Facebook's Oversight Board Just Announced Its First Cases, But It Already Needs An Overhaul

Dipayan Ghosh & Justin Hendrix

19.12.2020 doi: 10.17176/20201219-172911-0

On the 1st of December, Facebook Director of Governance and Global Affairs Brent Harris published a blog post announcing the first cases the newly constituted Oversight Board will consider. The announcement was a long time coming; it took Facebook two years to develop the Board, and since, both the company and the Board have been criticized for its slow progress. Indeed, in the interim, a rival group calling itself the "Real Facebook Oversight Board" formed in protest. The "real" board is concerned that the official Oversight Board lacks true independence from Facebook. This criticism is rooted in bylaws that appear to leave Facebook firmly in control of which cases the Board will take up.

The Oversight Board was never going to be a panacea for the complex problem of content moderation on a platform that hosts billions of users, but it is clear already that the Board's governance model requires an overhaul if it is to achieve meaningful success.

Background

More than two years ago, chief executive Mark Zuckerberg published a Facebook post titled "Blueprint for Content Governance and Enforcement," recognizing the need for "independent oversight and transparency into our systems" and a more "rigorous policy-making process." Several months later, in September 2019, Facebook announced it would launch an Oversight Board to handle questions about how to deal with problematic content, to operate for at least years, funded by $130 million. Earlier this year, after much attention on the company's decision-making over the Board's structure and membership, Facebook finally announced the first twenty Board members, which included "a former prime minister of Denmark, a former European Court of Human Rights judge, and a Nobel Peace Prize laureate," amongst other luminaries and experts.

The Board has two primary functions. The first is to take up cases of potential offending content and adjudicate whether or not that content should stay on Facebook. Such judgments are intended to help develop jurisprudential guideposts on content policy and moderation for the company. Like a supreme court, the Board's judgments are binding. The second is to openly discuss within the Board policy priorities the Board believes Facebook should undertake, and deliver related guidance to the company to inform the corporate policy making process. This guidance is not binding.

The Cases

A somewhat common throughline of the initial slate of six cases before the Board is a fundamental mismatch between the vagaries of user intent and what an image or text more literally signifies:

- A case regarding the removal of a comment made by a public figure on a post about violence against French people.
- A photo of a dead child alongside commentary on China's treatment of Uyghur Muslims.
- A set of photos purportedly showing churches in Baku with a caption indicating disdain for Azerbaijani people and support for Armenia in the Nagorno-Karabakh conflict.
- A post of photos with female nudity (which is banned by Facebook) related to breast cancer symptoms.
- A post containing an alleged quote from Joseph Goebbels, the Reich Minister of Propaganda in Nazi Germany.

- A post in a Group claiming hydroxychloroquine and azithromycin are a cure for COVID-19 and criticizing France's response to COVID-19.

Intriguingly, after the announcement of these cases, one user apparently deleted the content in question, regarding violence against French people. As a result, Facebook announced that "this case has been deleted from our platform and from the Board's systems, which means the Board will no longer be able to view or issue a decision on this case." The Board announced as a replacement a 7th case:

- A photo depicting what appears to suggest violence against Indian Muslims, depicts President Emmanuel Macron of France as "the devil," and calls for the boycott of French products.

It is worth noting that the circumstances around the replacement of a case resulting from user deletion of content raises novel concerns around the influence that individual users could have on the deliberations of the Oversight Board.

Criticism

Critics of the Board and its selections have not been shy. While Columbia University professor Emily Bell noted that perhaps "the most striking feature of the Board's first set of cases is the lack of ambition in their subject matter," a key concern is that no matter what it decides in these exemplary cases, the Board will not address the issue of Facebook's disclosure of its methods and reasoning on content moderation decisions more broadly. "The OSB is an attempt to address some of these systemic shortcomings and inconsistencies, though it does not correct the fundamental lack of transparency," Danish human-rights advocate Jacob McHangama wrote . Likewise, Dia Kayyali, associate director of advocacy at Mnemonic, a human rights group, is "pretty concerned that ultimately we are not going to have a good sense of what information they're basing the decisions on". The Board has chosen in this first tranche a handful of cases from tens of thousands submitted to it, representing a fraction of a fraction of the billions of content moderation choices Facebook is making on a regular basis, algorithmically or by direct human intervention.

This is indeed the fundamental problem with the Oversight Board; it has the potential to accomplish far more than we would project it will. Facebook's well over 2 billion users collectively generate billions of instances of content every day. An especially high ratio of the content that ultimately achieves virality, is questionable and potentially offending to many people. As scholars at MIT have established, falsehoods shared on social media spread much faster and farther than the truth, and we can likely extend that analogy to other forms of offending content.

Long-term, we can recognize that the Board, operating much like a private, platform-specific supreme court of content policy, can help direct Facebook's content moderation policies toward a socially acceptable outcome, just as the European Court of Justice or US Supreme Court establishes precedent. But how much closer are we to that outcome with six cases heard after many months, on a set of issues that does not appear to reflect on the core, fundamental harms perpetuated over and facilitated by the company's platforms – including hateful conduct, incitement to violence, the spread of misinformation, the coordinated disinformation problem, election interference, exploitative content and more?

