

# Oversight Board Submission to the Special Rapporteur on Freedom of Opinion and Expression: Challenges in Times of Conflicts and Disturbances

July 2022

## I. Introduction

The Oversight Board thanks the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression for the opportunity to address challenges of social media companies respecting human rights when governing content relating to conflict situations. The [Oversight Board](#) (“the Board”) was created to improve how Meta treats people and communities around the world by ensuring respect for free expression, through independent judgement. The Board is made up of 23 members of diverse disciplines and backgrounds from around the world. They include academics, civil society leaders, former judges and mandate-holders from the UN and regional human rights bodies. The Board’s charter and bylaws were informed by public consultation and a [human rights review](#) conducted by the organization Business for Social Responsibility (BSR).

People can appeal content decisions made by Meta to the Board, including both decisions to remove content and decisions to leave content up. Meta can also refer cases to the Board. The Board selects significant and difficult cases from around the world to determine whether Meta acted in accordance with its policies, values and human rights responsibilities. The Board issues binding decisions and makes recommendations to Meta. The Board also issues [policy advisory opinions](#), through which it reviews Meta’s policies at the company’s request, and makes recommendations for how they should be improved. While the Board’s recommendation power is non-binding, the company is required to respond publicly to them.

From October 2020 to December 2021, the Board received over 1.1 million appeals. More than 8 in 10 user appeals to restore content concerned Meta’s rules on bullying, hate speech or violence and incitement. To date, the Board has issued 25 case decisions, many of which related to conflict situations, as well as other crises or forms of civil unrest, including coups, riots, and government suppression of protests. The Board believes these cases and recommendations shine a light on the challenges social media companies like Meta face when moderating content from conflict regions. This submission provides an overview of the Board’s cases and recommendations, as well as Meta’s responses to them.

## II. Respecting human rights in content governance related to conflict and civil unrest

Through its cases, the Board has made numerous recommendations to bring Meta’s policies and their enforcement into greater compliance with the company’s values and human rights responsibilities.

### a) The need for a comprehensive crisis protocol

The Board has stressed the importance of Meta developing a crisis protocol to ensure the company is prepared to respond to crises and novel situations consistently across the globe, with an emphasis on prioritizing situations where Meta’s adverse human rights impacts are most severe.

i. Relevant Oversight Board Decisions

On May 5, 2021, the Board published its decision on [Former President Trump’s Suspension](#). Following the riot at the US Capitol on January 6<sup>th</sup>, 2021, Meta indefinitely suspended Former President Trump from its platforms for two posts shared on his official pages that violated the Dangerous Individuals and Organizations (DIO) Community Standard. The Board found that the two posts constituted praise or support of people engaged in violence, and that in maintaining an unfounded narrative of electoral fraud and persistent calls to action, Mr. Trump created an environment where there was a serious risk of violence. The posts served to legitimize the ongoing violent actions of the January 6 rioters. In issuing its decision, the Board referred to the UN Rabat Plan of Action and 2020 Joint Statement of international freedom of expression mandate-holders to emphasize the duty of state actors to condemn violence and to “provide accurate information to the public on matters of public interest, while also correcting misinformation.” The Board also stressed the need to develop less restrictive means of regulating such speech, including by “developing effective mechanisms to avoid amplifying speech that poses risks of imminent violence, discrimination, or other lawless action, where possible and proportionate.”

On June 13, 2022, the Board published its decision in the [Sudan graphic video](#) case. The case concerned the removal of a video depicting a dismembered body posted following the military coup in Sudan in October 2021 and the start of protests (for more detail, see section (d) below). In the decision, the Board stressed the importance of the company being “prepared to respond quickly and systematically to conflicts and crisis situations around the world” and condemned the slow progress the company had made on developing and publishing a crisis protocol.

The Board is currently considering a case from Ethiopia, [Tigray Communication Affairs Bureau](#), that raises similar issues of the company’s policies and processes in conflict situations. This case decision will be published in August 2022.

ii. The Board’s Recommendations and Meta’s responses

In the Former President Trump’s Suspension decision, the Board recommended that Meta “[d]evelop and publish a policy that governs Facebook’s response to crises or novel situations where its regular processes would not prevent or avoid imminent harm. This guidance should set appropriate parameters for such actions, including requirement to review its decision within a fixed time.”

