

Mandates of the Working Group of Experts on People of African Descent; the Working Group on the issue of human rights and transnational corporations and other business enterprises; the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression; the Special Rapporteur on the promotion and protection of human rights in the context of climate change; the Special Rapporteur in the field of cultural rights; the Special Rapporteur on the right to development; the Special Rapporteur on the right to education; the Special Rapporteur on the right to food; the Special Rapporteur on the rights to freedom of peaceful assembly and of association; the Special Rapporteur on the independence of judges and lawyers; the Special Rapporteur on the human rights of internally displaced persons; the Independent Expert on the promotion of a democratic and equitable international order; the Independent Expert on human rights and international solidarity; the Working Group on the use of mercenaries as a means of violating human rights and impeding the exercise of the right of peoples to self-determination; the Special Rapporteur on the human rights of migrants; the Special Rapporteur on minority issues; the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance; the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism; the Special Rapporteur on violence against women and girls, its causes and consequences and the Working Group on discrimination against women and girls

Ref.: AL OTH 125/2022
(Please use this reference in your reply)

29 December 2022

Mr. Elon Musk,

We have the honour to address you in our capacities as Working Group of Experts on People of African Descent; Working Group on the issue of human rights and transnational corporations and other business enterprises; Special Rapporteur on the promotion and protection of human rights in the context of climate change; Special Rapporteur in the field of cultural rights; Special Rapporteur on the right to development; Special Rapporteur on the right to education; Special Rapporteur on the right to food; Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression; Special Rapporteur on the rights to freedom of peaceful assembly and of association; Special Rapporteur on the independence of judges and lawyers; Special Rapporteur on the human rights of internally displaced persons; Independent Expert on the promotion of a democratic and equitable international order; Independent Expert on human rights and international solidarity; Working Group on the use of mercenaries as a means of violating human rights and impeding the exercise of the right of peoples to self-determination; Special Rapporteur on the human rights of migrants; Special Rapporteur on minority issues; Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance; Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism; Special Rapporteur on violence against women and girls, its causes and consequences and Working Group on discrimination against women and girls, pursuant to Human Rights Council resolutions 45/24, 44/15, 48/14, 46/9, 51/7, 44/3, 49/13, 43/4, 50/17, 44/8, 50/L4, 45/4, 44/11, 42/9, 43/6, 43/8, 43/36, 49/10, 50/7 and 50/18.

We are sending this letter under the communications procedure of the Special Procedures of the United Nations Human Rights Council to seek clarification on the

Twitter Inc.

information we have received¹. 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.

In this connection, we would like to bring to your attention the **harmful impact of hate speech, including misogynistic hate speech, incitement to hatred and racism targeting people of African descent and spreading on social media such as Twitter. It is especially alarming given the inadequacy of the responses and solutions put in place to counter them.**

According to the information received:

Twitter, Inc. plays a primary role in the digital ecosystem. It has been reported that in the early days of your acquisition of Twitter, the use of the hateful and racist “N” word on the platform increased by almost 500 percent within a 12-hours period compared to the previous average.

The measures actually in place to counter hate speech, including misogynistic hate speech, incitement of hatred, modern slavery and other online conducts targeting people of African descent are proven to be ineffective. The above-mentioned facts, in their seriousness, are blatant examples of it. We have been informed that some attempts have been made to regulate and monitor content upon publication. However, these measures are not sufficient if they do not ensure sustained, systematic and real checks aimed at effectively assessing the harmfulness of the contents in line with international standards for freedom of expression. Furthermore, a thorough review of business models is necessary to ensure that the business operations, including in relation to content moderation and the amplification of potentially harmful content are compliant with international human rights standards, including freedom of expression.

It is also reported that almost the entire human rights oversight department of Twitter has been fired and there is currently a lack of capacity to monitor human rights compliance on the platform.

