

Ranking Digital Rights' response to Facebook on the Oversight Board bylaws, trust, and human rights review

January 28, 2020

Today, Facebook released the highly anticipated [bylaws](#) for its planned Oversight Board for content decisions. We at RDR think this experiment in internet governance shows real progress toward building new models of content moderation that protect and promote human rights. At the same time, it reveals that much work remains to be done for these models to succeed and endure.

This document contains our full response to the company's plans for the Oversight Board to-date, along with our comments on a third-party human rights review conducted by [Business for Social Responsibility](#) (BSR).

Facebook's plans for the Oversight Board respond to a pressing need for increased transparency about the company's policies and practices pertaining to the human rights of its users. This need was clearly reflected in the [2019 RDR Corporate Accountability Index](#), published in May 2019. The 2019 RDR Index revealed critical gaps in Facebook's [grievance and remedy](#) mechanisms, which [remained among the weakest of any ranked company](#). RDR also [called on](#) Facebook to improve, expand, and disclose more information about its human rights due diligence mechanisms, including human rights impact assessments on its products and services, enforcement of terms of service, and targeted advertising policies and practices.

We welcome several elements of Facebook's most recent updates on the development of the Oversight Board, including:

- **Creating an independent Trust and disclosing the Trust's foundational documents.** Since the [very first public consultations](#) on the Oversight Board and [other multi-stakeholder models](#), RDR has been a strong and consistent proponent of independent governance structures for the Oversight Board, advancing the idea of a trust as a means for it to achieve financial autonomy and fulfill its fiduciary duties to Facebook's users.
- **Commissioning and publicly disseminating BSR's independent human rights review.** In particular, we commend Facebook's proactive decision to release the assessment prior to the launch of the Oversight Board. This sets an example for other tech companies to follow. Public disclosure of in-house and third-party human rights due diligence is sorely lacking in the tech field. The 2019 RDR Index found that most companies showed [next to no evidence](#) of conducting regular, comprehensive, and credible due diligence. Even fewer companies publish the results of their assessments. This endemic opacity can be countered with publications such as this one. Along with Facebook's [previous human rights impact assessment on Facebook in Myanmar](#), this document shows further commitment to expanding accountability and visualizing internal

due diligence efforts for the benefit of the public. These disclosures are exceptions in the industry: the methods and processes used to conduct such assessments are [almost never revealed](#) (assuming they are conducted at all) – and when they are, information about their nature, scope, and process is rarely shared with the public.

- **Disclosing clear timeframes for remedy and other procedural aspects.** The 2019 RDR Index found that Facebook was [not fully transparent](#) about the timeframes of their content review processes. This is not uncommon — many other companies provide only ballpark estimates or nothing at all. We are glad that Facebook has taken note of these concerns in the course of planning the Oversight Board. We commend Facebook for integrating clear timeframes across most aspects of the Oversight Board’s operations (e.g., decision, implementation, notification, public response) into the Bylaws.
- **Incorporating an advisory role on policy development for the Oversight Board.** In order to remain dynamic and relevant, the Oversight Board must have some authority to provide advice on Facebook’s content policy, in addition to its primary role of making binding decisions on content. We appreciate that the Bylaws explicitly confer this authority, in the form of “advisory policy statements,” to the Oversight Board. Nonetheless, Facebook’s approach to this could be improved (see Recommendations).

These steps in the direction of tangible transparency should be emulated and advanced in future efforts.

In addition, we have several comments and recommendations regarding Facebook’s recent updates and further recommendations for the Oversight Board’s creation.