In its [response](#), Meta agreed to develop a Crisis Policy Protocol “which will be informed by various frameworks that we use to address risk, imminent harm, and integrity challenges. The protocol

will focus on the threshold for when context specific policies are deployed, deactivated, and reassessed.” As of July 1<sup>st</sup>, 2022, the company has yet to publish the crisis protocol. In its [Q4 2021 Update](#), Meta reported that the company has prepared a proposal for a new Crisis Protocol in response to the Board’s recommendation and that the protocol was adopted. Meta also stated that it would soon provide information on this protocol on its Transparency Center.

**b) The need for human rights due diligence in conflict and crisis situations**

The Board has highlighted the importance of human rights due diligence for how Meta’s platforms may have adverse impacts in armed conflict situations, as outlined in the UN Guiding Principles on Business and Human Rights, and the Working Group’s additional guidance on this issue in its 2020 report to the UN General Assembly ([A/75/212](#)).

**i. Relevant Cases**

The [Shared Al Jazeera post](#) decision concerned a user in Egypt who shared a post, in May 2021, by the news outlet Al Jazeera reporting on a statement made by Al-Qassam Brigades military spokesman. Al-Qassam Brigades, the military wing of the Palestinian group Hamas, and their spokesperson are both designated dangerous under Meta’s DIO policy. The DIO policy prohibits praise, support or representation of designated entities but does allow reporting or neutral discussion on them. The list of designated individuals and organizations is not public, though the Oversight Board recommended in its [Nazi Quote](#) decision that Meta make the list public or provide a list of examples. The post related to the May 2021 armed conflict between Israeli forces and Palestinian militant groups in Israel and Gaza. The conflict broke out after weeks of rising tensions and protests in Jerusalem tied to a dispute over ownership of homes in Sheikh Jarrah neighborhood of East Jerusalem. Following the raid by Israeli forces of a Mosque during Ramadan prayers, which injured hundreds of worshippers, Al-Qassam Brigades issued an ultimatum demanding that Israeli soldier withdraw. Once the deadline had expired, Al-Qassam Brigades and other militant groups in Gaza launched rockets at the civilian center of Jerusalem, which began 11 days of armed conflict. The Board found the content should not have been removed because it neutrally discussed a designated entity, noting the public interest value of sharing reporting on a threat and ensuring access to information for people in the region.

In its decision, the Board noted numerous public comments that highlighted that Facebook is among the primary means that Palestinians have to communicate news and opinion and express themselves freely. The Board took note of the severe limitations on freedom of expression in territories governed by the Palestinian Authority and Hamas ([A/75/532](#), para. 25) and by the Israeli government (Working Group on the Universal Periodic Review, [A/HRC/GW.6/29/ISR/2](#), paras. 36-37).

Finding that Meta has a heightened due diligence responsibility in conflict-affected regions (A/75/212, para 13), particularly to respect non-discrimination, the Board called attention to public comments and publicly available information alleging that Facebook has disproportionately removed or demoted content from Palestinian people and content in the Arabic language, especially in comparison to its treatment of anti-Palestinian or anti-Arab posts within Israel.

In the case of [Alleged Crimes in Raya Kobo](#), the Board again reiterated that Meta has a heightened due diligence responsibility in conflict-affected regions in order to protect the right to life. The case concerned a post alleging civilian aid by ethnic Tigrayans in Raya Kobo in the commission of crimes by Tigray forces during the ongoing armed conflict in Ethiopia (for more details about the case, see part (c) below).

ii. The Board’s recommendations and Meta’s responses

In the Shared Al Jazeera post decision, the Board recommended that Meta “[e]ngage an independent entity not associated with either side of the Israeli-Palestinian conflict to conduct a thorough examination to determine whether Facebooks’ content moderation in Arabic and Hebrew, including its use of automation, have been applied without bias.” In its [response](#), Meta said it had engaged BSR to perform human rights due diligence in line with this recommendation and to publish the outcome in the first quarter of 2022. In a more [recent update](#), Meta stated that the due diligence report is still underway and that they “hope to publicly communicate insights from this work by Q3 2022.”

In the Alleged Crimes in Raya Kobo case, the Board recommended that Meta “commission an independent human rights due diligence assessment on how Facebook and Instagram have been used to spread hate speech and unverified rumors that heighten the risk of violence in Ethiopia. The assessment should review the success of measures Meta took to prevent the misuse of its products and services in Ethiopia. The assessment should also review the success of measures Meta took to allow for corroborated and public interest reporting on human rights atrocities in Ethiopia. The assessment should review Meta’s language capabilities in Ethiopia and if they are adequate to protect the rights of its users.”