Hate speech based on gender and misogyny often intersects with other forms of hate speech and xenophobia based, for example on sexual orientation, nationality, race, ethnicity, religion, migration status, and disability. The rise of social media platforms revealed the ubiquity and intersectionality of online hatred directed against women, girls, minorities and gender diverse people. Despite the lack of comprehensive data, it has been estimated that 23 per cent

¹ Further information about the communication procedure is available at: <http://www.ohchr.org/EN/HRBodies/SP/Pages/Communications.aspx>.

of women have reported having experienced online abuse or harassment at least once in their life, and that 1 in 10 women has experienced some form of online violence since the age of 15.

Studies have shown that women of color are 34 per cent more likely to be the targets of harassment than white women. Black women are targeted most of all: One in every 10 tweets sent to them was abusive or problematic, whereas for white women it was one in 15. Although abuse is targeted at women across the political spectrum, women of color were much more likely to be impacted and women of African descent are disproportionately targeted. Twitter's failure to crack down on this problem means it is contributing to the silencing of already marginalized voices.

These phenomena are so worrisome that it has raised the concern of the United Nations High Commissioner for Human Rights. On 5 November 2022, Mr. Volker Türk issued an open letter to you, as the new owner of Twitter, emphasizing the responsibilities involved in managing such an influential platform in terms of respecting human rights and addressing any adverse human rights impact related to it. While stressing the need to uphold freedom of expression worldwide, The High Commissioner further stressed that freedom of expression stops at hatred to incite discrimination, hostility and/or violence.

Without prejudging the accuracy of these allegations, we express our deep concern that the business model of social media platforms provides an unrestrained and fertile space for racism against people of African descent, with women of African Descent and other minorities experiencing the brunt of both racism and sexism. Digital technologies are being deployed to advance xenophobic and racially discriminatory treatment and exclusion of migrants, refugees, and stateless persons. Of particular concern is that hate speech and incitement to hatred not only are able to harm the mental health of those affected but can also result in further discrimination, particularly in the enjoyment of human rights, violence and physical harm to people and, in the worst cases, to death. Broadly speaking, online racism not only affects targeted groups of people, such as people of African descent and minorities, but exercises greater influence on society at large, emphasizing divisions, exacerbating fractures and strengthening polarization within the society. Judges and other legal operators who play a role in protecting human rights have also become the targets of racist online harassment and attacks. The above-mentioned elements become more meaningful when we consider the rising importance that young people attach to cyberspace, and the latter's potential to influence their choices and values. We regret to note how much white supremacy and misogyny are interlinked and embedded in the culture of many countries. We would also like to express serious concerns at the existing gender bias in some AI systems and applications that contributes to gender-based discrimination and violence. In view of the transformative power of social media to advance gender equality and combat any form of discrimination, your company has a particular responsibility to take all necessary measures to ensure it operates in compliance with the UN Guiding Principles on Business and Human Rights.

We urge you to put international human rights standards at the centre of your business model, including in the management of content on your platforms. This can be done through carrying out regular human rights impact assessments of the

products, defining and prohibiting hate speech on your platform, in accordance with international human rights standards, adopting policies to counter hate speech based on human rights law while upholding international standards freedom of expression, of peaceful assembly and of association. We underline the importance of contextual analysis and evaluation in the context of content moderation, and the importance to allocate sufficient capacity in this regard, including through human oversight and expertise, including linguistic and cultural expertise, and the involvement of communities most affected by hate speech, including people of African descent, minorities, women and girls.

As per the United Nations Strategy and Plan of Action on Hate Speech, all efforts to address and respond to hate speech should integrate an intersectional gender analysis and identify gender-based hate speech as well as the gendered elements in broader hate speech trends. The promotion and protection of gender equality and women's and girls' rights must be central to all efforts to prevent gender-based hate speech. Responses aimed at preventing gender-based hate speech should be holistic, address the root causes and drivers of gender inequality and the increasing backlash against women's and girls' rights and be developed alongside those most affected by gender-based hate speech.

Legal provisions, frameworks and policies to combat gender-based hate speech and its intersections with other forms of hate speech should be implemented and strengthened. Comprehensive policies should include establishing safe mechanisms for persons to report hate speech, harassment and incitement to violence and to be able to access effective support and gender-responsive remedy, without fear of discrimination and further victimization, and system to record and publish statistics on these violations, disaggregated by age, gender and all prohibited grounds of discrimination.