BSR Human Rights Review

RDR agrees with several of the broader concerns and recommendations put forth in BSR’s Human Rights Review of the Oversight Board. A non-exhaustive list of these areas of overlap includes:

- **Anchoring in human rights.** Facebook’s content policies and values, as defined in the Oversight Board’s [Charter](#), are simply insufficient as a foundation for the Oversight Board’s operations. Universal standards are paramount. We are glad to note that BSR recommends anchoring Facebook’s Oversight Board in the international human rights framework on both the substantive (e.g., content decisions) and procedural (e.g., grievance mechanisms) side, in addition to referencing other international conventions. We strongly support this focus as a foundation for the Oversight Board. **As it stands, the Bylaws do not contain an explicit commitment to using universal human rights principles as a foundation or core component of the review process.** In fact, although the Bylaws include a commendable and previously unincorporated commitment to include human rights analyses in the Oversight Board’s annual report, neither the Bylaws nor the Charter refer to the UN Guiding Principles on Business and Human

Rights (UNGPs) or discuss exactly how human rights will be built into the operations of this new entity. We recognize the progress from the draft Charter, which makes no reference to human rights norms, to the final Charter and Bylaws, where the Oversight Board's decision-making process includes assessing the impact of content removal on the right to free expression. But this neither covers the full spectrum of human rights nor equates to accepting human rights principles as the cornerstone. BSR's Human Rights Review clearly demonstrates how the UNGPs can be used as a practical roadmap for Oversight Board members and be operationalized to support its work. We also welcome the recommendation to apply the [Global Network Initiative \(GNI\) Principles](#) and [Implementation Guidelines](#) in the Oversight Board's operations and the use of human rights language in the Charter.

- **Consideration for secondary, cumulative, and downstream harms.** We welcome BSR's recommendation to consider the full spectrum of possible harms stemming from content-related decisions. RDR believes that it is imperative to go beyond direct harms; this is explicitly stated in the UNGPs and widely accepted in best practices on due diligence. We also support adopting severity criteria in line with the UNGPs and BSR's recommendations, and welcome the specific examples provided in BSR's Review. Particularly notable is the inclusion of individuals who are not users of Facebook or Instagram. The precipitous spread of hate speech and disinformation on Facebook-owned platforms has had cascading effects in several countries, [directly contributing](#) to violence and loss of life in Myanmar, Sri Lanka, and India, among others. The victims of these information cascades did not necessarily maintain a presence on the platforms involved, but were certainly rightsholders who should have been much better protected. Indeed, their lack of awareness of targeted campaigns against them often rendered them more vulnerable to the harms that ultimately befell them. Such cases provide the best proof that scoping for harms should go well beyond those who are already connected.
- **Scope of authority and its evolution.** We strongly agree with BSR's assessment that the Oversight Board should have the authority to recommend changes to Facebook's Community Standards as well as to the processes and mechanisms for enforcing them, as we have expressed in [previous commentary](#). We believe that all such recommendations must be made public. We are glad to see a recommendation on tackling novel cases and emerging trends, which will contribute to a more dynamic due diligence process within Facebook itself. We also support the future gradual expansion of the Oversight Board's scope to encompass algorithmic policies and practices, changes to the visibility of content, and the rights of non-users potentially affected by content on Facebook and Instagram (see Section 3 of the BSR review). We feel it is especially important to introduce additional oversight to algorithmic content curation systems and targeted advertising. In the interest of transparency, it is vital that any alterations to the scope of the Oversight Board's work be publicly disclosed and clearly communicated in a manner similar to Facebook's [existing mechanism](#) to track changes to its Community Standards. It is reassuring that the Bylaws outline several content

types (e.g., groups, events) and decision types (e.g., content rated “false” by fact-checkers) that will be within the Oversight Board’s purview in the future.

- **Engagement of a “user advocate.”** BSR recommends that a “user advocate” should be provided to support users who wish to make their case to the Oversight Board. We agree with this recommendation. [Research and reporting](#) on the roles and advantages of ombudsperson institutions in fostering open government culture has demonstrated that these entities are an important channel of access for citizens. Similarly, the proactive participation of a user advocate in the surfacing and appeals process would humanize an experience that, for many users, would otherwise remain a black box - regardless of other efforts to equip them with resources. A user advocate would transform the passive provision of information into active engagement.

Recommendations

Transparency: Publish comprehensive data about the Oversight Board’s decisions and implement systemic transparency.

