The Promise and Pitfalls of the Facebook Oversight Board
A Human Rights Perspective
The Promise and Pitfalls of the Facebook Oversight Board
A Human Rights Perspective

Carr Center for Human Rights Policy
Harvard Kennedy School, Harvard University
May 6, 2021

Flynn Coleman, Brandie Nonnecke, Elizabeth M. Renieris
Technology and Human Rights Fellows, Carr Center for Human Rights Policy

The views expressed in the Carr Center Discussion Paper Series are those of the author(s) and do not necessarily reflect those of the John F. Kennedy School of Government or of Harvard University. Faculty Research Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). Papers may be downloaded for personal use only. Cover photo by NeONBRAND on Unsplash
1. The Facebook Oversight Board recently issued its first decisions on content removals by Facebook. What are the benefits of external oversight boards for platform governance and accountability?

**FLYNN COLEMAN** It is understandable why one would look at the Facebook Oversight Board (FOB) through the rose-tinted “it’s better than nothing” lens—it may not be. Facebook (FB), an unprecedentedly wealthy global social platform and technology business, wields enormous, unchecked influence. Responding corporately to the public’s demand to curb an overwhelming spread of misinformation and hate speech, propelled by their micro-targeted algorithms, much like an autocracy might react, FB established its own “independent” oversight system.

Although the FOB, populated by well-qualified and respected members, can provide a certain measure of moral guidance to FB, it cannot escape the inherent conflict that all members are on the payroll of the conglomerate. And by its very design, the FOB cannot provide truly impartial global governance and accountability, and thereby allows FB to sidestep responsibility.

The short-term benefit of the FOB is that its existence is a tacit acknowledgement of the serious threat that FB’s scale and algorithmic architecture poses to the democratic process. It underscores the dilemma we confront in coordinating the appropriate geopolitical response necessary to formulate real solutions that will inure to the common good and not just transfer more power to the technology companies.

While any kind of meaningful oversight and internal reform can be appreciated, it is essential that we think more holistically and systematically, within a broader human rights framework, and direct the resources, brainpower, and media space presently allocated to self-policing toward grassroots organizing, legal and policy reform, and human rights advocacy to foster an international consensus on how we can build more inclusive, humane, and equitable societal infrastructure.

**BRANDIE NONNECKE** An external oversight board can serve as a robust check on a platform’s myopic viewpoint, providing alternative perspectives that can critically question the platform’s actions not through the typical lens of what is best for the bottom line, but what is best for society. External oversight boards can play an important role in guiding appropriate technology and policy strategies by enabling greater transparency and accountability.

However, in order for the benefits to be realized, external oversight boards should at minimum be:

1. **Inclusive**—to the greatest extent possible—of the diversity of stakeholders who will be affected by its decisions;
2. **Transparent** in its structure and process;
3. **Independent** from corporate pressure; and
4. **Effective** in its ability to influence a company’s practices.

These criteria are often written into the by-laws of external oversight boards, however none are easy to implement in practice.

**ELIZABETH RENIERIS** As its name would suggest, Facebook’s Oversight Board is an experiment in corporate governance that aims to provide external oversight over Facebook’s internal decisions on content. The problem is that it’s only half-baked.

Corporate governance generally involves balancing the competing interests of a variety of stakeholders that are both internal and external to a company. Key internal stakeholders with a role in corporate governance include a company’s chief executives, upper management and employees, corporate boards and board-level committees, and individual and institutional shareholders, among others. External stakeholders may include individual and enterprise customers, governments and regulatory bodies, consumer protection groups, independent auditors, the media, judiciary, and civil society, among others.

While internal corporate stakeholders will prioritize company growth and shareholder value above all other interests, external stakeholders provide balance by incentivizing corporate actors to address environmental-, social-, and governance-related (or “ESG”) concerns, the civil and human rights-related risks of a company’s products and services, and the broader societal impacts of a company’s activities. In order to do this, external stakeholders must be truly external and independent of the company’s direct or indirect influence.

Though novel in the context of content moderation-related decisions by globally dominant platforms like Facebook, external oversight is commonplace in other matters of corporate governance. For example, the federal securities laws require public companies to have their financial statements examined and reported on by external, independent auditors, and stipulate specific requirements for what constitutes “independence.” However, no such standards or requirements exist for the FOB.
2. What are the risks of external oversight boards for platform governance and accountability?