Social media companies should contribute to facilitate the exercise of the rights of expression, association and peaceful assembly, to public advocacy and education campaigns, to advance respect and non-discrimination and combat the root causes of racist and gender-based hate speech, including discriminatory social norms and harmful racial and gender stereotypes, and to actively partner and show public support for feminists and women human rights defenders working on all anti-discrimination issues. Gender-based hate speech, including misogyny, combined with direct discrimination against women, girls and gender-diverse people through laws and policies, are dominant factors in censoring and restricting women's freedom of expression. Growing trends of populism, nationalism and fundamentalism worldwide have further accentuated patriarchy and misogyny and enhanced discrimination against women and girls, as well as the suppression of their ability to express themselves.

Hate speech based on gender and misogyny must be addressed within the international human rights framework. As racist and gendered hate speech proliferates on social media platforms, there are increasing calls to prohibit or criminalize such hate speech. Addressing hate speech does not mean limiting or prohibiting the rights to participate, to access information, to speak out and to mobilize. It means keeping speech from escalating into incitement to discrimination, hostility or violence, which is prohibited under international law. In line with the United Nations Guiding Principles on Business and Human Rights, as a matter of due diligence, social media companies should carry out regular human rights and gender impact assessments to

identify and mitigate systemic risks affecting people of African descent and minorities, including women and girls.

From a human rights perspective, the phenomenon of violence against women and girls particularly those who are also member of a minority is facilitated by new technologies and digital spaces. Acts of online violence may force women to retreat from the Internet. Research indicates that 28 per cent of women who had suffered ICT-based violence intentionally reduced their presence online. Other common outcomes are social isolation, whereby victims or survivors withdraw from public life, including with family and friends, and limited mobility, when they lose their freedom to move around safely. We would also like to emphasise that the development of new social media digital spaces is especially critical for new generations of girls and boys, who are starting their lives extensively using new technologies to mediate in their relationships, affecting all aspects of their lives. As such, we express grave concerns regarding the manifest amplification of a phenomenon of intersecting violence perpetrated against women and girls of African descent on online platforms, including Twitter.

Furthermore, there is an increasing recognition of the deep impact of systemic oppressions on mental health.

Electronic communication services, social media platforms and search engines provide an ideal environment both for the anonymous delivery of targeted threats, sexual harassment and extortion and for the mass dissemination of intimidating, defamatory, degrading, deceptive or discriminatory narratives. Individuals or groups systematically targeted by cyber-harassment are generally left without any effective means of defence, escape or self-protection and, at least in this respect, often find themselves in a situation of “powerlessness” comparable to physical custody. Depending on the circumstances, the physical absence and anonymity of the perpetrator may even exacerbate the victim’s emotions of helplessness, loss of control and vulnerability, not unlike the stress-augmenting effect of blindfolding or hooding during physical torture. Likewise, the generalized shame inflicted by public exposure, defamation and degradation can be just as traumatic as direct humiliation by perpetrators in a closed environment. As various studies on cyberbullying have shown, harassment alone in comparatively limited environments can expose targeted individuals to extremely elevated and prolonged levels of anxiety, stress, social isolation and depression and significantly increases the risk of suicide, which may amount to psychological torture. (Report of the Special Rapporteur on Torture A/HRC/43/49, paras 74-75).

In connection with the above-alleged facts and concerns, please refer to the **Annex on Reference to international human rights law** attached to this letter, which cites international human rights instruments and standards relevant to these allegations.

