

Mandate of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression

REFERENCE:
OL OTH 24/2019

1 May 2019

Dear Mr. Zuckerberg:

I am writing to provide preliminary reactions to **Facebook’s initiative to create an Oversight Board for Content Decisions** (“the Board”). I welcome any genuine attempt to enhance accountability and oversight of the content moderation policies of social media platforms, and I have been pleased to engage with Facebook on these and other issues in recent years. I am especially gratified by Facebook’s open comment process concerning the Board and, given my confidence in the professionalism and commitment of your content policy team, I am certain that company decisions will take into account stakeholder equities.

For background, it is my responsibility as the UN Special Rapporteur on freedom of opinion and expression, under Human Rights Council resolution 34/18, to evaluate how governments, non-state actors and companies protect and promote everyone’s right to seek, receive and impart information and ideas worldwide. I report to the UN Human Rights Council and the General Assembly, conduct official country missions, and communicate regularly with governments, civil society and private industry. Digital rights lie at the center of much of this work, with my formal reporting to the UN often focusing on the obligations of governments to ensure protection of rights online and the concomitant responsibilities of companies.

To its credit, Facebook has engaged closely with human rights mechanisms worldwide in recent years, particularly those associated with the Human Rights Council. Richard Allan promisingly stated last summer that Facebook “looks for guidance in documents like Article 19 of the International Covenant on Civil and Political Rights (ICCPR), which sets standards for when it’s appropriate to place restrictions on freedom of expression.” This is an important recognition and I do not intend the comments below to detract from my appreciation for the company’s engagement or its move towards human rights framing in its public pronouncements.

From a human rights perspective, oversight is but one mechanism – albeit a critically important one – to ensure that social media platforms, in the words of the UN Guiding Principles on Business and Human Rights, “avoid infringing on the human rights of others and . . . address adverse human rights impacts with which they are involved.” (A/HRC/17/31, Principle 11, hereinafter “UN Guiding Principles”) Oversight can help refine *internal* content policies and ensure their consistent implementation. But oversight should also, in my view, reinforce company responsibilities to protect human rights. Indeed, Facebook is not ‘just another company’. Its global position in the information space is unprecedented. With that position, and that power, comes enormous impact – and unusual responsibility to get it right, to protect individuals, to avoid undermining

public institutions, to innovate. Rooting Facebook’s content decisions, the Board’s review of these decisions, and the structure of that process in international human rights standards, such as those reflected in the ICCPR, is critical to discharging this responsibility. Governments have an obvious role to play in ensuring company compliance, but that does not obviate company policies and practices as the first line of defense and protection.

With that in mind, I would like to share the following preliminary reactions to the draft charter and the Board:

A. Standards

The draft charter indicates that the Board would oversee appeals of Facebook’s content decisions. It indicates that the standards of review will be premised on a “set of values” that “encompass concepts like voice, safety, equity, dignity, equality and privacy.” I have no quarrel with these values, but they are evidently reflective of Facebook’s Community Standards, and suggest that these standards will be the primary framework for the Board’s review of content decisions.

I would strongly encourage that the Board’s review standards integrate international human rights law, the absence of which is concerning. Article 19(2) of the ICCPR states that “[e]veryone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice.” Article 19(3) recognizes that freedom of expression “may . . . be subject to certain restrictions, but these shall only be such as are provided by law and are necessary . . . (a) [f]or respect of the rights or reputations of others; [or] (b) [f]or the protection of national security or of public order (*ordre public*), or of public health or morals.” These global standards are also reflected in European, inter-American and African legal instruments and jurisprudence.