BRANDIE NONNECKE An external oversight board with inadequate scoping or lack of authority risks creating false accountability and “ethics washing” where its mere creation may turn away the scrutinizing eye of the public and policymakers and allow systemic issues to remain largely unchecked. The FOB is charged with reviewing content moderation decisions for high-profile or contentious cases of content or accounts that have been taken down. Narrowly scoping the board’s remit to a set of specific cases and limiting its authority to inform the platform’s broader strategies and policies risks undermining more impactful work to address the systemic issues creating the content moderation challenges in the first place. In order for the FOB to have a lasting effect, it should also have the ability to weigh in on the features of the platform that have led to these situations, such as its content curation and user engagement strategies (see Karen Hao’s recent article).

While Facebook sought to limit the authority of the FOB by requesting that it “focus on the outcome of enforcement, not the method,” the FOB’s recent set of rulings and especially its ruling on President Trump’s ban from the platform, make it clear that it intends for its work to have far-reaching effects beyond each particular case it reviews by informing the platforms’ policies and practices. As Evelyn Douek pointed out, the FOB has requested Facebook make its community standards and moderation tactics more transparent and that it implement an auditing process on the accuracy of its automated content moderation tools. If Facebook complies, the FOB’s scope and authority may be focusing more tightly on where it matters—the policies and corresponding methods of content moderation.

ELIZABETH RENIERIS The main risks of an external oversight board like Facebook’s Oversight Board (FOB) are that it is neither external nor capable of providing independent oversight. While FOB members do not work for Facebook, they are subject to Facebook’s internal operational challenges. For example, the FOB was delayed in getting up and running due in part to logistical matters, such as obtaining their laptops and other equipment from Facebook.

The extent to which the FOB can do its job also depends on the extent to which it can access information about how Facebook operates. As demonstrated in its recent ruling about former President Trump, Facebook refused to answer or provide evidence in response to a number of the FOB’s critical questions about the company’s handling of Trump’s account.

Moreover, unlike with respect to external oversight over a company’s financial affairs, there are no clear rules or standards for what constitutes “independence” in the context of external oversight over a platform’s content moderation decisions or for platform governance and accountability more generally.

Without such independently developed and enforced rules or standards, a company like Facebook is left to set their own standards for what counts as “independence,” and there is reason to doubt their approach. For example, FOB members are appointed and compensated by Facebook, albeit indirectly through an independent legal trust established and wholly funded by Facebook, whose trustees were also appointed by Facebook. While the individual members represent broad geographic and professional diversity, they mostly share Facebook’s singular commitment to “free expression” as a mechanism for unfettered growth. Finally, without sufficient independence or oversight capabilities, an external oversight board like the FOB risks distracting from and displacing meaningful efforts and resources devoted to platform governance and accountability, including important legal reforms and regulatory interventions. This, in turn, can provide a false sense of safety or progress, while human rights-related harms continue without redress.

FLYNN COLEMAN The primary risk of oversight boards like the FOB is that they will distract and divert us from instituting comprehensive safeguards and reforms, conceivably certifying a myriad of private, commercially motivated, and unaccountable deciders as the standard instead. Furthermore, we will not get uniformity or unanimity amongst the big firms. As of today, neither FB, nor Google, nor Twitter can agree on how or if Section 230 should be revised, as they all have separate self-interests. FB thinks that platforms can self-regulate and police content; Google suggests that any changes to the rule may produce “unintended consequences,” and Twitter believes in its process for labeling Tweets and is proposing such ideas as crowdsourcing oversight (“Birdwatch”).

As large as FB is, it represents only a single platform in the digital media ecosystem that traffics information, which includes cloud services, content networks, domain registrars, and internet service providers, all of which are intertwined and linked, and all which have varying, rival business models. Do we want to entrust the private sector with designing the rules they have to follow, while also relying on them for protection and enforcement of core human rights that are central to our democratic process and to the future of our elections? No, but we do want them responsible, like oil companies, for cleaning up their disasters.

The issues here are complex and global. The big technology companies do have outsized roles to play, starting with providing more transparency about how their businesses operate. Presently, we do not have much more than a peek inside their “black boxes”; in the past, we failed to adequately recognize how tobacco companies distorted scientific findings (“doubt is our product”), which delayed appropriate interventions. We cannot make the rules when we don’t know how the game is being played. The FOB and occasional de-platforming actions are only band-aids on a thoroughly toxic system. Giving them oversight authority without re-imagining the landscape as a whole sets a bad precedent.
3. What are alternatives to external oversight boards for platform governance and accountability?

BRANDIE NONNECKE Dozens of bills and regulatory reforms targeting social media platforms have been introduced by state legislatures (e.g., California’s Social Media Transparency and Accountability Act of 2021), Congress, the European Commission (e.g., the Digital Services Act and Digital Markets Act), and governments around the world. Platforms are well aware of the impending threat of legislation/regulation and are eager to stymie or shape the rules that will govern them.