In [response](#), Meta stated that complying with all elements of the recommendation may not be feasible, because “methodologies are largely qualitative, and rights holders in conflict zones may have security or other concerns that inhibit their participation.” In its [Q1 2022 Quarterly Update](#) (p21-22), Meta also cited privacy and safety risks as a reason why the company cannot publish the full due diligence it has undertaken in Ethiopia, while saying nothing about an independent assessment.

In the Former President Trump’s Suspension decision, the Board recommended that Meta “Undertake a comprehensive review of Facebook’s potential contribution to the narrative of electoral fraud and the exacerbated tensions that culminated in the violence in the United States

on January 6. This should be an open reflection on the design and policy choices that Facebook has made that may allow its platforms to be abused.”

In [response](#), Meta stated that the company regularly “review [their] policies and processes in response to real world events” and that they have expanded their research initiative focused on the effect of their platforms on elections. Meta also said that it ultimately believes that “independent researchers and our democratically elected officials are best positioned to complete an objective review of these events.” According to [media reporting](#), the company has shifted resources away from election integrity efforts to focus more on the metaverse.

c) The need to address unverified rumors endangering the right to life in conflict situations

The Board highlighted the danger that unverified rumors pose in contributing to harmful narratives during an armed conflict driven by sectarian violence.

i. Relevant Case

The case of [Alleged Crimes in Raya Kobo](#) involved a post from a user in Ethiopia containing allegations that the Tigray People’s Liberation Front (TPLF), with the assistance of ethnic Tigrayan civilians, killed and raped women and children, and looted the properties of civilians in Raya Kobo and other towns in Ethiopia’s Amhara region. The Board began by identifying the tension between protecting freedom of expression and reducing the threat of sectarian conflict. This content was posted during ongoing civil and ethnic war in a region with a history of lethal ethnic conflict and in a country where Meta’s platforms play a key role in providing information about the conflict to the population. The Board also noted that it was not possible to verify the allegations made in the post during this period, given the communication blackout in the Amhara region of Ethiopia.

The Board held that removing the content was in line with the company’s human rights responsibilities as a business under UNGP Principle 13. And that “in a heated and ongoing conflict, unverified rumors may lead to grave atrocities, which the experience in Myanmar has indicated.” The Board noted the Ethiopian government officials’ conduct in instigating or spreading hate speech targeting Tigrayans and that “unverified rumors can feed into hateful narratives and contribute to their acceptance.” While unverifiable rumors may not individually lead to imminent harm, “when such content appears on an important, influential and popular social media platform during an ongoing conflict, the risk and likelihood of harm become more pronounced.” The Board also referred to the findings in *Nahimana* case, [Case No. ICTR-99-52-T](#), in support of the finding that “cumulative impact can amount to causation through a ‘gradual build-up of effect’.”

ii. The Board’s recommendations and Meta’s responses

In the Alleged Crimes in Raya Kobo case, the Board recommended that “Facebook’s Community Standards should reflect that in the context of war and violent conflict, unverified rumors pose

higher risk to the rights of life and security of persons. This should be reflected at all levels of the moderation process.”

In [response](#) to the Board’s recommendation, Meta stated that it would continue to work with its trusted partners and independent fact checkers to identify and remove misinformation that may contribute to the risk of imminent harm. However, Meta also stated that it did not agree that “the appropriate way to balance voice and safety is to remove more reports from conflict zones as “unverified rumors” when we have no signal from a trusted partner or a third-party fact-checker that those reports are false or could contribute to a risk of harm.... Especially in context of war and violent conflict, it is often not possible to verify information quickly. Removing everything that is unverified could lead to the removal of accurate claims by observers or victims of crimes against vulnerable people.”

d) The need to safeguard the use of social media to raise awareness of and document human rights abuses

The Board has recommended that Meta amend its policies to ensure content raising awareness of and documenting abuses can remain on the platform. It has reiterated that Meta has a responsibility to collect, preserve and, where appropriate, share information of content that has been properly removed that may provide evidence of violations of international criminal, human rights, and humanitarian law.

i. Relevant Cases

The [Sudan graphic video](#) was a case referred by Meta, involving a Facebook post which appeared to depict a civilian victim of violence in Sudan, which was posted following the military coup in the country on October 25, 2021. The video shows a person lying next to a car with a significant head wound and a visibly detached eye. A caption accompanying the video calls on people to stand together and not trust the military. According to the [UN High Commissioner for Human rights](#), following the military takeover of civilian government in Sudan and the start of civilian protests, security forces in the country fired live ammunition, used tear gas, and arbitrarily arrested and detained protesters. Security forces also targeted journalists and activists, severely restricting freedom of the press and access to information. Widespread and periodic internet shutdowns occurred at the start of the coup and during the months of unrest that followed.