It is our responsibility, under the mandates provided to us by the Human Rights Council, to seek to clarify all cases brought to our attention. In this regard, we would be grateful to have your cooperation and comments on the following matters:

1. Please provide any additional information and/or comment(s) you may have on the above-mentioned allegations.
2. Please indicate specific remedial measures that your company has taken or is considering taking to prevent racialized hate targeting people of African descent, including women and girls belonging to this group who suffer the brunt of sexism and racism.
3. Please explain how your company addresses the spread of White supremacist, anti-Black and misogynistic content on social media platforms, and please clarify whether any inclusive appeal systems for moderation and remove abusive contents have been included in your policy.
4. Please describe the measures that your company has taken or is considering taking to screen advertisements and/or posts that disseminate disinformation causing polarization and racial discrimination within society, while ensuring that these measures meet international human rights standards, including for the promotion and protection of freedom of expression.
5. Please enumerate research and studies that you are conducting to assess the intended and unintended impacts that your platforms may have on people of African descent.
6. Please share any measure taken by your company in the design and management of your platform to address the existing or potential existence of racial and gender bias in AI that contributes to racial and gender-based discrimination and violence.
7. Please provide any additional information on regulation and prevention, particularly Community Guidelines or self-regulation, content moderation and integrity efforts that may be relevant, including any public human rights assessments carried out by your company (internally or by third parties) in the development of products/services/platforms, as well as internal or external bodies that are tackling these issues.
8. Please indicate how international standards for freedom of expression, association and peaceful assembly are protected and guaranteed in your content moderation policies and other policies aimed at countering hate speech and disinformation.
9. Please provide information as to what human rights due diligence steps, as set forth in the UN Guiding Principles on Business and Human Rights, have been undertaken by your company to identify, prevent, mitigate and account for human rights abuses caused by or contributed to your own activities, or directly linked to your operations,

products or services by your business relationships (including abroad). This includes the adverse impact of the services described in this letter.

10. Please provide information on whether your company has put in place human rights training for personnel in relevant business functions.
11. Please provide information on steps taken by your company to establish, or participate in effective operational-level grievance mechanisms, in line with the UN Guiding Principles on Business and Human Rights, to address adverse human rights impacts caused by your company throughout your operations globally.
12. Please explain what monitoring and evaluation systems you have in place to ensure the effectiveness of actions taken to mitigate and prevent the allegations mentioned in this letter.

This communication and any response received will be made public via the communications reporting [website](#) within 60 days. They will also subsequently be made available in the usual report to be presented to the Human Rights Council.

While awaiting a reply, we urge that all necessary interim measures be taken to halt the alleged violations and prevent their re-occurrence and in the event that the investigations support or suggest the allegations to be correct, to ensure the accountability of any person(s) responsible for the alleged violations.

We may publicly express our concerns in the near future as, in our view, the information upon which the press release will be based is sufficiently reliable to indicate a matter warranting immediate attention. We also believe that the wider public should be alerted to the potential implications of the above-mentioned allegations. The press release will indicate that we have been in contact with your company to clarify the issue/s in question.

Please note that letters expressing similar concerns will also be sent to the chief executive officers of other relevant social media companies.

Please accept, Mr. Elon Musk, the expression of our most distinguished consideration.

Catherine Namakula
Chair-Rapporteur of the Working Group of Experts on People of African Descent

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

Ian Fry
Special Rapporteur on the promotion and protection of human rights in the context of climate change

Alexandra Xanthaki
Special Rapporteur in the field of cultural rights

Saad Alfarargi
Special Rapporteur on the right to development

Farida Shaheed
Special Rapporteur on the right to education

Michael Fakhri
Special Rapporteur on the right to food

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

Clement Nyaletsossi Voule
Special Rapporteur on the rights to freedom of peaceful assembly and of association

Margaret Satterthwaite
Special Rapporteur on the independence of judges and lawyers

Paula Gaviria
Special Rapporteur on the human rights of internally displaced persons

Livingstone Sewanyana
Independent Expert on the promotion of a democratic and equitable international
order

Obiora C. Okafor
Independent Expert on human rights and international solidarity

Ravindran Daniel Justin
Chair-Rapporteur of the Working Group on the use of mercenaries as a means of
violating human rights and impeding the exercise of the right of peoples to self-
determination

Felipe González Morales
Special Rapporteur on the human rights of migrants

Fernand de Varennes
Special Rapporteur on minority issues

K.P. Ashwini
Special Rapporteur on contemporary forms of racism, racial discrimination,
xenophobia and related intolerance

Fionnuala Ní Aoláin
Special Rapporteur on the promotion and protection of human rights and fundamental
freedoms while countering terrorism

Reem Alsalem
Special Rapporteur on violence against women and girls, its causes and consequences

Dorothy Estrada-Tanck
Chair-Rapporteur of the Working Group on discrimination against women and girls

Annex

Reference to international human rights law

In connection with the above-alleged facts and concerns, we would like to draw your attention to the applicable international human rights norms and standards, as well as authoritative guidance on their interpretation.

