

onezero.medium.com

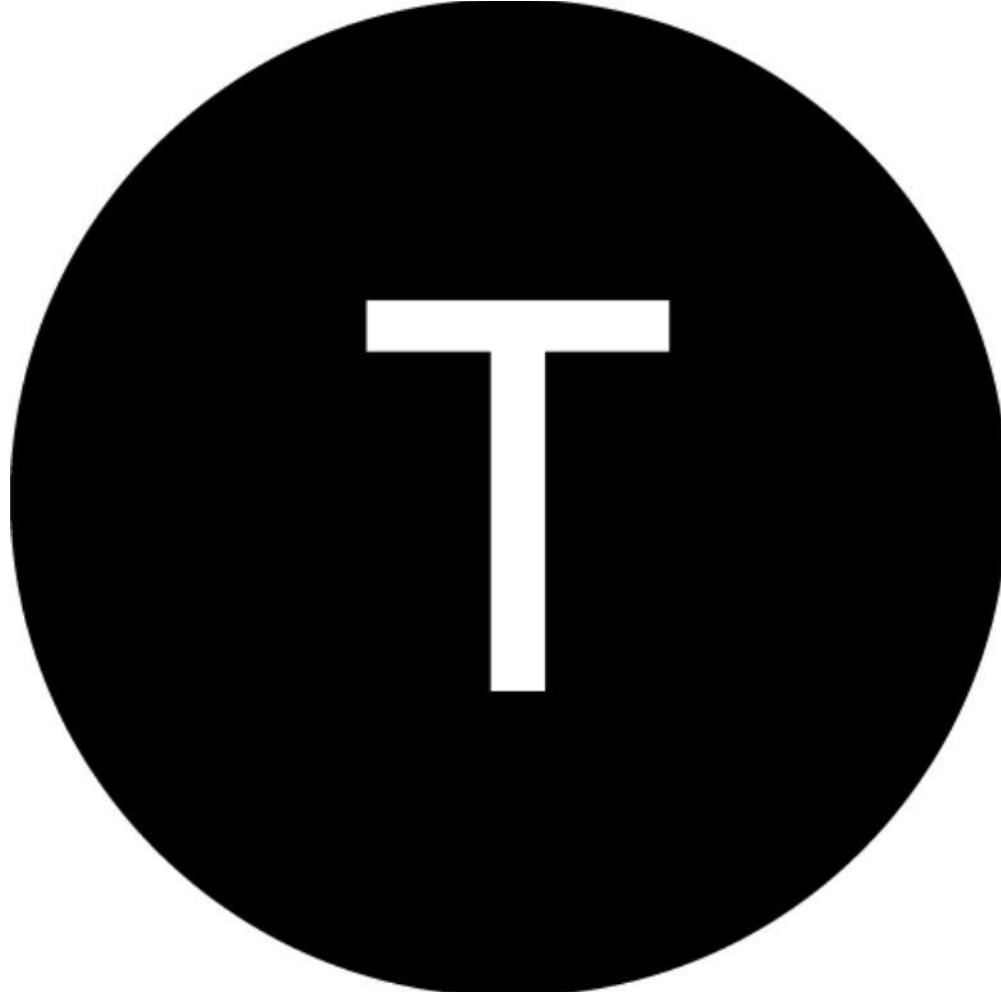
A New Constitution for Content Moderation – OneZero

David Kaye

12-15 minutes

Social media platforms should look to human rights law to govern speech on the internet





Today, a few private companies, driven to expand shareholder value, control social media. And yet the rules of speech for public space, in theory, should be made by relevant political communities, not private companies that lack democratic accountability and oversight. If left alone, the companies will gain ever greater power over expression in the public sphere.

Governments see that corporate power and are jealous of it, as they should be. French President Emmanuel Macron said directly when he appeared before the UN's Internet Governance Forum in November 2018 in Paris: "I deeply believe that it is necessary to regulate." Macron is not alone. Other democratic governments are also making explicit demands that social media companies regulate their platforms in accordance with national laws or assertions of

public security.

