

Mandates of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression and the Working Group on the issue of human rights and transnational corporations and other business enterprises

REFERENCE:
OL OTH 37/2020

15 May 2020

Dear Mr. Yiming:

We have the honour to address you in our capacities as Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression and Working Group on the issue of human rights and transnational corporations and other business enterprises, pursuant to Human Rights Council resolutions 34/18 and 35/7.

We are independent human rights experts appointed and mandated by the United Nations Human Rights Council to report and advise on human rights issues from a thematic or country-specific perspective. We are part of the special procedures system of the United Nations, which has 56 thematic and country mandates on a broad range of human rights issues. We are sending this letter under the communications procedure of the Special Procedures of the United Nations Human Rights Council to seek clarification on information we have received. Special Procedures mechanisms can intervene directly with Governments and other stakeholders (including companies) on allegations of abuses of human rights that come within their mandates by means of letters, which include urgent appeals, allegation letters, and other communications. The intervention may relate to a human rights violation that has already occurred, is ongoing, or which has a high risk of occurring. The process involves sending a letter to the concerned actors identifying the facts of the allegation, applicable international human rights norms and standards, the concerns and questions of the mandate-holder(s), and a request for follow-up action. Communications may deal with individual cases, general patterns and trends of human rights violations, cases affecting a particular group or community, or the content of draft or existing legislation, policy or practice considered not to be fully compatible with international human rights standards.

As independent experts appointed by the UN Human Rights Council, it is our responsibility to gather information and report on how governments and non-state actors, including private companies, promote and protect every individual's right to seek, receive, and impart information and ideas of all kinds, regardless of frontiers and through any media. In discharging the mandate, we communicate regularly with governments and private industry and investigate the extent, nature and severity of restrictions and violations, reporting our findings to the Human Rights Council and General Assembly. As digital space has become a primary medium for debate, discussion, and exchange of information, the work of the Special Rapporteur on the freedom of opinion and expression has paid particular attention to expression on online platforms, with recent reporting on the role of States and the information and communications technology ("ICT") industry in providing an enabling environment for online expression.

TikTok

In October 2019, TikTok announced its plan to “[c]reate a committee of outside experts to advise on and review content moderation policies covering a wide range of topics, including child safety, hate speech, misinformation, bullying, and other potential issues,” and “[f]urther increase transparency around our content moderation policies and the practices we employ to protect our community” as part of the “ongoing commitment to maintaining the app experience users expect while providing them the protections they deserve.”¹ In March 11, 2020 TikTok announced refined plans to launch a “Transparency Center for moderation and data practices.”² Welcoming TikTok’s initiative to engage with legal experts to refine its content moderation policies and to increase transparency around content decisions, we would like to provide some comments to encourage the incorporation of human rights standards into your policies and would welcome the opportunity to further engage with TikTok on the subject.

Content moderation by companies may range from down-ranking content to permanent removal of content or accounts, with significant impact on freedom of expression and other fundamental rights such as the right to freedom of association and public participation. At the same time, companies face pressure from governments to restrict expression, as well as public pressure to meet business goals and expectations. In dealing with these challenges, internal policies and guidelines, terms of service agreements and community standards have often fallen short of providing adequate solutions. This has resulted in unpredictable environments for users, subjecting the companies to criticism of bias and arbitrariness.³

Given the complexity of the task of content moderation, we urge you to use international human rights standards as your guidepost and integrate the following suggestions into your policies. It is our view that putting human rights at the centre of your content moderation infrastructure will allow TikTok to stand on a firm, principled ground that may prevent and mitigate adverse human rights impacts while maintaining an environment that accommodates the varied needs and interests of its users.

I. Standards for Content Moderation

Human rights law pertaining to freedom of opinion and expression is outlined in the International Covenant on Civil and Political Rights (“ICCPR”). Article 19(1) of the ICCPR protects everyone’s right to freedom of opinion without interference. Article 19(2) of the ICCPR guarantees freedom of expression, which it defines as the “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

¹ TikTok, Our Commitment to Our Users and the TikTok Experience (15 October 2019),

<https://newsroom.tiktok.com/en-us/our-commitment-to-our-users-and-the-tik-tok-experience>.