A common tactic to avert the scorn of policymakers is to promise greater transparency. During the March 2021 congressional hearing on social media’s role in spreading dis/misinformation, CEOs Dorsey, Pichai, and Zuckerberg mentioned “transparency” a collective 13 times in their written testimonies. While transparency is necessary for appropriate platform governance and accountability, it is not sufficient. With greater transparency must come focused efforts to identify appropriate public and private sector policy and technology strategies to address the benefits and harms revealed. I believe independent public interest research is critical to achieve this.

Independent research initiatives such as Social Science One are promising, but have faced significant challenges in gaining access to platform data. This is due in part to platforms’ overly cautious interpretation of nascent data privacy legislation such as the EU GDPR and the California Consumer Privacy Act (CCPA) to restrict third-party access to users’ data (see chapter 13 in Persily and Tucker’s co-edited book, “Social Media and Democracy”). To really support transparency and accountability in platform governance, industry-wide principles and operational guides for the appropriate sharing and use of data for public interest research are needed. We’re leading some of these efforts at the CITRIS Policy Lab through the Public Interest Research Alliance.

ELIZABETH RENIERIS At this stage in the evolution of companies like Facebook, there is simply no alternative to regulation. Effective regulatory interventions will require a multi-faceted approach that includes comprehensive privacy- and data protection-related legislation, enhanced consumer protection rules, mandated corporate governance reforms (such as mandatory reporting and board composition and requirements), and certain competition- and antitrust-related measures, among other tools. Given Facebook’s global reach, effective regulatory interventions will also require cross-border coordination and cooperation, and the application of more globally-relevant normative frameworks such as international human rights law (over more US-centric standards, e.g. the First Amendment).

And even regulation on its own may not be enough, given the size and scale of Facebook’s operations and influence — particularly given the company’s hand in shaping the rules. In 2020, Facebook spent more than $20 million on lobbying Congress, more than any other individual corporation or technology company, outspending large digital platform companies like Amazon; defense contracting giants like Lockheed Martin, Boeing, and Raytheon; and powerful interests like the American Bankers Association (the company also spent a record EUR 4.5 million to lobby lawmakers in Brussels in 2020). In order to implement meaningful governance and oversight over a company of Facebook’s scale, we also have to otherwise limit its ability to influence lawmakers, including through meaningful corporate lobbying reforms.

Human rights are simply ineffective when the balance of power between sovereign nations and private corporations is as distorted as it is by an entity like Facebook. The result is a kind of epistemic coup, which scholar Shoshana Zuboff describes as a phenomenon by which oligopolistic companies “overrid[e] democratic governance with computational governance by private surveillance capital.”
FLYNN COLEMAN  Our digital platforms, networks, and communications mirror human behavior. Hate speech, propaganda, and conspiracy theories are nothing new. It is only the speed, accessibility, and scale of the delivery technology that has changed.

We need to take a hard look at the exponentially advancing designs of technology itself that preys upon the allure of the sensational and the extreme. In the near term, deepfakes will be very difficult to distinguish from reality, and deep learning language models, such as GPT-3 (that mimics human text), will be able disseminate information without attributable authorship. The genie may already be out of this bottle, but there is an opportunity to place guardrails on the technology to, at least, slow down the proliferation of harmful content while we come up with solutions for this new age.

In the short term, we must demand transparency and compliance with existing regulations and encourage the Department of Justice and state attorneys general to pursue the purveyors of dangerous lies, much like some are doing to stop the spread of COVID misinformation. We must also form global alliances to collectively develop ways to better patrol our borderless cyber-communities, bar bad actors, and encourage the tech companies to ensure that their own Internal Review Boards aspire to more rigorous standards. In the longer term, we make digital literacy a core component of early and higher education, recognizing that we cannot cede the moral high ground to algorithms. We also train public interest technologists, collaborating across disciplines and at the intersections of democracy, social science, and technology; and we explore such ideas as independent citizen assemblies to help review policies. The generations who have grown up with artificially intelligent components of early and higher education, recognizing that we cannot cede the moral high ground to algorithms. We also train public interest technologists, collaborating across disciplines and at the intersections of democracy, social science, and technology; and we explore such ideas as independent citizen assemblies to help review policies. The generations who have grown up with artificially intelligent technologies need to be ethically prepared for the world we are leaving to them.