Authoritarian governments are taking cues from the loose regulatory talk among democracies. They are doing what they have long wanted to do—taking control of online expressive space from corporations and punishing individuals for criticism and reporting. Most authoritarian regimes will do what they want to restrain online speech, but there is a serious risk that states in what we might think of as Freedom House’s “partly free” category—transitional ones that seesaw between openness and control, that could tip into blossoming democracy or creeping authoritarianism—will borrow Macron’s language of regulation and deploy it to constrain debate and dissent.

Rebecca MacKinnon, one of the leading thinkers and activists of the digital age, [has warned](#) that internet freedom is threatened not only by authoritarians “but also by Western companies and democratically elected politicians who do not understand the global impact of their actions.” Activists and individual users struggle to have a voice in what has largely been a behind-the-scenes effort to define the rules for online expression.

Indeed, often forgotten are the users, the individuals who have grown to rely on social media for communication, commerce, and access to information of all kinds. A recent interaction brought this home to me. After a work trip to Bangkok, where I met with activists from across South and Southeast Asia, my family and I took a few days to visit Angkor Wat, the stunning complex of ancient Khmer temples in Cambodia. We hired a 40-ish Cambodian man, whom I will call S, to drive us from the heavily touristed central sites to Beng Melea, an abandoned Angkor-period temple about an hour away that jungle has consumed.

S spoke some English. He wasn't gregarious but, as we headed to the Cambodian Landmine Museum not far away, he shared with us his story of surviving the Khmer Rouge. He began with the genocide and his family's experience and ended with contemporary Cambodia, whose government has become [increasingly authoritarian and repressive of speech, media, protest, and opposition politics](#). We hadn't prompted a discussion of politics, but S went there as we passed countless posters of the prime minister, Hun Sen, telling us that everybody he knew wanted a different government. We wanted to know: In a country like Cambodia, where the media face intense restrictions and online protest can result in detention, prosecution, and harassment, how did he know this? How did he get his information?

"Facebook," he responded instantly. Facebook, he told us, is where people learn things, where they share information. Some of it is rumor, some from neighbors and friends, and some of it is reporting people get from the outside world about Cambodia. It has become, he said, the alternative to state media. Without it, he was not sure what sources he would have.

The platforms have facilitated ethnic cleansing and racist attacks against the Rohingya in Myanmar. They have permitted disinformation, even in Cambodia where there are allegations that [Hun Sen manufactured his popularity](#) on the platform. Yet Facebook in Cambodia also offers an outlet for those who desperately want to know the truth about public authorities. It's a reality that should, at least in part, shape ideas about how to promote human rights and democratic values online.

So to bring back the question at the center of the debate over online speech: Who is in charge? What are the tools available to

ensure that online speech benefits from democratic control, promoting and protecting freedom of expression, privacy of communications, rights of association and assembly, and other values of free societies? This is not an easy agenda to satisfy.

The moment calls for a serious rethink, an approach not simply to the rules governing online speech but to the public's participation in making, interpreting and enforcing them. It requires action by companies, governments, and civil society to protect speech in the digital age, not merely incremental changes at the margins of social media's management of the public square. We need new models of content moderation and public oversight, supported and promoted by rights-protecting government regulation, and with a long-term vision of public investment to sustain the infrastructure of freedom of expression in a social media age.

The following ideas sketch out the kinds of changes that would help companies and governments meet the challenges of policing content:

Decentralized decision-making

The companies are not built to moderate content at global scale. They often alienate and flatten the cultures across the markets where they operate. They have addressed problems of scale by hiring (or promising to hire) more moderators with language skills and local or regional political knowledge. This is tinkering, and while important, far from enough. Local civil society activists and users should have an explicit role in company policymaking. Wherever the companies enjoy a market presence, they should develop multi-stakeholder councils, members of which they would compensate, to

help them evaluate the hardest kinds of content problems, to evaluate emerging issues, and to dissent to the highest levels of company leadership.

Human rights standards as content moderation norms

Facebook and Twitter both made claims in 2018 that their standards are or should be rooted in the human rights of their users. Richard Allan, the lead executive in Europe, wrote that Facebook looks for “guidance” from human rights law, such as Article 19 of the International Covenant on Civil and Political Rights, one of the two central treaties of human rights law. Facebook and Google are both members of the Global Network Initiative, an effort of activists and companies to ensure that companies adhere to basic principles of protection of freedom of expression and privacy. While Twitter has refused to join GNI, the company’s CEO Jack Dorsey has said that the platform should integrate the [values of human rights in its rules](#). These are positive steps—but they are hardly even a start.