We would like to highlight the United Nations Guiding Principles on Business and Human Rights, which were unanimously endorsed in 2011 by the Human Rights Council in its resolution (A/HRC/RES/17/31) after years of consultation with governments, civil society and the business community. The Guiding Principles have been established as the authoritative global standard for all States and businesses to prevent and address business-related adverse human rights impacts. These Guiding Principles are based on the recognition of:

- a) "The existing obligations of States to respect, protect and fulfil human rights and fundamental freedoms;
- b) The role of business enterprises as specialised bodies or corporations performing specialised functions, which must comply with all applicable laws and respect human rights;
- c) The need for rights and obligations to be matched by appropriate and effective remedies when they are violated".

The Guiding Principles also make clear that companies have an independent responsibility to respect human rights. Principles 11-24 and 29-31 provide guidance to companies on how to meet their responsibility to respect human rights and to provide remedies where they have caused or contributed to adverse impacts. The Guiding Principles have identified two main components of the corporate responsibility to respect human rights, which require "business enterprises to:

- a) Prevent their own activities from causing or contributing to adverse human rights impacts and address those impacts when they occur;
- b) Seek to prevent or mitigate adverse human rights impacts directly related to operations, products or services provided through their business relationships, even where they have not contributed to those impacts". (Guiding Principle 13)

The commentary to Guiding Principle 13 notes that companies can be affected by adverse human rights impacts, either through their own activities or as a result of their business relationships with other parties (...) The 'activities' of business enterprises are understood to include both actions and omissions; and their 'business relationships' include relationships with business partners, entities in their value chain and any other non-State or State entities directly linked to their business operations, products or services.

To meet their responsibility to respect human rights, companies should have in place policies and procedures appropriate to their size and circumstances:

- a) A political commitment to uphold their responsibility to respect human rights;
- b) A human rights due diligence process to identify, prevent, mitigate and account for how they address their human rights impact;

- c) Processes to redress any adverse human rights impacts they have caused or contributed to (Guiding Principle 15).

According to Guiding Principles 16-21, human rights due diligence involves:

- a) Identifying and assessing actual or potential adverse human rights impacts that the enterprise has caused or contributed to through its activities, or that are directly related to the operations, products or services provided by its business relationships;
- b) Integrate the results of impact assessments into relevant business functions and processes, and take appropriate action in accordance with their involvement in the impact;
- c) Monitor the effectiveness of the measures and processes adopted to address these adverse human rights impacts in order to know whether they are working;
- d) Communicate how adverse effects are addressed and demonstrate to stakeholders - particularly those affected - that appropriate policies and processes are in place to implement respect for human rights in practice.

This process of identifying and assessing actual or potential adverse human rights impacts should include substantive consultation with potentially affected groups and other stakeholders (Guiding Principle 18).

Where an enterprise causes or is likely to cause an adverse human rights impact, it should take the necessary steps to end or prevent that impact. “The establishment of operational-level grievance mechanisms for those potentially affected by corporate activities can be an effective means of redress provided they meet certain requirements listed in Principle 31 (Guiding Principle 22).

Furthermore, business enterprises should remedy any actual adverse impact that they cause or to which they contribute. Remedies can take a variety of forms and may include apologies, restitution, rehabilitation, financial or non-financial compensation and punitive sanctions (whether criminal or administrative, such as fines), as well as the prevention of harm through, for example, injunctions or guarantees of non-repetition. Procedures for the provision of remedy should be impartial, protected from corruption and free from political and other attempts to influence the outcome (commentary to Guiding Principle 25).