The video was initially removed for violating the Violence and Graphic content policy, but then restored under the newsworthiness exception. The [newsworthiness exception](#) is a general exception that can be applied to all Community Standards to allow violating content to remain on the platform where the company judges the public interest value in the content to outweigh harms it may cause. Meta restored the post with an age-gated warning screen, though the Violence and Graphic Content policy does not allow for the use of a warning screen for video content of this kind. Unlike other policies, like on Hate Speech or Dangerous Individuals and Organizations, the

Violence and Graphic Content policy does not have a “raising awareness” exception that would cover the documentation of human rights abuses and violations. Meta informed the Board that it had documented 17 uses of the newsworthiness exception in connection to the Violent and Graphic Content policy over a 12-month period. By comparison, Meta had removed 90.7 million pieces of content under this community standard in the first three quarters of 2021. This disclosure, alongside similar disclosures about the newsworthiness allowance in the Board’s first Policy Advisory Opinion, goes to show how “exceptional” the newsworthiness exception is.

The Board held that the newsworthiness exception did not provide an adequate mechanism for preserving content of this nature on the platform and that, in order to avoid removing protected expression, the company should amend the Violent and Graphic Content policy itself to allow such content to remain on the platform. Regular content reviewers assessing content at-scale are not permitted to apply the newsworthiness exception, they must escalate content for additional review for the exception to be applied. However, as the Board stated, Meta does not provide clear criteria or process for escalation to facilitate the use of this policy as a regular part of the company’s scaled content moderation system. The Board stated that “in context of war or political unrest, there will be more graphic and violent content captured by users and shared on the platform for the purposes of raising awareness of or documenting abuses. This content is important for promoting accountability.”

In the [Former President Trump’s Suspension](#) case, the Board noted that “the removal of content or disabling of accounts, while potentially reducing the risk of harm, may also undermine accountability efforts, including by removing evidence.” The Board noted that Meta has a responsibility to “collect, preserve and, where appropriate, share information to assist in the investigation and potential prosecution of grave violations of international criminal, human rights and humanitarian law by competent authorities and accountability mechanisms.”

ii. The Board’s recommendations and Meta’s responses.

In the Sudan graphic video case, the Board recommended that “Meta...amend the Violent and Graphic Content Community Standard to allow videos of people and dead bodies when shared for the purpose of raising awareness of or documenting human rights abuses. This content should be allowed with a warning screen so that people are aware that content may be disturbing.” The Board recommended that Meta “undertake a policy development process that develops criteria to identify videos of people or dead bodies when shared for the purpose of raising awareness of or documenting human rights abuses.” Meta has not provided its response to the Board’s recommendation in the Sudan graphic video case at the time of submission, but those responses are due in August 2022.

In the Former President Trump’s Suspension decision, the Board recommended that Meta “make clear in its corporate human rights policy how it collects, preserves and, where appropriate, shares

information to assist in investigation and potential prosecution of grave violations of international criminal, human rights and humanitarian law.”

In [response](#) to the Board’s recommendation, Meta stated that “[i]nternational privacy laws create a layer of complexity to collecting, preserving, and sharing user content and/or personal information or personally identifiable information. These laws also contain requirements about, among other things, data storage and data deletion requirements, which must be considered before we are able to fully address the board's recommendation.” The company has not provided any further updates.

e) Need for transparency on government requests to remove content

The Board has noted the practice of government’s reporting content to silence dissent or criticism of the government and recommended that Meta formalize a more transparent process for receiving, responding to, and providing transparency around such government requests. This concern has arisen in a number of the Board’s cases addressing conflict situations or other crises, where governments may be perceived as attempting to shape the public perception of events by pressuring Meta to remove content.

i. Relevant Cases

In its [Öcalan’s Isolation](#) decision, the Board noted the limited transparency on government involvement in content removal. “While [Meta] includes statistics on government legal requests for the removal of content based on local law, it does not include data on content that is removed for violating the Community Standards after a government flagging the content.”

In the [Shared Al Jazeera post](#) case, the Board asked Meta questions on whether the company had received official or unofficial requests from Israel to remove content related to the April-May conflict. The question was prompted by allegations that Facebook had censored Palestinian content due to Israeli government demands. The company responded that it “has not received a valid legal request from a government authority related to the content the user posted in this case, Facebook declines to provide the remaining requested information.”