Human rights law gives companies a language to articulate their positions worldwide in ways that respect democratic norms and counter authoritarian demands

The companies should make human rights law the explicit standard underlying their content moderation and write that into their rules. They are global companies dominating public forums worldwide. International human rights law provides everyone with the right to seek, receive, and impart information and ideas of all kinds, regardless of frontiers. It protects everyone’s right to hold opinions without interference. Just as important, human rights law gives

companies a language to articulate their positions worldwide in ways that respect democratic norms and counter authoritarian demands. It is much less convincing to say to authoritarians, “We cannot take down that content because that would be inconsistent with our rules,” than it is to say, “Taking down that content would be inconsistent with the international human rights our users enjoy and to which your government is obligated to uphold.”

It is not a risk-free answer; governments may ultimately block access to their platforms. But that itself will cause risks for governments, given the popularity of the platforms among their citizens and the authoritarian signal website blocking sends to the world.

Some argue that human rights law applies only to governments and not to companies. But that is rapidly becoming an archaic way of thinking about the structure of international governance. There is a growing recognition that [corporations have responsibilities](#) not to interfere with the rights individuals enjoy, whether it is a multinational company involved in mineral extraction that helps fuel conflict or undermine worker rights, or an internet company sharing user data with an authoritarian regime.

I have also heard it argued that human rights law would permit all sorts of bad behavior that undermines user experience, such as misogynistic harassment and bullying, or that human rights principles would make it more difficult for the companies to address disinformation and, for instance, racism, anti-Semitism, Islamophobia, and homophobia. But while human rights law promotes free expression, it also permits restrictions under certain rule-of-law guidelines. Restrictions must be “provided by law”—or in the context of social media companies, have fixed rules, not

subject to their discretion, that are publicly accessible and understandable. Restrictions must be necessary and proportionate to protect the rights or reputations of others, public order or national security, or public health or morals.

The companies should explain why they require the adoption or enforcement of certain rules. It means that restricting expression should be the last resort, not the first option, particularly where there may be other tools available to deal with a perceived problem. Under these rules, companies can act to protect all of their users' rights and thus protect against activities that, for instance, seek to silence others' voices, release private information, or use the platforms' tools to incite violence.

Finally, some might say that human rights law is too general for the companies to apply. But companies have ample jurisprudence to draw from, based on court decisions interpreting and applying human rights law. This jurisprudence can be found in the European Court of Human Rights, the Inter-American Court for Human Rights, the emerging jurisprudence of regional and sub-regional courts in Africa, national courts in democratic societies, treaty bodies that monitor compliance with their norms, and the work of UN and regional human rights mechanisms. It is not an answer to say the law does not exist because some look down upon it as a lesser form of law, or because of ignorance that this body of law even exists.

That said, human rights law alone cannot fix the problems of corporate dominance. Two other tools are necessary.

Radically better transparency

Companies should open up their processes and proposals to public comment and, when they adopt new rules about content, explain clearly how they arrived at the changes. They also need to disclose to their users why they take certain kinds of content actions—what was the basis of a decision and how the user can appeal it. Clarity into algorithmic decision-making—the inputs into the A.I. that stands to control expression, not some broadly opaque math—would provide bases for individuals and academics to register serious challenges to company enforcement.

Industry-wide oversight and accountability

The second tool to maximize the value of human rights norms and local engagement is to subject company rules and decisions to industry-wide oversight and accountability. The companies should work with civil society leaders, activists, and academics to develop what Article 19 has called “social media councils.” Facebook has already tiptoed into a better appeals process, as Mark Zuckerberg has started talking about a Supreme Court for Facebook and an external appeals mechanism. Early in 2019, the company released a draft charter for a [Facebook Oversight Board](#) that could permit appeals to an independent body, with whose decisions Facebook would commit to comply. This seems promising, and Facebook has committed to consult widely as it develops this tool. (I raised some concerns about it [here](#).)