In our [feedback](#) on the Draft Charter for the Oversight Board (May 2019), we called for data on appeals that are escalated to the Oversight Board to be released on a regular basis, drawing on a clear prior commitment from Facebook to do so. Such data should be made public to the extent that this does not violate the right to privacy of the user(s) involved and/or pose a threat to other fundamental human rights. To that end, we welcome Facebook’s intention to create a **public repository of cases and decisions** made by the Oversight Board. We recommend that this repository be as comprehensive as possible, depending on the volume of cases.

The categories of data that we expect to be included in this reporting remain unchanged. **We are glad to see that several of our recommended standards on this issue have been built into the Bylaws**, including the number and type of cases heard, summary of responses, and breakdowns by region, source of referral, and platform. We strongly agree with Facebook’s initiative to include additional information in the annual report, namely an analysis of how the Oversight Board’s decisions have considered the specific human rights implicated by a case and summary reports on the timeliness of Facebook’s response to the Oversight Board’s decisions. The remaining categories of data that we expect to be disclosed are:

- A breakdown of cases by **which Community Standard they violated** or were related to.
- A breakdown of cases by **format or content** at issue (e.g. text, image, video, livestream).
- The **number of accounts and pieces of content** covered by the cases considered by the Oversight Board (some cases may involve multiple accounts).
- The number of accounts and pieces of content that were **taken down or otherwise actioned** as a result of the Oversight Board’s decision in each of the cases it has considered.

We call on Facebook to clarify the steps that the Oversight Board will take to **protect the privacy of the user(s) whose content will be reviewed through this mechanism**. We are glad to see that the Bylaws require the Oversight Board to seek the consent of users before releasing personally identifiable information. We also welcome the decision to conduct a legal review within three days if such consent is not provided. However, it remains to be seen how such decisions will be disclosed publicly in cases where consent is not given and the Oversight Board decides to retain or restore the content in question.

We also invite Facebook to **embed transparency into other parts of the process**.

Transparency is much more than a buzzword; Facebook should seize the opportunity to set a standard and demonstrate to other actors in the industry how transparency can be furthered through concrete actions. BSR's Human Rights Review, for instance, discusses the process of collecting statements from potentially affected vulnerable groups (p. 44). Given the importance of this step for human rights impact assessments anywhere in the world, the procedural guidelines for the collection of testimonies should be transparent, easy to understand, and publicly available. We recommend that Facebook create a **repository of documents** to make their commitment to transparent processes more tangible, having already set the precedent with the release of the materials pertaining to the independent Trust.

Due diligence: [Specify the relationship between the due diligence mechanisms of the Oversight Board and those used by Facebook.](#)

RDR believes that the Oversight Board should have the authority to provide input on human rights impact assessments (HRIAs) pertaining to broad policy changes. This should apply particularly when these impact assessments are conducted on overarching issues such as removing or downranking broad categories of content. As its scope and mandate expands (as envisioned in BSR's Human Rights Review), it should extend to providing advisory opinions on the HRIA process itself. While HRIAs within Facebook are governed by a separate process and don't usually pertain to individual pieces of content, it will be important to clarify any planned coordination activities between the HRIA teams and the Oversight Board. Facebook should make clear whether there will be ongoing communication between the Oversight Board and Facebook's human rights teams in general—and if so, how it will be structured.

On a broader level, the Bylaws confirm that the Oversight Board would have an additional advisory role on the company's policy, as specified above. However, they imply (e.g., section 2.1) that it will only be able to issue these statements when called upon by Facebook. **We strongly believe that the Oversight Board should have the authority to present advisory opinions on content policy to Facebook of its own accord, without prior solicitation by the company.**

Decision-making options: Clarify the decisions available to the Oversight Board on individual pieces of content.

Neither the Bylaws nor BSR's Human Rights Review explicitly state what the options available to the Oversight Board would be for any given piece of content. At a minimum, we understand that there are three basic options: keep up, take down, or demote.

However, it is unclear from the final Charter, the Bylaws, and the Human Rights Review whether the Oversight Board's authority will extend beyond the binary decision of removing content or keeping it up. Even demotion (reducing visibility) is not explicitly considered as an option, let alone additional measures such as tagging and quarantining content. Indeed, even demotion itself is not a binary decision, as the visibility of content can be reduced to different degrees and the reasons for demotion may vary widely. Both suppressing a piece of content and allowing unrestricted access to it entail human rights risks, depending on the circumstances. Thus, procedurally speaking, it is critical to both make the Oversight Board's decision-making options explicit and explain what happens to a piece of content once an appeal reaches it..

Autonomy of Oversight Board and Trust: Clarify remaining sources of uncertainty regarding the operations of the independent Trust and Facebook's ability to influence the Oversight Board.

We believe the Trust is a vital ingredient for the Oversight Board's success. However, there are limits to the Trust's independence given that Facebook will be responsible for appointing the trustees. This is a concern that extends to the overall structure of the Oversight Board, as the current process involves Facebook itself selecting the first co-chairs (officers) and then collaborating with them to identify and recruit members. This carries the risk of percolating bias.

In the case of the Trust, the Bylaws give the trustees the power to approve or reject candidates, but do not identify specific criteria trustees should use to make these decisions. It remains uncertain whether a simple majority of trustees would be sufficient to reject a candidate. Similarly, the trustees must have clear grounds for rejection, which the Bylaws do not require them to present. Furthermore, the Bylaws state that the trustees will process requests to remove Oversight Board members from the Oversight Board itself, its Director, or "the public." It is imperative that Facebook clarify what "the public" means here and what the process will be for hearing such requests.

Finally, the Bylaws indicate that Facebook would have the prerogative to amend a section (titled 'People') that includes the case submission procedure for users and content types that fall under the Oversight Board's authority. Unlike the bylaws for another section ('Facebook'), these amendments would not require any form of approval from the Oversight Board. **Given that amendments to this section could reduce the scope of the Oversight Board's operations and the content it can tackle, we strongly recommend that Facebook seek concurrence before implementing such changes.**

Role of governments: Clarify the role of governments and the Oversight Board’s ability to provide transparency and accountability around government orders that threaten human rights.

In our view, governments should never be able to submit nominations for Oversight Board membership or submit cases for review. The risk of government capture has been a fixture of debates surrounding multistakeholder models of content governance. It is important to keep this in mind as the Oversight Board evolves. In addition, as per Facebook’s Community Standards, BSR’s Human Rights Review, and the Bylaws, the Oversight Board will take up “no cases in which reversing Facebook’s decision would violate the law.” Because the restrictiveness of laws varies widely by regime and circumstance, it is critical to explain what steps Facebook or the Oversight Board would take in cases where discriminatory or harmful practices are embedded or tacitly endorsed in national law—for instance, through advisory opinions.

RDR understands that transparency around cases involving national law is often an intractable problem due to companies’ duty of compliance and potential threats from repressive governments. However, increased transparency about specific cases that the Oversight Board is not able to address and why it is not able to address them is likely to have a tangible, positive effect. Facebook and the Oversight Board (once constituted) should jointly consider how the Oversight Board could address government censorship practices in a manner that would protect the safety and security of Facebook’s employees, Oversight Board members, and the users of social networking services owned by Facebook.

Conclusion

Like all digital platforms serving an enormous user base, Facebook has a fiduciary duty to its users and their environment—a responsibility to act in their best interest, to serve as a responsible steward of their data, and to provide them with options whenever their human rights are at play. If properly implemented, the planned Oversight Board could represent a paradigm shift in the direction of meaningful transparency and the protection of users’ rights in ways that are independent from companies’ operations. We welcome this opportunity to give input on ways we believe will ensure this mechanism can become a successful model for the future.

Ranking Digital Rights is a nonprofit research initiative that works to promote freedom of expression and privacy on the internet by creating global standards and incentives for companies to respect and protect users’ rights. We do this by ranking the world’s most powerful internet, mobile ecosystem, and telecommunications companies—including [Facebook](#)—on relevant commitments, policies, and practices, based on international human rights standards.