The responsibility of business enterprises to respect human rights applies to all enterprises regardless of their size, sector, operational context, ownership and structure.

We would further like to draw your attention to article 19 of the International Covenant on Civil and Political Rights, which, while enshrining the right to freedom of expression, also stipulates that this right carries with it special duties and responsibilities, and may therefore be subject to certain restrictions, when provided by law and necessary for respect of the rights and reputation of others, as well as for the protection of national security or of public order (*ordre public*), or of public health or morals. The provision is complemented by Article 20 of the Covenant, which prohibits “racial or religious hatred that constitutes incitement to discrimination, hostility or violence”. The Human Rights Committee explained that articles 19 and

20 of the Covenant “are compatible with and complement each other”, Human Rights Committee, general comment No. 34 (2011), para. 50.

We wish to highlight also the provisions of the International Convention on the Elimination of All Forms of Racial Discrimination, in particular Article 2 and Article 5, calling upon the eradication of “incitement to, or acts of” racial discrimination, with due regard to other rights protected by human rights law, including the freedom of expression. Furthermore, Article 4 of the Convention expressly condemns dissemination of ideas of racial superiority. General comment No. 35 to the International clarifies that, although the term “hate speech” is not mentioned in the Convention, “[t]he drafters of the Convention were acutely aware of the contribution of speech to creating a climate of racial hatred and discrimination, and reflected at length on the dangers it posed”. The mentioned General Comment also recalls that “the right to freedom of expression is not unlimited but carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but only if they are provided by law and are necessary for the protection of the rights or reputations of others and for the protection of national security or of public order, or of public health or morals. Freedom of expression should not aim at the destruction of the rights and freedoms of others, including the right to equality and non-discrimination”.

We would also like to make reference to the Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence, known as Rabat Plan of Action and issued on October 5, 2012, which represents a useful guidance in distinguishing between permissible speech and speech that may amount to incitement. The Rabat Plan also addresses the role of all stakeholders, reminding that “self-regulation, where effective, remains the most appropriate way to address professional issues relating to the media. In line with principle 9 of the Camden Principles, all media should, as a moral and social responsibility and through self-regulation, play a role in combating discrimination and promoting intercultural understanding, including by considering the following:

- a) Taking care to report in context and in a factual and sensitive manner, while ensuring that acts of discrimination are brought to the attention of the public.
- b) Being alert to the danger of furthering discrimination or negative stereotypes of individuals and groups in the media.
- c) Avoiding unnecessary references to race, religion, gender and other group characteristics that may promote intolerance.
- d) Raising awareness of the harm caused by discrimination and negative stereotyping.
- e) Reporting on different groups or communities and giving their members the opportunity to speak and to be heard in a way that promotes a better understanding of them, while at the same time reflecting the perspectives of those groups or communities” (para. 58).

We would also like to draw our attention to the report of the Special Rapporteur on minority issue hate speech, social media and minorities (A/HRC/46/57), issued in March 2021. The report specifically addresses the targeting of minority women, including persons of African-descent in social media, and the obligations of private companies and other actors:

84. The Special Rapporteur also invites States and other interested parties to refer to existing guidance on what constitutes forms of expression that may be restricted under international human rights law in the area of impermissible hate speech, in particular the six-part threshold test in the Rabat Plan of Action and principle 12 of the Camden Principles on Freedom of Expression and Equality.

87. The Special Rapporteur invites State agencies and social media platforms to put in place in procedures and mechanisms for the mandatory collection of data on hate speech, and at a minimum on incitement to genocide and advocacy that constitutes incitement to discrimination, hostility or violence. Such data should be disaggregated according to the basis of the hatred, whether national, ethnic, racial or religious hatred, which all relate to vulnerable minorities, or hatred affecting other vulnerable groups, on the basis of such factors as gender and sexual orientation.

88. The Special Rapporteur recommends that national legislation be in place that requires social media platforms to adopt policies and protocols to identify hate speech on the grounds of national origin, ethnicity, religion and language, as well as other factors such as gender and sexual orientation, while respecting the right to privacy of individuals.