In my June 2018 report to the Human Rights Council, I noted that company standards based on “vague assertions of community interests” has “created unstable, unpredictable and unsafe environments for users and intensified government scrutiny” - the very problems that the creation of the Board seeks to address. In contrast, international human rights standards “enable companies to create an inclusive environment that accommodates the varied needs and interests of their users while establishing predictable and consistent baseline standards of behavior.” Furthermore, a consistently human rights-based approach to content decisions will enable Facebook to “stand on firmer ground when [it] seek[s] to hold States accountable to the same standards,” and resist State attempts to exploit its content policies to censor content. (All citations to my June Report are to UN Document A/HRC/38/35, which may be found at the website of my mandate.)

The jurisprudence of Article 19 of the ICCPR (not to mention other related jurisprudence, such as that of the European and Inter-American human rights courts)

would provide the oversight board with a set of tools and a “common vocabulary” for addressing and resolving hard questions around the moderation of online content. The Human Rights Committee, the body charged with monitoring implementation of the ICCPR, has broadly interpreted expression protected under Article 19(2) to include “political discourse, commentary on one’s own and on public affairs, canvassing, discussion of human rights, journalism, cultural and artistic expression, teaching, and religious discourse” as well as “expression that may be regarded as deeply offensive.” (General Comment 34, CCPR/C/GC/34) Article 19(3)’s requirements of legality, necessity and legitimacy of objectives provide guidance on how Facebook should develop and implement content standards in a manner that respects users’ freedom of expression.

Measures Facebook has adopted, for instance in the face of anti-vaccination disinformation campaigns, are often understandable responses to unfolding crises, but their ad-hoc development may be susceptible to criticisms of bias and arbitrariness. Aligning these measures with human rights standards, however, can place them on a more principled footing. Under Article 19(3), restrictions on expression may be validly imposed if they are “provided by law” and “necessary” to serve a legitimate objective, such as the protection of public health. The Human Rights Committee has found that “law” must be “formulated with sufficient precision to enable an individual to regulate his or her conduct accordingly.” Even though Facebook does not make laws, the general principles of legality should nevertheless guide Facebook’s development of its rules and policies. In the context of its response to vaccine misinformation, for example, these principles would at least require Facebook to provide more information about how it defines “vaccine misinformation,” the processes it has developed for flagging such content, and the types of consultations it conducted in developing these measures and with whom it consulted. These are also the kinds of considerations that the Board, to provide genuine oversight, should be equipped to assess in reviewing appeals of content decisions.

Article 19(3) also provides concrete metrics for assessing the impact of particular forms of expression on its platform, and calibrating a proportionate response to address such impacts. Under the requirement of legitimacy of objectives, it is incumbent on those advocating for restrictions to explain the “precise nature of the threat” and assess whether there is a “direct and immediate connection between the expression and the threat.” (CCPR/C/GC/34) In this example, these principles should lead Facebook to assess and explain how the spread of vaccine misinformation on its platforms raises public health concerns. Under the requirement of necessity, restrictions on expression must be “appropriate to achieve their protective function,” the “least intrusive instrument amongst those which might achieve their protective function” and “proportionate to the interest to be protected.” (*Id.*) Considerations of proportionality provide Facebook with a principled and internationally recognized framework for evaluating its decision to demote and de-emphasize anti-vaccination content rather than categorically ban such content on its platforms. Again, these are also the kinds of questions that the Board could be authorized to address in its review of content decisions.

B. Independence of the Board

A substantial portion of the draft charter reflects on the need to safeguard the independence of the Board and its decision-making. While this is encouraging, I urge Facebook to consider how human rights standards concerning judicial independence and other forms of external oversight may inform the composition, governance and decision-making authority of the Board.