² TikTok, TikTok to launch Transparency Center for moderation and data practices (11 March 2020)

<https://newsroom.tiktok.com/en-us/tiktok-to-launch-transparency-center-for-moderation-and-data-practices>

³ David Kaye, *Rep. of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, ¶ 41, U.N. Doc. A/HRC/38/35 (Apr. 6, 2018), [hereinafter 2018 Report on Content Moderation], available at <http://undocs.org/A/HRC/38/35>.

choice.”⁴ Article 19(3) establishes permissible limitations on freedom of expression: restrictions on the right to freedom of expression must be “provided by law,” and necessary “for respect of the rights or reputations of others” or “for the protection of national security or of public order (*ordre public*), or of public health and morals.”⁵ The General Assembly, the Human Rights Council and the Human Rights Committee have concluded that the same rights available offline apply online as well.⁶

In his 2018 report to the Human Rights Council on content moderation, the Special Rapporteur for the freedom of opinion and expression explained the applicability of these principles to private companies through the United Nations Guiding Principles on Business and Human Rights (“the Guiding Principles”), which establishes that “‘global standard[s] of expected conduct’ should apply throughout company operations and wherever they operate.”⁷ The Human Rights Council endorsed the Guiding Principles in 2011 and they have become a key tool for the private sector’s implementation of human rights principles and civil society’s evaluation of the conduct of private companies. Accordingly, many companies have acknowledged their responsibility to respect human rights and demonstrated commitment to engage closely with human rights mechanisms in conducting content moderation.

The legality, or “provided by law”, requirement provides that any rules limiting expression must be “formulated with sufficient precision to enable an individual to regulate his or her conduct accordingly.” Vague and overbroad law “triggers complaints of inconsistent policy enforcement that penalizes minorities while reinforcing the status of dominant powerful groups,”⁸ resulting in “unstable, unpredictable and unsafe environments for users.”⁹ Furthermore, any restriction on expression must be public and transparent.¹⁰ These general principles of legality should guide the development of platform rules. Therefore, while certain content may be legitimately restricted under human rights law, platforms must engage in rigorous efforts to define categories of restricted content. This will require sufficient precision and clarity so that users can predict with reasonable certainty and consistency what content may be subject to removal or other form of restrictions.

⁴ International Covenant on Civil and Political Rights, Mar. 23, 1976, available at <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>.

⁵ *Id.*

⁶ See G.A. Res. 68/167 (Dec. 18, 2013),

http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/68/167; Human Rights Council Res. 26/13, U.N. Doc. A/HRC/RES/26/13 (June 26, 2014), http://hrlibrary.umn.edu/hrcouncil_res26-13.pdf; U.N. Comm’n on Human Rights.

⁷ 2018 Report on Content Moderation, *supra* note 2, ¶9.

⁸ 2018 Report on Content Moderation, *supra* note 3, ¶ 27

⁹ *Id.*

¹⁰ David Kaye, Rep. of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, ¶ 6, U.N. Doc. A/74/486 (Oct. 9, 2019), [hereinafter 2019 Report on Hate Speech], available at https://www.ohchr.org/Documents/Issues/Opinion/A_74_486.pdf.

Restrictions on content must also have a legitimate objective. Legitimate grounds include protection of one or more interests specified in Article 19(3), that is, the rights or reputations of others, national security, public order, public health or morals.¹¹

Finally, the necessity prong places the onus on companies to explain the “precise nature of the threat” requiring a restriction.¹² Additionally, platforms should demonstrate that there is a “direct and immediate connection between the expression and the threat.”¹³ It follows that restrictions must be more than merely useful, reasonable or desirable. Any restrictive measure implemented 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”.¹⁴

Considerations of proportionality provide companies with a principled and internationally recognized framework for evaluating content. They should identify a continuum of options and select the least restrictive approach to deal with the specific harm, reserving the most severe penalties such as the banning of a user from an online service as the last resort. These rules will enable your company to protect users’ rights while responding effectively to activities that seek to silence certain voices, infringe on privacy, or use the platforms’ tools to incite violence.¹⁵

Given the size of TikTok’s user base and the impact of your company on human rights of users, we believe that commitment to respect human rights must be clearly articulated in your terms of service and community standards, and that your decision-making processes should be built upon the standards of legality, legitimacy, and necessity and proportionality.