89. The Special Rapporteur encourages social media companies to manage hate speech on their platforms with reference to the human rights implications of their products, including algorithms and artificial intelligence programs such as chatbots. They must therefore have in place human rights review processes that refer to and focus specifically on the groups most susceptible to hate speech in the States concerned, and in particular minorities, women and lesbian, gay, bisexual, transgender and intersex communities.

92. The Special Rapporteur emphasizes that social media's content moderation systems and community standards and any oversight or appeal entity should clearly commit to protecting vulnerable and marginalised minorities and other groups. Minorities should specifically be identified as priorities for social media platforms.

93. The Special Rapporteur invites owners of social media platforms to apply the Guiding Principles on Business and Human Rights in their own operations. Human rights standards should be systematically integrated into the content policies and decision mechanisms of social media platforms, so that, as specified in the Guiding Principles, their owners “comply with all applicable laws and respect internationally recognized human rights wherever they operate”, and “treat the risk of causing or contributing to gross human rights abuses as a legal compliance issue wherever they operate”, for which they could be liable.

96. The Special Rapporteur encourages States and social media platforms to comprehensively address distortion and systemic bias against Jews and Muslims, as evidence suggests that antisemitism and Islamophobia are pressing challenges for minority rights.

98. In order to improve mechanisms and content vetting policies for the handling of hateful content, and to ensure incorporation of the concerns of the main targets of hate speech in social media, the Special Rapporteur urges that minorities, as the most targeted and vulnerable groups, be represented in advisory and other relevant boards.

100. The Special Rapporteur encourages all stakeholders to look into innovative, educational and preventive strategies that focus on protection of and respect for diverse communities in order to counter hate speech.

The Special Rapporteur on minority issues also noted in his August 2022 country mission report to the United States (A/HRC/49/46/Add.1) that “African Americans are among the most marginalized minorities in the country in socioeconomic terms, are by far the most likely to be denied the right to vote in federal and State elections, to be incarcerated, to be the targets of hate speech on social media and to be disproportionately discriminated against”, that “African Americans are reported to be the minority most affected by hate crimes and hate speech”, and that “that minority women are particularly vulnerable to some of the most violent and dangerous forms of hate speech on social media”

We would also like to draw your attention to the report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression on online hate speech (A/74/486), issued on October 9, 2019. The report provides for a set of recommendations, in addition to the principles adopted in earlier reports and in keeping with the Guiding Principles on Business and Human Rights, to private companies, especially to those active ICT sector:

- a) Evaluate how their products and services affect the human rights of their users and the public, through periodic and publicly available human rights impact assessments;
- b) Adopt content policies that tie their hate speech rules directly to international human rights law, indicating that the rules will be enforced according to the standards of international human rights law, including the relevant United Nations treaties and interpretations of the treaty bodies and special procedure mandate holders and other experts, including the Rabat Plan of Action;
- c) Define the category of content that they consider to be hate speech with reasoned explanations for users and the public and approaches that are consistent across jurisdictions;
- d) Ensure that any enforcement of hate speech rules involves an evaluation of context and the harm that the content imposes on users and the public, including by ensuring that any use of automation or artificial intelligence tools involve human-in-the-loop;
- e) Ensure that contextual analysis involves communities most affected by content identified as hate speech and that communities are involved in identifying the most effective tools to address harms caused on the platforms;
- f) As part of an overall effort to address hate speech, develop tools that promote individual autonomy, security and free expression, and involve de-amplification, de-monetization, education, counter-speech, reporting and training as alternatives, when appropriate, to the banning of accounts and the removal of content.

In a report A/74/486 to the UN General Assembly, the then Special Rapporteur freedom of expression provides important guidance on how internet companies can address hate speech on their platforms, while respecting international standards for freedom of expression. Companies should evaluate how their products and services affect the human rights of their users and the public, through periodic and

publicly available human rights impact assessments. Furthermore, companies should adopt content policies based on international human rights law and enforcing the rules in line with international human rights standards, including relevant UN treaties, guidance from treaty bodies and special procedures mandate holders and the Rabat Plan of Action. Furthermore, it is crucial that any enforcement of hate speech rules involves an evaluation of context and the harm that the content imposes on users and the public, including by ensuring that any use of automation or artificial intelligence are subject to human oversight to minimize the risk of legitimate content being removed. Communities most affected by content identified as hate speech must be involved in the contextual analyses and in identifying tools to address harms caused on the platforms.