In the case of [Punjabi Concern over the RSS in India](#), the Board asked Meta about possible communications from Indian authorities to restrict content around the farmer’s protests, specifically content critical of the government over its treatment of farmers, or content concerning the protests. The company declined to answer the Board’s question. The company explained that its moderation in India is independent of government influence, that “its staff receive training specific to their region, market, or role as part of the Global Ethics and Compliance initiative, which fosters a culture of honesty, transparency, integrity, accountability and ethical values. Further, Facebook’s staff are bound by a Code of Conduct and an Anti-Corruption Policy.” In emphasizing the importance of reviewing and auditing content moderation processes, the Board said that such



“assessments should take into account the potential for coordinated campaigns by government and non-state actors to maliciously report dissent.”

ii. The Board’s Recommendations and Meta’s responses

In its Öcalan Isolation decision, the Board recommended that Meta should “include information on the number of requests Facebook receives for content removals from governments that are based on Community Standards violations (as opposed to violations of national law), and the outcome of those requests.” The Board also said that the company should “ensure users are notified when their content is removed. The notification should note whether the removal is due to a government request or due to a violation of the Community Standards or due to a government claiming a national law is violation (and the jurisdictional reach of any removal.)”

In its [response](#), Meta stated that if content violates its Community Standards, the company will remove it and notify the user. However, because “these reports are reviewed under a standardized process in the same way and against the same policies as reports from any other source, we are not currently able to provide a different notice based on the source of the report. In addition, we may receive reports of a piece of content from multiple sources at the same time—for example, from a government and from user reports on Facebook. Such situations create additional challenges in determining whether content should be considered as removed in response to a government report.” The company cited the same challenges in providing greater transparency in its public reporting on enforcement. If the content is removed due to a national law, the company stated that the user is notified of this unless the company is legally prohibited from doing so.

In the Shared Al Jazeera Post decision, the Board recommended that Meta “formalize a transparent process on how it receives and responds to all government requests for content removal, and ensure that they are included in transparency reporting. The transparency reporting should distinguish government requests that led to removals for violations of the Community Standards from requests that led to removal or geo-blocking for violating local law, in addition to requests that led to no action. ”

In [response](#), Meta said that when a request to remove or restrict content is made by a government, the company first assesses the content for compliance with its policies. If the content violates the policies, the company removes it. If the content does not violate Meta’s policies, then Meta conducts a legal review, including a human rights due diligence review, and “may restrict access to the content in the jurisdiction where it has been reported as unlawful.” While the company publishes [a report](#) detailing these requests and instances of removal, the report does not include content flagged by government authorities and removed for violating Meta’s policies. Meta also explained that “In some cases, governments, law enforcement agencies, or those acting on their behalf may report content in ways that do not allow us to clearly identify them as such, or identify whether they are acting in an official capacity. For example, a government official or law enforcement officer may use our in-product reporting tools to report content in the same way as

any Facebook user. We are not able to distinguish these reports, as they are treated in the same way as any other user report.”

f) The need for greater transparency on enforcement of content policies, broken down by country and language

The Board has noted the importance of ensuring minority and opposition voices are not silenced due to mistakes or inadequate enforcement systems. [Transparency on enforcement](#) of content policies that provide more granular removals data broken down by region and language can help the public evaluate whether enforcement processes lead to greater mistakes in some regions or for some languages than others. The impact of enforcement errors on the sharing of important information during conflicts or other crises has been a recurring concern for the Board.

i. Relevant Cases and Context

The case of [Punjabi Concern over the RSS in India](#) concerned a post from Punjabi-language online media on allegations of discrimination against minorities and silencing of opposition in India by “Rashtriya Swayamsevak Sangh” (RSS) and the Bharatiya Janata Party (BJP). The content was posted during India’s mass farmer protests and briefly touched on the reasons behind the protests and praised them. Meta determined that the post violated the DIO policy and removed the content. The Board found the content did not violate the DIO policy and that inadequate time or attention was given to reviewing the content. The Board highlighted the importance of uninhibited public debate, especially concerning political figures and discussion on human rights (General Comment 34, paras 11, 34). The Board also noted the particular emphasis human rights law places on nondiscrimination in the realization of rights and the importance of independent and diverse media, especially for ethnic and linguistic minorities (General Comment 34, para. 14).