II. Obligation of Non-Discrimination

In his 2018 report, the Special Rapporteur on the freedom of opinion and expression discussed the importance of providing meaningful guarantees of non-discrimination in content moderation.¹⁶ Non-discrimination is enshrined in Article 26 of the ICCPR as a stand-alone guarantee of equality before the law and equal protection of the law. Under Article 26, States have the obligation to “prohibit any discrimination and guarantee to all persons equal and effective protection against discrimination on any ground such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status.”¹⁷ As such, Article 19 in conjunction with Article 26 entails the right to freedom from discrimination in the holding and forming of opinions, the expression of and access to ideas and information, and the exercise of

¹¹ See General Comment 34, available at <https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>

¹² General Comment 34, available at <https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>

¹³ General Comment 34, available at <https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>

¹⁴ General Comment 34, available at <https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>

¹⁵ See Generally 2019 Report on Hate Speech, *supra* note 12, ¶6.

¹⁶ See 2018 Report on Content Moderation, *supra* note 3, ¶48.

¹⁷ International Covenant on Civil and Political Rights, Mar. 23, 1976, available at <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>.

privacy and the protection of personal data.¹⁸

Content moderation systems, which are designed to assess a huge volume of user-generated content in a short period of time, often have inherent limitations that can perpetuate bias and discrimination. For example, companies use artificial intelligence-driven systems as a means to quickly filter through content across the range of their rules.¹⁹ The Special Rapporteur on the freedom of opinion and expression has raised concern that such moderation algorithms are particularly problematic in dealing with cultural, language or gender-based contexts and sensitivities, or public interest in the content.²⁰ Other commonly used tools, such as user flagging, may also prove inadequate. User flags, which give individuals the ability to log complaints, do not encourage nuanced discussion of appropriate boundaries, and are easily manipulated to coordinate harassment and trolling to remove legitimate content.

To create a system in line with human rights standards, we urge TikTok to embrace a fundamental commitment to non-discrimination. Such an approach would require the company to move away from the formalistic and uniform treatment of all content and allow modifications and variances to the policies and practices to reflect the particularities different communities.²¹ TikTok's content moderation policies should be malleable, taking into account linguistic and cultural nuances and sensitivities. The Transparency Center indicates that the company may wish to build such a system. Meaningful implementation though, will require meaningful and regular conversations with various stakeholders from different regions who can evaluate emerging issues and provide insight into difficult, context-based content problems. Process will be key.

III. Human rights Due Diligence

In the 2018 report of the Working Group on the issue of human rights and transnational corporations and other business enterprises (Working Group) to the General Assembly the working Group noted that “The Guiding Principles clarify that business enterprises have an independent responsibility to respect human rights and that in order to do so they are required to exercise human rights due diligence. Human rights due diligence refers to the processes that all business enterprises should undertake to identify, prevent, mitigate and account for how they address potential and actual impacts on human rights caused by or contributed to through their own activities, or directly linked to their operations, products or services by their business relationships.”²² Effective due diligence involves identifying and assessing actual or potential adverse human rights impacts that the enterprise may cause or contribute to through its own activities, or which

¹⁸ David Kaye, Rep. of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, ¶ 36, U.N. Doc. A/73/348 (Aug. 29, 2018), available at <https://freedex.org/wp-content/blogs.dir/2015/files/2018/10/AI-and-FOE-GA.pdf>.

¹⁹ Id. ¶¶ 13-16.

²⁰ Id.

²¹ See 2018 Report on Content Moderation, *supra* note 3, ¶48.

