued that the tech companies should control content. These sentiments reached a high point following the revelations about Russia's meddling in the 2016 US elections to sow social discord and division through coordinated social media misinformation campaigns. However, Congress has been reluctant to regulate anything involving speech and has in effect shifted responsibility to the tech companies.

Having heard the message, tech companies have taken several key steps to control content. They hired tens of thousands of moderators to continuously review posts and remove violent, lewd, hateful, and misleading material. The companies are also increasingly using artificial intelligence (AI). Facebook uses image matching, language-understanding software, cluster group targeting, and fake account detection to crack down on repeat offenders espousing radical ideology. Facebook developed DeepText, which it describes as "a deep learning-based text understanding engine that can understand with near-human accuracy the textual content of several thousand posts per second, spanning more than 20 languages." Instagram trained the DeepText algorithm to identify words and comments deemed toxic and to filter them out without human input. Google uses AI to scrub terrorist videos from YouTube, reporting that it is faster and more accurate than human moderators.

Finally, tech companies are also relying on community engagement, with users flagging offensive speech for review by companies' moderators. And some companies are partnering with third-party fact-checkers, such as Snopes and Politifact, to review flagged fake news content.

Administrative challenges

All these censoring activities present major challenges for tech companies due to the large volumes involved—billions of unique posts every day. In some cases human moderators have as little as 10 seconds to review a post. They can hardly take much longer given the astronomical number of posts that must be reviewed. The work pays poorly and is psychologically stressful. Furthermore, different companies have different standards for what is

> **Tech companies seem to be limiting online speech much more than Congress and the courts restrict offline content.**

considered a violation of their user agreement. Thus, for a while Alex Jones was silenced on Facebook, but his voice was carried by Twitter.

Companies face difficulty moderating material in countries where they do not have sufficient language experts to properly review content. Facebook often relies on translation software in such countries, but that is far from foolproof. In Bosnia, a failure to properly translate text resulted in perpetrating the falsehood that an imprisoned war criminal was still a fugitive.

Another pressing problem is that companies face

different demands in different countries. For instance, hate speech is banned in many western democracies but not in the United States. Hence, citizens in countries in which some content is banned can nevertheless access it on the World Wide Web, using various applications to circumvent censorship.

No wonder the results often seem capricious both in terms of what is allowed to stand and what is blocked.

The tech companies seem to be limiting online speech much more than Congress and the courts restrict offline content. Although there are no comprehensive statistics on how much online content is being removed, the data that are emerging indicate the expansive scope of the action involved. The types of content being removed include:

Hateful content. In one typical month (September 2018) YouTube removed 94,400 videos it deemed to violate guidelines on "violent or graphic" content. Between January and June of 2018 Twitter suspended or banned more than half a million accounts for abusive or violent content. In the third quarter of 2018 Facebook took action on 15.4 million posts with violent or graphic content, meaning they removed content, put a warning screen over content, disabled offending accounts, notified law enforcement, or used some combination of these measures.

Specific examples include:

- Right-wing conspiracy theorist Alex Jones was banned from Facebook, Apple, YouTube, Twitter, and Spotify.
- Beatrix von Storch, a lawmaker with the far-right Alternative for Germany party, was blocked from Facebook and Twitter for disparaging comments against Muslims.
- Alt-right blogger Chuck Johnson was banned by Twitter for asking for donations to "take out" a political activist.
- Roger Stone, a right-wing political operative, was banned by Twitter for apparent threats against CNN hosts.
- Facebook banned the violent right-wing group Proud Boys.
- Facebook removed a misleading video about immigrants produced by Donald Trump's political team.

Note that all these voices are tolerated in other media because the United States does not ban hate speech, and political speech is considered to command an especially high level of protection.

Sexually charged content. Between July and September of 2018 YouTube took down more than 200,000 videos that violated nudity or "adult content standards"; between January and June of 2018 Twitter took action on 21,000 accounts for violating its child sexual exploitation policy;

and during the third quarter of 2018 Facebook removed more than 8.7 million pictures containing child nudity.

Specific examples include:

- Actress Chelsea Handler's nude pictures were removed from Instagram, but not Twitter.
- Mother Jordan-Lee Jones had a picture of her naked daughter playing on the beach removed from Instagram.
- Provocateur Martin Shkreli was permanently suspended from Twitter for harassing a female journalist.
- Rapper R. Kelly's music was removed from Spotify and Apple as a result of his sexual abuse charges.
- Actress Rose McGowan was temporarily suspended from Twitter after she repeatedly tweeted about the producer Harvey Weinstein's alleged sexual misconduct, including toward her. Twitter explained that McGowan's account had violated its privacy policy because one of her tweets included a private phone number.

Fake news. After the 2016 US elections and in the run-up to the 2018 midterm elections, several social media companies made a concerted effort to prevent the dissemination of false information on their platforms. As a result, the following actions were taken:

- Facebook removed more than 800 accounts associated with the production and dissemination of fake news.
- Twitter removed thousands of accounts that originated in Iran, Russia, Bangladesh, and Venezuela, which were tied to producing fake news.
- Twitter removed 2,752 troll accounts linked with Russia's Internet Research Agency.

Troubled results

No one seems satisfied with the current state of affairs. Republicans complain that tech companies' censorship practices reflect a liberal bias, as many more right-wing posts are removed than left-leaning ones. Senator Ted Cruz (R-TX) warned Facebook's chief executive, Mark Zuckerberg, during a 2018 congressional hearing that his company would lose protections provided by the Communications Decency Act if its moderators revealed political bias.