In conducting its human rights analysis, the Board noted the political context in India during this time. There were mass anti-government farmer protests and increasing governmental pressure on social media platforms to remove content. The Board acknowledged that mistakes are inevitable when moderating content at scale. However, the Board held that Meta’s responsibility to prevent, mitigate and address adverse human rights impacts requires learning from these mistakes, citing UNGPs, Principles 11 and 13.

The Board emphasized the importance of transparency for public scrutiny and assessment of how the company enforces its policies and whether there are any signs that minority language speakers are treated differently. The Board also stressed the importance of processes for reviewing moderation decision making, including auditing, to check for and correct any bias, especially in relation to places experiencing periods of crisis or unrest.

ii. The Board’s recommendations and Meta’s responses

The Board recommended that Meta “improve its transparency reporting to increase public information on error rates by making this information reviewable by country and language for each Community Standard. The Board underscores that more detailed transparency reports will help the public spot areas where errors are more common, including potential specific impacts on minority groups, and alert Facebook to correct them.”

In [its response](#), Meta stated that it is assessing the feasibility of implementing this recommendation. The company noted “several challenges to sharing data about enforcement actions broken down by region and country” including, that “bad actors might create fake accounts to mask the country in which they’re located”, use a VPN to appear to be in another location, and that the challenge is greater for Groups and Pages with multiple members and administrators.

The Board also recommended that Meta translate its “Community Standards and Internal Implementation Standards into Punjabi. Facebook Should also aim to make its Community Standards accessible in all language widely spoken by its users.” Prior to this case, the company did not provide its community standards in Punjabi. Meta committed to publishing its policies in Punjabi, which it now has, and stated it is assessing translating its policies to Urdu and other Indic languages.

### **III. Public Outreach and Engagement**

For each case the Board considers, it invites public comments, mapping stakeholders to reach out to for insights on the issues a case represents. For the 25 decisions published to date, the Board received 10,022 public comments from individuals and organizations around the world. The overwhelming majority of these were submitted for the board's decision on Former President Trump's Suspension. The public comments have covered a broad range of themes, including: the need to adopt a more context-sensitive approach to removing violent and graphic content that would set a higher threshold for removal of content in regions subject to armed conflicts; the need to preserve materials for potential future investigations or to hold violators of human rights accountable; the problem of insufficient language capacity for human review of content moderation decisions in conflict zones; concerns about the alleged opaque relationship between governments and Facebook; concerns that messages from designated terrorist organizations were allowed on the platform; concerns that Facebook removes journalistic content; and the company's policies and practices on assessing off-platform context in enforcing its Community Standards, particularly if the content may incite violence.

The Board also holds regular stakeholder engagements on specific content moderation issues where it invites stakeholders to take part in a discussion under the Chatham House Rule, in order to ensure frank discussion and to protect participants.

As described in the Sudan Graphic video decision, in March 2022, the Board spoke with approximately 50 advocacy organization representatives and individuals working on reporting and documenting human rights abuses. The Board heard from stakeholders about: the vital role of social media within countries controlled by repressive regimes for documenting human rights violations and bringing international media and public attention to state-sanctioned violence; concerns that a universal standard on violent and graphic content is in practice a US- focused standard; and the usefulness of warning screens to address the real problem of trauma, though some organizations reported that warning screens may limit the reach of their content.

In April 2022, the Board spoke with representatives of advocacy organizations, academics, professionals from international and inter-governmental organizations and other human rights experts on content moderation in conflict zones. The discussion touched on a number of themes including the need for principled moderation practices that still account for different conflict situations and the ways in which social media is being used; concerns about balancing the need for impartial or unbiased moderators during an armed conflict with the need to ensure moderators understand the local context; the importance of transparency from Meta on the number of moderators by country and language; statistics on moderation broken down by country; the value of providing real examples of how the company enforces its policies in different conflict situations; and the acute danger of hate speech and incitement in an armed conflict context.

#### **IV. Conclusion**

This submission highlights the most relevant Oversight Board case decisions and recommendations for ensuring that Meta respects freedom of expression in conflict situations and in other crises.

The Board is currently considering [two cases](#) that raise key issues of content moderation in conflict situations. One is from Ethiopia, involving a post from one of the parties to the armed conflict calling on the other side to surrender or die. The second concerns the removal of a post that mentions the Taliban in news reporting removed for violating its Dangerous Individuals and Organizations Community Standard. Those case decisions are due for publication in August 2022. For more details on these and other cases and a fuller understanding of the Board's recommendations to date, see the [Board's decisions](#) and [Meta's responses](#).