The farther-reaching alternative to sanctioning corporate, extrajudicial mechanisms is to find the political will to create bold, meaningful public policy legislation and to more effectively apply the laws already on the books. We start by using antitrust principles to mandate diluting the concentrated power of the technology behemoths by breaking them into smaller, more governable, less dominant companies, like we did with Standard Oil and Ma Bell to good effect, creating healthier competition and improved service.

It is also feasible that emerging decentralized technologies, such as blockchain, will change the social media landscape, preempting regulation and internal reform in ways we cannot yet forecast.

4. What is the role of human rights in platform governance and accountability?

BRANDIE NONNECKE  In its rulings, the FOB often applied international human rights standards to justify its positions—and for good reason. International human rights standards provide a global framework of human values that can be used to evaluate and shape ethical business practices. For example, the UN Guiding Principles on Business and Human Rights, which draw upon internationally recognized human rights principles contained within the Universal Declaration of Human Rights and corresponding instruments such as the International Covenant on Civil and Political Rights, provide useful guidance to companies on how to identify and mitigate human rights risks.

In a recent Carr Center report by Mark Latonero and Aaina Agarwal, the authors explore the increasing application of human rights impact assessments (HIAs) to identify and mitigate human rights risks of AI. In their analysis, they cite the application of an HRI to evaluate Facebook’s role in fueling genocide in Myanmar. Their findings are telling. When applied appropriately, HIAs can be extremely effective at identifying and mitigating human rights harms and enabling remedies. If applied inappropriately, HIAs can perform “ethics washing” and diminish an individual’s rights through dismissal of corporate culpability.

Platforms should use international human rights standards as a means to better understand and mitigate harms of their products and services. We’re witnessing a rise in the use of HIAs among platforms, which is promising if done appropriately, but extremely problematic to the human rights they seek to support if done inappropriately.

ELIZABETH RENIERIS  Although Facebook is a U.S. company, its operations and impact are global with its platform offerings, now reaching more than 3 billion users worldwide combined across its Facebook, Instagram, WhatsApp, and Messenger platforms. Despite the different legal and normative frameworks in many geographic regions where it operates, Facebook’s activities demonstrate a clear bias for U.S. norms and values, in part pander to regulators at home.

This is especially pronounced in the context of speech and expression, which Facebook has largely interpreted through an American lens vis-à-vis the First Amendment of the U.S. Constitution. Whereas the First Amendment is largely focused on the speaker of speech or the right to impart information, free expression under international human rights law is much broader and includes the right to not only impart, but also to receive information.
Nevertheless, Facebook’s algorithmic amplification of certain content, as well as the proliferation of mis- and disinformation, harassment, and hate speech on its platforms, inhibit the rights of many of its users to freely impart and receive information, with particularly outsized impact on minority and marginalized groups. As such, a more expansive and globally relevant normative framework is necessary.

While the company recently published its first official Corporate Human Rights Policy, genuine adherence to substantive and procedural human rights standards will require Facebook to undertake meaningful human rights impact assessments of its products and services, implement robust human rights due diligence throughout the company, add effective board-level human rights expertise and oversight, and introduce a variety of other measures to give it teeth.

FLYNN COLEMAN There are “three fundamental human rights—freedom of expression, safety, and privacy.” Guess who said this? Twitter CEO Jack Dorsey. I don’t doubt the sincerity of some of those trying to create oversight and reform from within, but we also already have an excellent global human rights framework to interpret, and those same rights should accrue to our lives online and be applicable to Big Tech.

International human rights law revolves around an agreement that all human beings are equally worthy and inherently valuable, and a consensus of sovereign states is a key part of this understanding. This accord is foundational to protecting people’s agency and dignity and is the existing axis around which we should compose an ecosystem of rights and responsibilities applicable to our information technologies. Free speech, at the forefront of the governance and accountability debate, is not revered and defended universally in the “U.S. sense” of the term; whereas freedom of opinion and expression, as set forth in Article 19 of the Universal Declaration of Human Rights, is generally more widely embraced. This language provides a better roadmap.

The question is how can we best enshrine the same rights in the digital sphere, now that the internet has grouped us together “like neurons in a giant brain” (Stephen Hawking). For those fortunate enough to have access, the global public square is open 24/7, and we should no more privatize the whole of that sacred space than we can privatize democracy.

Although Big Tech has an important role to play in regulation, the acts of labeling, fact-checking, or censoring alone cannot solve the divisions among us, nor will they stop the spread of dis- and misinformation. It’s about trust and belonging; and until we can construct forums of engagement that encourage humility, accommodate tolerance, and foster genuine, open dialogue, our best collective future will not be reflected in these spaces.