The concerns over the Board's efficacy are not necessarily irresolvable. Facebook might work with the Board to harness its members' expertise on wider issues that actually concern the public. For instance, the Board could deliver judgements on contentious pieces of content that appear during important election cycles. However, its role as it stands is far too narrow for the Board to be able to have a corrective impact on the company in the long run. We need a Board that is empowered to take on the company's decisions around dis- and misinformation, misleading political speech, and contentious incitement in real time for it to have a truly positive impact. Meanwhile, the Board's current function – which is to deliver judgements on vanishingly few, relatively insignificant instances of content every few months – is not additive to Facebook's existing capacities. Over the past fifteen

years, the company was already taking this function on internally, and had already established policies and protocols to gather public feedback as it moved forward with these judgements to establish a global company-wide approach. Now it is the Oversight Board that takes on the accountability for content decisions, not the company directly.

This bleeds into another frustration many have with the company: The Facebook Oversight Board does not offer general "Oversight" over the company. It offers narrow oversight to a few, rather uncontentious instances of content that will overall have no substantial impact on the company's innovation on questions of content moderation. It is misleading to the public to regard it as a general Oversight Board; it should be appropriately titled to reflect what it is – a Content Moderation Advisory Committee. Indeed, this is akin to what TikTok has done already.

Strategic changes - a path to success?

We are not keen only to point out the gaps in the board's governance and structure – we think there is a path to success, and we would suggest some strategic changes so as to make the board's work meaningful and impactful on the company's operation and thus the general public.

*Create a system by which the contentious user-generated content that the public actually cares about is tackled.*The six initial cases do not necessarily address what the general, global public cares about the most. There is a necessary process in whittling down what the Board can consider, due to its capacity constraints. But are the six cases before the Board really the six that the public wishes for a public judgement on more than the other tens of billions of instances of content? Facebook should instead provide the Board some mechanism to determine what the public desires judgement on, by developing a system to infer public attention to an issue and a piece of content, and suggest that the Board tackle the content that rises to the top of such an ordering.

Establish algorithmic transparency measures for content moderation. The vast number of content moderation decisions never pass under a human eye. On some level, if the Oversight Board is to answer its mission it will need to publish code. Or, if not code, at least some representation of the logic of Facebook's content moderation systems. In order to understand how the company decides what sees the light of day and what does not, the world needs to see the algorithmic decision-making processes, and alterations that are made as a result of the Board's deliberations. Until the public – or, in its stead, interested regulatory authorities, researchers and civil society – can understand and comment on how the code ties to the decisions taken under this new form of governance, it will not be entirely apparent how the Oversight Board's decisions affect the handling of moderation decisions on the site. Indeed, understanding the relationship between automated and human content moderation was a focus of an external Data Transparency Advisory Group that issued a report last year.

Establish a binding structure for feedback on the company's general policy process. Facebook has created a system by which the Board can suggest policy feedback on any issue to help inform its approach. However, these suggestions – unlike the Board's judgements – are non-binding. The company and Board should consider ways to either make them binding, or if that is not possible, generate accountability on the company so that it must take into serious consideration the perspectives of the Board. For instance, there could be an internal corporate stipulation that each policy recommendation the Board makes will be openly discussed at the next all-hands meeting, or that there will be ample opportunity for the company's ranks or the public to engage directly with the company's executive content policy leadership to discuss the Board's policy perspectives. The company could also establish mechanisms to gauge public support of the Board's policy recommendations, and a corporate commitment that it will enact such recommendations pending an internal corporate ballot procedure or other determinative process.

Resolve the capacity issue. One of the key concerns that the public and Facebook alike are likely to raise is the Board's capacity issue. Overseeing a company as dominant over the media ecosystem as Facebook is

both a difficult challenge and time-consuming work. If (or as) capacity constraints become genuine obstacles to the Board's progress, more expertise should be brought into the composition of the Board itself, or perhaps even more critical, it is necessary to hire more staff to work for the Board, depending on where the bottlenecks lie.

Absent these much-needed changes to restructure the Board and thus heighten its impact on Facebook, the media ecosystem, and the public, we should recognize the Board for what it is, driven by the function it currently serves – one of corporate propaganda.

ABOUT THE AUTHORS

Dipayan Ghosh
Dipayan Ghosh is co-director of the Digital Platforms & Democracy Project at the Mossavar-Rahmani Center for Business and Government at the Harvard Kennedy School, and author of Terms of Disservice (2020). He was a technology and economic policy advisor in the Obama White House, and formerly served as an advisor on privacy and public policy issues at Facebook.
Justin Hendrix
Justin Hendrix works on problems at the intersection of technology and democracy. He is a cofounder of Tech Policy Press, a startup nonprofit media venture, and an associate research scientist and adjunct professor at NYU Tandon School of Engineering.

SUGGESTED CITATION

Ghosh, Dipayan; Hendrix, Justin: Facebook's Oversight Board Just Announced Its First Cases, But It Already Needs An Overhaul, Verfassungsblog, 19.12.2020, doi: 10.17176/20201219-172911-0.

Licensed under CC BY-SA 4.0.

Langzeitarchiviert bei IntRechtDok unter: doi.org/10.17176/20201219-172911-0.