Breitbart News claims that "YouTube is purging right-wing and independent commentators in the wake of the Parkland High shooting while admitting that it is mistakenly banning conservatives," and runs headlines such as "Facebook Censors Pro-Trump Page as Company Denies Censorship Before Congress." President Trump tweeted "Social Media is totally discriminating against

Republican/Conservative voices. Speaking loudly and clearly for the Trump administration, we won't let that happen. They are closing down the opinions of many people on the RIGHT, while at the same time doing nothing to others."

According to the Pew Research Center, "A majority of Republicans say technology firms support the views of liberals over conservatives and that social media platforms censor political viewpoints.... Fully 85% of Republicans and Republican-leaning independents think it likely that social media sites intentionally censor political viewpoints, with 54% saying this is very likely. And a majority of Republicans (64%) think major technology companies as a whole support the views of liberals over conservatives."

Distrust of tech companies is not limited to the right. An increasingly common argument on the left is that AI is infected with bias. Some on the left fear it is used to perpetuate racism. Representative Alexandria

> **The most generous way to look at the current situation is to view it as a period of experimentation, with different responses to a vexing and complicated problem.**

Ocasio-Cortez (D-NY) stated that algorithms "always have these racial inequities that get translated, because algorithms are still made by human beings, and those algorithms are still pegged to basic human assumptions. They're just automated. And automated assumptions—if you don't fix the bias, then you're just automating the bias."

Finally, defenders of free speech do not think the power to censor should be placed in the hands of tech companies. Vera Eidelman, a staff attorney for the ACLU Speech, Privacy, and Technology Project, holds that "Facebook has shown us that it does a bad job of moderating 'hateful' or 'offensive' posts, even when its intentions are good. Facebook will do no better at serving as the arbiter of truth versus misinformation, and we should remain wary of its power to deprioritize certain posts or to moderate content in other ways that fall short of censorship."

The most generous way to look at the current situation is to view it as a period of experimentation, with different responses to a vexing and complicated problem. One could argue that tech companies should follow the same standards that offline platforms (print,

TV, radio) are required to adhere to, but given that a few tech companies control so much of the market (about two-thirds of people in the United States get their news on social media), they have much more power in the world of ideas and communications than any offline corporation and hence are a source of special concern.

Possible solutions

Out of the confusion and inconsistencies and changing standards, it is possible to form some preliminary conclusions. One is that the treatment of the problem is fragmented, with different parts of the content treated in different ways. Most important, although the tech companies can play a key role in regulating content, the public ought to be involved, and elected officials should have the final say in what is allowed and banned. Tech companies face a separate set of challenges because they operate in authoritarian countries such as China and Iran that have no commitment to the principle of free speech, but I am not addressing that here. The following suggested guidelines for action are intended for democratic societies:

- Tech companies should be required (if need be, by legislation) to remove any communication that is illegal offline, for instance, child pornography.
- Companies should be prevented from facilitating illegal acts such as sex trafficking and terrorism. For example, Craigslist took down certain classified pages because they were determined to contribute to sexual violence.
- Content that directly incites violence or threatens violence should be removed. For example, the likes of Alex Jones should be banned not because they are spreading right-wing conspiracy theories, but because their actions directly lead to severe harassment. After Jones called the Sandy Hook massacre in Newtown, Connecticut, a hoax "false flag" operation with "crisis actors," several families who lost children during the shooting received death threats and were forced to move multiple times.
- Continue the practice of labeling posts rather than removing them, thus allowing users to ignore the warning and click through. This seems to be a sound approach to sexually explicit content that some people might deem pornographic and others not.
- Standards for including and excluding content should be clearly stated. Criteria used by moderators and AI should be made public and be subject to congressional oversight. Public review would ensure the standards are not too sweeping and not partisan.
- Removing fake news may seem at first a rather straightforward task. For instance, it's a no-brainer to delete reports that Hillary Clinton ran a sex ring out of a Washington, DC, pizzeria. But in many cases the line between fake news and the rest is blurry. Many news stories that grossly mislead and manipulate have a small kernel of truth, but its significance is vastly overstated and misinterpreted. To protect the public from manipulation by foreign sources as well as extremist domestic ones, the best that can be done is to insist that the source be disclosed. Then the public can decide which sources to trust.
- For many reasons, people active on the internet need to have vetted IDs. For instance, this is necessary if they seek to withdraw funds or sign contracts. For the purposes at hand, if such e-IDs were widely available, tech companies could provide a forum in which only those who identify themselves will be able to post. That is, tech companies would vet the source, not the content. The internet security expert Eugene Kaspersky suggested that such e-IDs could be used for what he calls "red zones" on the internet, which include "voting in elections, online banking, interactions with official bodies, and other critical transactions." At the same time, the companies will continue to provide the opportunity for people to post anonymously, which is essential for a vibrant democracy.
- Congress should regularly review the guidelines the tech companies use and tighten them if there is compelling evidence that abuse is rampant. At the same time, Congress might demand making some standards less stringent in order to protect free speech. Currently, too much of the responsibility to control speech in cyberspace rests in private hands. In an op-ed published in the Washington Post on March 30, 2019, about government regulation of the internet, Facebook CEO Mark Zuckerberg wrote: "I believe we need a more active role for governments and regulators…. Lawmakers often tell me we have too much power over speech, and frankly I agree. I've come to believe that we shouldn't make so many important decisions about speech on our own." A combination of tech companies setting standards and Congress reviewing them seems to be a promising way to move toward a cyberspace that is less dangerous and wild, but not tamed.

Amitai Etzioni is University Professor at The George Washington University and author, most recently, of Reclaiming Patriotism.