Article 14(1) of the ICCPR states that all persons “shall be equal before the courts and tribunals,” and that “everyone shall be entitled to a fair and public hearing by a competent, independent and impartial tribunal established by law.” The Basic Principles on the Independence of the Judiciary, which were endorsed by the UN General Assembly in 1985, stipulate that “[a]ny method of judicial selection shall safeguards against judicial appointments for improper motives.” Over the years, multiple UN Special Rapporteurs on the independence of judges and lawyers have recognized that the most effective method for securing judicial independence during the appointments process is through a “well-entrenched independent mechanism ... responsible for the appointment, promotion, transfer and dismissal of judges.” (See UN Documents E/CN.4/1995/39; A/HRC/11/41; A/HRC/35/31)

I am mindful that the Board is not a court of law and do not mean to suggest that these standards are strictly applicable or obligatory on companies. However, these standards could inform Facebook’s choices on initial appointments and future selection process. For example, the draft charter suggests that Facebook will “select the first cohort” of board members “based on a review of qualifications that will be made public.” However, the experience of international human rights law indicates that another option under consideration – delegating initial appointments to a selection committee – may be a more robust guarantee of the Board’s independence. Even though the draft charter indicates that determining that committee may “create its own selection challenges,” it would still function as critical layer of separation between Facebook and the Board.

Consistent with international best practice on ensuring independent oversight, the draft charter would empower the board with final decision-making authority over the selection of Board members. To strengthen this assurance of independence, the charter should also establish objective criteria for the selection of future Board members that details the qualifications required for the position, including relevant standards of “integrity, ability and efficiency.” (A/HRC/11/41) In the interests of transparency and public accountability, public consultations should also be a mandatory component of the selection process.

I am, however, concerned that the draft charter preserves the authority of Facebook to remove members if they are deemed to have “violated the terms of his or her appointment.” If the terms of appointment are vaguely formulated (for example, if they contain broad prohibitions on disclosing confidential information), Facebook would effectively exercise broad discretion to remove Board members, to the detriment of their independence. Decisions to remove Board members should remain solely within the

purview of the Board. Furthermore, the criteria for removals should be based only on “reasons of incapacity or behavior that renders them unfit to discharge their duties,” such as “serious grounds of misconduct or incompetence.” (A/HRC/11/41)

Finally, given that the vast majority of Facebook’s users are outside the United States, I am encouraged that Facebook is committed to ensuring that the Board reflects “geographic and cultural balance as well as a diversity of backgrounds and perspectives.” To honor this commitment, special effort should be made to recruit Board members from the Global South, and from groups or communities that have experienced chronic, well-documented discrimination on Facebook’s platforms and apps.

C. Independent Review

The effectiveness and legitimacy of the Board will also depend on its capacity to conduct an independent and comprehensive review of content decisions. I urge Facebook to provide the board with fact-finding capacities and resources that will enable it to effectively assess whether content decisions are in line with international human rights standards.

Meaningful Review of Company Actions

Independent review of content decisions will require access to information concerning the company policies, processes, deliberations and actions leading to those decisions. To assess the proportionality of a content removal or account suspension, for example, the Board may require information pertaining to previous content actions taken against the user(s) at issue, the availability and feasibility of other content-related measures (e.g. demoting rather than removing the content at issue), and factors that may amplify the content at issue (e.g. recommendation algorithms, bot accounts, ad policies).

The Board should also be permitted to interview relevant Facebook staff and contractors, including content moderators and members of the policy, legal, product design and engineering teams. To address privacy, confidentiality and safety considerations that may arise from these investigations, Facebook should consider the implementation of relevant safeguards, such as the anonymization of sensitive information and guarantees of non-reprisals against staff or contractors who provide critical feedback to the Board.

Local Engagement and Participation

The willingness and capacity to engage with civil society and other external stakeholders are also essential components of an independent review process. I appreciate that the draft charter proposes that the “board will be able to call upon experts to ensure it has all supplementary linguistic, cultural and sociopolitical expertise necessary to make a decision,” and that “Facebook users and pertinent stakeholders may also submit arguments and material to the panel.” However, I urge Facebook to ensure that the deliberative process of the Board prioritizes engagement with communities and groups

historically at risk of censorship and discrimination, and particularly those that may not be represented within mainstream civil society (such as representatives of certain indigenous groups).