²² United Nations Guiding Principles on Business and Human Rights, Principles 17-19, available at https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf.

may be directly linked to its operations, products or services by its business relationships; (b) Integrating findings from impact assessments across relevant company processes and taking appropriate action according to its involvement in the impact; (c) Tracking the effectiveness of measures and processes to address adverse human rights impacts in order to know if they are working; (d) Communicating on how impacts are being addressed and showing stakeholders – in particular affected stakeholders – that there are adequate policies and processes in place.²³

As TikTok continues to refine its content moderation framework, the company must be mindful that it is necessary to conduct robust due diligence, including periodic impact assessments and reassessments of the policies and practices. The ICT sector evolves rapidly, and new practices or trends can result in unintended or disproportionate limitations on users' rights. To mitigate such harms, we recommend TikTok to develop clear and specific criteria for identifying activities that trigger reassessments, followed by thorough rights-based evaluations of such activities. The triggering activities may range from revisions of content moderation policies and processes to introduction of new features or services, implementation of automation technologies, and market-entry decisions specific to certain countries. Periodic review of internal policies and practices must be complemented by regular training of content policy teams, general counsel and especially content moderators in the field who conduct the actual work of restriction.²⁴

In addition, as articulated in the 2018 report of the Special Rapporteur on the freedom of opinion and expression, companies should seek to mitigate human rights risks posed by national laws or demands that are inconsistent with the Article 19(3) principles.²⁵ While companies have duty to comply with the local laws where they conduct business, companies responding to government removal requests must also keep in mind that any restrictions on the right to freedom of expression must be exceptional, and that States bear the burden of demonstrating the consistency of such restrictions with the human rights principles. When faced with problematic requests, for example requests based on vaguely formulated laws subject to varying interpretations, “companies should seek clarification or modification; solicit the assistance of civil society, peer companies, relevant government authorities, international and regional bodies and other stakeholders; and explore all legal options for challenge.”²⁶ While we recognize the complexity of balancing content moderation and national laws, we express general concern that companies are giving too much weight to national demands at the expense of the internationally recognized human rights of individuals. We encourage TikTok to refer to the Guiding Principles for additional guidance on minimizing the adverse impacts on human rights of national laws and government requests.²⁷

²³ Summary of the report of the Working Group on Business and Human Rights to the General Assembly, October 2018 (A/73/163), available at

https://www.ohchr.org/Documents/Issues/Business/A_73_163_ExecutiveSummary_EN.pdf

²⁴ See Guiding Principle 16, commentary.

²⁵ 2018 Report on Content Moderation, *supra* note 2, ¶¶ 22-25.

²⁶ *Id.* ¶¶ 49-51.

²⁷ *Id.* ¶ 23.

Finally, we strongly urge all companies to prioritize engagement with users and civil society, for example through open comment processes or consultations. We are encouraged by the recent announcement of the Transparency Center for moderation and data practices. This could be an important step towards gaining input from affected rights holders and relevant local or subject matter experts. Communities historically at risk of censorship and discrimination will be especially valuable in accurately assessing the human rights impacts from diverse perspectives and understanding the hyper-local impacts of the rules, we encourage the transparency center to listen to users and experts from those communities.²⁸

IV. Transparency

On December 30, 2019, TikTok released its first transparency report on legal requests for user information and content removal that it received in the first half of 2019, as a part of “regular updates” that the company will provide for users.²⁹ While we welcome the report as a valuable step taken to increase transparency and accountability, we believe that human rights responsibilities require a more holistic approach to transparency. The Transparency Center is a positive step towards achieving a more holistic approach, however, the implementation of this center will require adherence to human rights principles.

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.”³⁰ Noting the lack of transparency as a major flaw in all the companies’ content moderation processes, the Special Rapporteur on the freedom of expression has stated 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.”³¹

In line with its human rights responsibilities, we encourage TikTok to develop measures that would enable the public to gain real insight into the company’s rulemaking process. This includes disclosing and updating all users on the company’s internal process for interpreting and applying content moderation rules so that users can sufficiently understand and predict the content decisions. Reports asserting that TikTok holds conferences with some of the more prominent users and the “safety team” to clarify community guidelines proves that TikTok is capable of educating its user base more widely on community guidelines, as it should. It is particularly important to provide meaningful and consistent transparency around development of policies governing contentious issues, such as hate speech to all users, not just a select few.

²⁸ Id. ¶ 54.

²⁹ TikTok Transparency Report (Dec. 30, 2019), <https://www.tiktok.com/safety/resources/transparency-report>.

³⁰ Guiding Principle 21.

³¹ 2018 Report on Content Moderation, *supra* note 2, ¶71.