In her report A/HRC/47/25, the Special Rapporteur freedom of expression called on social media companies to, in line with the Guiding Principles on Business and Human Rights, review their business models and ensure that their business operations, data collection and data processing practices are compliant with international human rights standards, including article 19 of the International Covenant on Civil and Political Rights, as well as data protection principles and relevant national consumer protection standards. They should also conduct human rights impact assessments of their products, particularly of the role of algorithms and ranking systems in amplifying disinformation or misinformation. Such assessments should be conducted regularly and ahead of and following significant events such as national elections or major crises like the COVID-19 pandemic.

In a report to the General Assembly (A/76/258), the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression has observed that in the digital age, the spate of online violence, hate speech and disinformation often compel women to self-censor, limit what they post or leave platforms and that social media platforms have failed to respond adequately to the risks and dangers of online violence, hate speech and disinformation that women confront in the digital space. The Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression has reminded social media companies that digital spaces are owned and managed by private actors, but they are public spaces accessed by millions of people and as such, taking into account the nature of their business, social media platforms should be guided by international human rights standards in their content moderation. Furthermore, in line with the United Nations Guiding Principles on Business and Human Rights, as a matter of due diligence companies should carry out regular human rights and gender impact assessments to identify and mitigate systemic risks affecting women and gender nonconforming people. In addition, guided by international human rights standards and the principles of gender equality, they should take the same coordinated approach to make digital spaces safe and gender inclusive. Policies on safety from online violence should be developed and made available, with full transparency in relation to algorithms, practices and decision-making processes, in an accessible, non-technical jargon in local languages.

The Special rapporteur on the rights to freedom of peaceful assembly and of association noted that international human rights norms should guide digital technology companies' governance and further recommend them "to prevent or mitigate the adverse human rights impacts of their involvement" (A/HRC/41/41, para 84). In the same report, he called them to "integrate early warning systems within business processes to identify human rights risks, and respond in a timely fashion

(...); support the research and development of appropriate technological solutions to online harassment, disinformation and propaganda, including tools to detect and identify State-linked accounts and bots (A/HRC/41/41, para 86 (b)).

In a report to the Human Rights Council (A/HRC/23/50), the Working Group on discrimination against women and girls has observed that given the complex obstacles deeply rooted in structural inequalities and multiple intersections of social-cultural exclusion, women's effective participation in political and public life can be achieved only when they have equal and full access to all mechanisms of independent human rights monitoring and redress, particularly in contexts where the risks are high. The Working Group has recommended to create enabling conditions for public recognition and acceptance for women by providing a positive image of diverse women, including minority women, indigenous women, women with disabilities, and other historically marginalized women (A/HRC/23/50). Online violence against women and girls from a human rights perspective was the focus of the report of the Special Rapporteur on violence against women and girls. The Special Rapporteur recommended member States to States should, in accordance with the principle of due diligence, enact new laws and measures to prohibit new emerging forms of online gender-based violence. Such laws should be grounded in international women's human rights law and standards, as outlined in the Convention on the Elimination of All Forms of Discrimination against Women (with due consideration for general recommendations Nos. 19 and 35 of the Committee on the Elimination of Discrimination against Women) and the Declaration on the Elimination of Violence against Women, and in other global and regional women's human rights instrument, such as the Inter-American Convention on the Prevention, Punishment and Eradication of Violence against Women (Belém do Pará Convention), the Council of Europe Convention on Preventing and Combating Violence against Women and Domestic Violence and the Protocol to the African Charter on Human and Peoples' Rights on the Rights of Women in Africa. Furthermore, States should ensure that their legal frameworks adequately protect all women's human rights online, including the right to life free from violence, freedom of expression and access to information, and the right to privacy and data protection (A/HRC/38/47).