Under the Guiding Principles, due diligence requires Facebook to identify, address and account for “actual and potential human rights impacts of their activities, including through regular risk and impact assessments, meaningful consultation with potentially affected groups and other stakeholders, and appropriate follow-up action that mitigates or prevents these impacts”(UN GUIDING PRINCIPLES, Principles 17–19). In my June 2018 report, I found that inconsistent enforcement of content policies tends to “penaliz[e] minorities while reinforcing the status of dominant or powerful groups.” I have urged companies to ensure that its content policies eschew “formalistic approaches” to non-discrimination and take into account the disparate impact of abuse, harassment and other forms of censorship on women, LGBTQ populations, racial, ethnic and religious minorities, migrant communities and other at-risk users.

To ensure that the Board addresses the frequently discriminatory impact of content decisions, I urge Facebook to ensure that the Board’s deliberations provide a meaningful avenue for engagement with relevant local communities and their representatives, particularly in the Global South. The draft charter should, at a minimum, provide clear direction on how external input will be solicited and integrated into the Board’s decision-making. Facebook should also think creatively about how to maximize local engagement: for example, it has been suggested that the draft charter should provide for “aggregated or ‘class action’ complaints” and opportunities for third party interventions.¹ The Board should also be given a budget to cover expenses associated with third party participation in its deliberations, such as travel-related expenses or the costs of interpreters and accommodations for participants living with disabilities. These standards should also guide Facebook’s own consultations for establishing the Board.

Implementation of the board’s decisions

Under the draft charter, Facebook would be “ultimately responsible for making decisions related to policy, operations and enforcement.” While Facebook has acknowledged that the board’s decisions will be binding on the “specific content brought for review,” the Board should retain oversight of the implementation of its decisions. For example, if content that the Board decides should be restored continues to be removed or otherwise restricted on Facebook’s platforms, the Board should be authorized to investigate the reasons for this inconsistency and propose broader policy guidance.

D. Transparency

The draft charter’s discussion of transparency focuses on the decisions of the Board, which it pledges to make public “with all appropriate privacy protections for users.” Decisions will be “issued on behalf of the board and will not be attributed to

¹ <https://www.article19.org/resources/facebook-oversight-board-recommendations-for-human-rights-focused-oversight/>

individual panel members,” and provide an opportunity for dissenting members to include their perspective “as part of the explanation shared.” While I welcome the proposal to issue publicly available explanations of the Board’s decisions, I believe that transparency of decision-making is merely part of a more holistic approach to transparency that Facebook should adopt in line with its human rights responsibilities.

Under the Guiding Principles, “business enterprises whose operations or operating contexts pose risks of severe human rights impacts should report formally on how they address them,” and “provide information that is sufficient to evaluate the adequacy of an enterprise’s response to the particular human rights impact involved.” (UN GUIDING PRINCIPLES, Principle 21) The Guiding Principles emphasize that transparency takes “a variety of forms, including in-person meetings, online dialogues, consultation with affected stakeholders, and formal public reports.” (*Id.*) In light of these standards, I explained in my June 2018 report that companies should “embark on radically different approaches to transparency at all stages of their operations, from rule-making to implementation and development of “case law” framing the interpretation of private rules.” Transparency also requires “greater engagement with digital rights organizations and other relevant sectors of civil society and avoiding secretive arrangements with States [or non-State actors] on content standards and implementation.”

Transparency should lie at the heart of all of the board’s activities, from its founding to its selection process, governance and decision making. At this critical stage of the Board’s development, Facebook should release timely updates and summaries of its consultations on the draft charter, including key findings and recommendations and its efforts to solicit geographically and culturally diverse input. The process of selecting members should be open to input from the public and third party experts, and changes to the rules of governance or the terms of appointment should be communicated to the public.