Furthermore, we urge TikTok to increase decisional transparency. When the content moderation committee, or another adjudicating body within the company, deliberates on certain issues or cases, it should publicly disclose the issues and explanations as to what standards and principles it has considered in reaching the decisions, as well as the role of moderation algorithms and other moderation tools in such processes. Recent reports note that content uploaded to the platform depicting ‘undesirable’ visuals is being removed. Down-ranking or removing content based on appearance of the user or their residence is simply unacceptable, as is removing content depicting low-income neighborhoods or natural disasters. Adherence to human rights standards would provide a framework for TikTok to avoid blatant human rights violations such as these, as well as increasing transparency and implementing decision-making criteria grounded in human rights norms. We would greatly appreciate if you could provide more information on how the content moderation process works and the criteria that is used.

Finally, in addition to information on government requests, TikTok should disclose information on private removal requests of user-generated contents and the company’s rate of compliance. Just as TikTok recognized in its recent statement,³² transparency is key not only in increasing the company’s accountability and protection for the users, but also in increasing the legitimacy of the company’s decisions.

V. Access to Due Process and Grievance Mechanism for Users

The company’s announcement of the Transparency Center provides positive reinforcement of the company’s previously stated goals of achieving greater transparency and legitimacy. However, the guidelines establishing the Transparency Center remain vague. The outlined goals of having civil society, policymakers, and the broader public understand TikTok’s decision-making process is positive. Implementation, however, will be difficult given the volume of content that is reviewed, removed, or demoted every day on the platform. As such, we recommend that TikTok release detailed guidelines regarding due process safeguards for its users including easy and convenient access to platform standards, notice to users in the case of violations of these standards, and an internal procedure for redress in the case of mistakes. This would provide a means for experts to impart useful and meaningful advice, utilizing the Transparency Center in the way it is seemingly designed to be used.

The responsibility to provide adequate remedies for the victims of human rights violation is articulated in Article 2 of the ICCPR, and the Guiding Principles also provide that companies should provide for or cooperate in remediation through legitimate processes where business enterprises identify that they have caused or contributed to adverse impacts for those whose rights have been violated from erroneous or excessive content decisions and establish or participate in effective operational-level grievance

³² TikTok, Our Commitment to Our Users and the TikTok Experience (Oct. 15, 2019), <https://newsroom.tiktok.com/en-us/our-commitment-to-our-users-and-the-tik-tok-experience>.

mechanisms for individuals and communities who may be adversely impacted. The 2018 report of the Special Rapporteur on the freedom of opinion and expression reiterated that companies should embrace a commitment to remedial measures by “institut[ing] robust remediation programmes, which may range from reinstatement and acknowledgment to settlements related to reputational or other harms.”³³

First, before or coterminous with the removal of content, users should receive notice and adequate explanations of the platform’s decision, including specific rules that have been broken, how content moderation guidelines were interpreted, and the specific action that will be taken. With the notice, users should be provided with a channel to appeal the decision and a guarantee to receive a prompt and meaningful response. In addition to meaningful due process, TikTok should recognize the obligation to provide access to remedy to the extent their takedown or other content decisions cause human rights harms. Company policies should not prevent users from pursuing legal recourse, nor force users to renounce such recourse.

Robust due process and grievance mechanisms are essential in ensuring consistency of the decisions, reducing the likelihood of discriminatory practices, and establishing sector-wide accountability.

In light of the above comments, we strongly encourage TikTok to take all steps necessary, including adopting the recommendations on this communication, to create a rights-based content moderation system that protects the ability of individuals to use its online platform as forum for free expression. We believe that a content moderation framework founded on human rights principles will equip TikTok with a set of tools to effectively resolve the challenges that it faces as a new company.

We welcome your comments in response to this letter and would enjoy the opportunity to engage in a dialogue with you and members of your team. We would like to inform you that, in keeping with our commitment to transparency in the work of our mandates, this communication, as a comment on pending or recently adopted policies, and any response received will be made public via the communications reporting [website](#) and on the [website page for the mandate of the Special Rapporteur on the right to freedom of expression](#). The publication will take place within 48 hours. The communication and the response 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

³³ See 2018 Report on Content Moderation, *supra* note 2, ¶59.

Githu Muigai
Chair-Rapporteur of the Working Group on the issue of human rights and transnational
corporations and other business enterprises