In the context of decisional transparency, I urged, in my June 2018 report, social media platforms to “develop a kind of case law that would enable users, civil society and States to understand how the companies interpret and implement their standards,” including through the creation of a “detailed repository of cases and examples would clarify the rules much as case reporting [in the courts] does.” This body of “platform law” should cover not only the board’s substantive review of content decisions but also its interpretations of the case selection criteria. The board’s decisions also provide an opportunity to clarify the extent to which it has considered and applied relevant human rights standards, including its interpretations of how these standards should be adapted to the realities of how online content is generated, shared and amplified, the role of Facebook in moderating such content, and regional or local particularities. Updates on the implementation of the Board’s decisions should also be provided.

E. A Holistic Approach to Due Diligence, Oversight and Accountability

The need for independent and external oversight of Facebook’s content decision-making practices may also require mechanisms of appeal and remedy that exist entirely

outside of Facebook. In this spirit, I urge you to seriously engage with and support proposals for public accountability led by civil society, such as the proposal for a multi-stakeholder, cross-industry Social Media Council.

More broadly, a rights-oriented approach to commercial content moderation “begins with rules rooted in rights, continues with rigorous human rights impact assessments [HRIAs] for product and policy development, and moves through operations with ongoing assessment, reassessment and meaningful public and civil society consultation.” (JUNE 2018 REPORT). In particular, HRIAs will enable Facebook to identify, prevent and mitigate adverse human rights outcomes that future revisions of its content moderation policies and processes may create. Facebook should also consider how it can harness the expertise of the Board to strengthen and mainstream HRIAs throughout its content moderation operations.

There are also areas of Facebook’s operations that, while outside the proposed jurisdiction of the board, nevertheless implicate freedom of expression and should be subject to human rights due diligence, oversight and accountability. Notably, the board “will not decide cases where reversing Facebook’s decision would violate the law.” I appreciate that the requirements of local law and other State pressures to restrict content may compel company action that is inconsistent with human rights standards. However, if Facebook is unable to prevent adverse human rights impacts in connection with State demands or requests, it has a responsibility to minimize these impacts to the “greatest extent possible.” (UN GUIDING PRINCIPLES, Principle 19) These prevention and mitigation strategies range from adopting interpretations of local laws that “resolve any legal ambiguity in favour of respect for freedom of expression [and other human rights],” challenging overbroad requests in court, to creative approaches to transparency that “disclose all relevant and publishable information” concerning government requests. (A/HRC/35/22) At a minimum, Facebook should seek independent and external review of its approach to State restrictions for consistency with these standards.

Facebook’s responsibility to respect freedom of expression is also triggered across a wide range of activities, including ad personalization and targeting, “the curation of user feeds and other forms of content delivery, the introduction of new features or services and modifications to existing ones ... and market-entry decisions such as arrangements to provide country-specific versions of the platform.” (JUNE 2018 REPORT) I am particularly concerned about the use of artificial intelligence and related technologies to moderate and curate content, in ways that “invisibly supplant, manipulate or interfere with the ability of individuals to form and hold their opinions or access and express ideas in the information environment.” (A/73/348) In my September 2018 report to the General Assembly, I called on the ICT sector to adopt both HRIAs and rights-based audits of their applications of artificial intelligence. I also urged innovative approaches to notice and consent that signal to users “when an artificial intelligence system is determining a user’s experience,” when their data is being collected by or to train AI applications, and the conditions under which such data will be used, stored or deleted. (*Id.*)

I would be pleased to have further discussion with you and your colleagues about the proposed Oversight Board. I am also happy to discuss how Facebook could implement its human rights responsibilities, and provide you with further information about my mandate.

In keeping with my own commitment to transparency in the work of my mandate, I will be making a copy of this letter available to the public and posting it on the website page for the mandate of the Special Rapporteur on the right to freedom of expression: (<http://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/LegislationAndPolicy.aspx>). This communication, as well as any response received, will also be made available in the communications reporting website of the OHCHR (<https://spcommreports.ohchr.org/Tmsearch/TMDocuments>) within two working days. It will also subsequently be made available in the usual report to be presented to the Human Rights Council.

Sincerely,

David Kaye

Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression