



## **Sudan's Rapid Support Forces Video Captive**

**2023-039-FB-UA**

### **Summary**

The Oversight Board has overturned Meta's original decision to leave up a video that shows armed men in Sudan, from the Rapid Support Forces (RSF), detaining someone in the back of a military vehicle. The video violates both the Dangerous Organizations and Individuals and Coordinating Harm and Promoting Crime Community Standards. The Board is concerned that Meta did not remove the content – which shows a prisoner of war and includes support for a group designated by the company as dangerous – quickly enough. This indicates broader issues around both effective content enforcement during armed conflicts and how content revealing the identity (“outing”) of a prisoner of war is reviewed. The Board calls on Meta to develop a scalable solution to proactively identify content outing prisoners of war during an armed conflict.

### **About the Case**

On August 27, 2023, a Facebook user posted a video of armed men in Sudan detaining a person in the back of a military vehicle. A man speaking in Arabic identifies himself as a member of the RSF and claims the group has captured a foreign national, likely a combatant associated with the Sudanese Armed Forces (SAF). The man goes on to say they will deliver him to the RSF leadership, and that they intend to find and capture the leaders of the SAF as well as any of the SAF's foreign associates in Sudan. The video includes derogatory remarks about foreign nationals and leaders of other nations supporting the SAF, while the accompanying caption states in Arabic, “we know that there are foreigners fighting side by side with the devilish Brotherhood brigades.”

In April 2023, an armed conflict broke out in Sudan between the RSF paramilitary group and the SAF, which is the official government's military force. Approximately 7.3 million people have been displaced because of the conflict, with more than 25 million facing severe food insecurity. Sudanese human rights organizations have reported that the RSF has detained more than 5,000 people, keeping them in inhumane conditions. There are reports that both sides have committed war crimes and crimes against humanity. Meta has designated the RSF under its Dangerous Organizations and Individuals Community Standard.

Shortly after the video was posted, three Facebook users reported the content but due to a low severity (likelihood of violating community standards) and low virality (number of views) score, they were not prioritized for human review and the content left up. One of the users appealed but this report was closed because of Meta's COVID-19 automation policies. The same user then appealed to the Oversight Board. After the Board brought the



case to Meta’s attention, the company removed the Facebook post under its Dangerous Organizations and Individuals Community Standard, also applying both a standard and severe strike to the profile of the person who posted the video.

## **Key Findings**

The content violates Meta’s Dangerous Organizations and Individuals Community Standard because it contains support for a group designated by the company as a Tier 1 dangerous organization – specifically by “channeling information or resources, including official communications” on the organization’s behalf. The man seen speaking in the video identifies himself as part of the RSF, describes its activities, speaks of the actions the group is taking and directly names the RSF commander, Mohamed Hamdan Dagalo. The Board finds that removal of this content, which includes threats to anyone who opposes or challenges the RSF, is necessary and proportionate.

In previous decisions, the Board has emphasized its concern around the lack of transparency of Meta’s designated organizations and individuals list. Given the situation in Sudan, where the RSF has de facto influence or control over parts of the country, civilians who rely on Facebook, including the RSF’s communications channels, for critical security and humanitarian information, could be at greater risk through the restrictions placed on those communications channels.

Additionally, the Board finds this content violates the Coordinating Harm and Promoting Crime policy because it shows a captured man who is fully visible and described in the video as a “foreign captive” associated with the SAF. Meta’s policy does not allow for the identity of a prisoner of war to be exposed during an armed conflict. Removing the video is necessary given the specific rules of international humanitarian law to protect detainees in armed conflict. The Board is concerned that this content was not identified and removed for violating Meta’s rule against outing prisoners of war. This lack of enforcement is likely because this rule is currently enforced on escalation-only, meaning human reviewers moderating content at-scale cannot take action themselves. In fact, the rule can only be enforced if brought to the attention of Meta’s escalations-only teams by some other means, for example, Trusted Partners or content with significant press coverage.

Finally, the Board is also concerned that Meta failed to remove this content immediately or shortly after it was posted. Meta’s automated systems failed to correctly identify a violation in this video, indicating a broader issue of enforcement. The Board believes that changes need to be made to allow more content supporting dangerous organizations to be sent for human review when it relates to armed conflicts.

## **The Oversight Board’s Decision**

The Oversight Board has overturned Meta’s original decision to leave up the video.

The Board recommends that Meta:



- Develop a scalable solution to enforce the Coordinating Harm and Promoting Crime policy that prohibits outing prisoners of war within the context of armed conflict. There should be a specialized team to prioritize and proactively identify content outing prisoners of war during a conflict.
- Audit the training data used in its video content understanding classifier to evaluate whether it has sufficiently diverse examples of content supporting designated organizations in the context of armed conflicts, including different languages, dialects, regions and conflicts.
- Include a hyperlink to the U.S. Foreign Terrorist Organizations and Specially Designated Global Terrorists lists in its Community Standards, where these lists are mentioned.

\*Case summaries provide an overview of cases and do not have precedential value.

## **Full Case Decision**

### **1. Decision Summary**

The Oversight Board overturns Meta’s original decision to leave up a video that shows armed men, who describe themselves as Rapid Support Forces (RSF) members, detaining a person in the back of a military vehicle. The RSF members describe the captive, whose face can be seen clearly, as a foreign national associated with the Sudanese Armed Forces (SAF). Meta has designated the RSF under its Dangerous Organizations and Individuals Community Standard. After the Board selected the case, Meta reviewed its original decision and removed the video for violating its Dangerous Organizations and Individuals Community Standard that prohibits “support” for designated entities, specifically “channeling information or resources, including official communications, on behalf of a designated entity or event.” The post also violates Meta’s Coordinating Harm and Promoting Crime policy, which prohibits content revealing the identity of a prisoner of war in an armed conflict; in this case, the person detained in the vehicle. The Board is concerned that Meta did not remove the content quickly enough, which could indicate there are broader issues of effective policy enforcement during armed conflicts.

### **2. Case Description and Background**

On August 27, 2023, a Facebook user posted a video showing armed men in Sudan detaining a person in the back of a military vehicle. In the video, a man, who is not the user who posted the content, identifies himself in Arabic as a member of the RSF paramilitary group. He claims the group has captured a foreign national, likely a combatant associated with the SAF, and that they intend to deliver him to the RSF



leadership. The man also states they intend to find the leaders of the SAF forces and their foreign associates in Sudan, that they will capture anyone working against the RSF and that they remain loyal to their own leader, Mohamed Hamdan Dagalo. The video includes derogatory remarks about foreign nationals and leaders of other nations supporting the SAF.

The video was accompanied by a caption, also in Arabic, that translates as “we know that there are foreigners from our evil neighbor fighting side by side with the devilish Brotherhood brigades.” Shortly after the video was posted, the user edited the caption. The edited caption translates as “we know that there are foreigners fighting side by side with the devilish Brotherhood brigades.” The post had fewer than 100 reactions, 50 comments and 50 shares, while the person who posted the content has about 4,000 friends and 32,000 followers.

Shortly after it was posted, other Facebook users reported the content, but these reports were not prioritized for human review and the post was kept up on the platform. One of these users appealed Meta's decision but the appeal was again closed without review. The same user then appealed to the Oversight Board. After the Board brought the case to Meta’s attention in October 2023, and following a review by Meta’s policy subject matter experts, the company removed the post from Facebook under its Dangerous Organizations and Individuals policy. Following removal of the content, Meta applied a severe strike in addition to a standard strike to the profile of the person who posted the content because a severe strike results in different, and stricter, penalties than a standard strike. The accumulation of standard strikes can lead to an ascending severity of penalties. When the content posted infringes Meta’s more severe policies, such as the Dangerous Organizations and Individuals policy, the company may apply additional, more severe restrictions on top of the standard restrictions. For example, [users may be restricted](#) from creating ads and using Facebook Live for set periods of time.

The Board considered the following context in reaching its decision on this case.

The armed conflict in Sudan started in April 2023 between the SAF – the military forces of the internationally recognized government, led by General Abdel Fattah al-Burhan – and the paramilitary group, the RSF, led by [Mohamed Hamdan Dagalo](#), generally known as “Hemedti.” Shortly after the beginning of the conflict, the SAF [declared](#) the RSF a rebel group and ordered its dissolution. The war in the country has been [classified](#) as a non-international armed conflict. As of November 2023, [according to Sudan War Monitor](#), the RSF was controlling most of West Darfur, the area around the capital Khartoum and parts of North and West Kordofan, while the SAF was in control of most of the Nile Valley and the country’s eastern provinces and ports.

The U.S. Treasury Department [sanctioned](#) Abdelrahim Hamdan Dagalo, an RSF figurehead and brother of Mohamed Hamdan Dagalo, on September 6, 2023. Meta independently designated the RSF as a Tier 1 terrorist organization almost one month



earlier on August 11, 2023, under its Dangerous Organizations and Individuals policy. At the time of publishing this decision, Meta's designation of the RSF remains in place.

According to the United Nations, since April 2023 approximately [7.3 million](#) people have been displaced because of the conflict, with women and children representing about half of that total. Over 25 million people, including more than 14 million children, are facing severe food insecurity and need [humanitarian assistance](#). Gender-based violence, sexual violence, harassment, sexual exploitation and trafficking are all escalating. It is estimated that disease outbreaks and the decline of the health system have resulted in around 6,000 deaths across Sudan. In October 2023, the UN Human Rights Council [adopted](#) a resolution to urgently establish an independent international fact-finding mission to Sudan, with a mandate to investigate and establish the facts and circumstances of alleged human rights and international humanitarian law violations committed during the conflict. Sudanese human rights organizations have [reported](#) that the RSF had detained more than 5,000 people in the capital Khartoum, keeping them in degrading, inhumane conditions of detention, with a lack of access to basic necessities essential for human dignity.

According to multiple sources, including the [International Criminal Court](#) and the [U.S. Department of State](#), there are reports that members of both the SAF and the RSF have committed genocide, crimes against humanity and war crimes in Sudan. Additionally, the reports mention that the RSF and allied militias have committed war crimes by ethnically targeting Masalit communities in Sudan and the Chad border. Experts specializing in Middle East and North Africa studies, consulted by the Board, highlighted reports that both sides are also responsible for widespread abuses against detainees, including [inhumane conditions](#), [illegal and arbitrary detentions](#), [ethnic targeting](#), [sexual violence](#), killing and using hostages as [human shields](#).

Experts consulted by the Board noted that Meta's designation of the RSF as a dangerous organization led to the organization's dissemination of information, including harmful narratives, being limited. However, this designation also encouraged the RSF to explore other tactics for sharing information, like resorting to the use of non-official personal pages and accounts, including to post content about detainees. This made it harder for observers to effectively monitor or counter the group's activities. Experts also noted the designation of the RSF contributed to information asymmetry and hampered access to information for civilians. For example, people would be less likely to receive RSF updates about the security conditions in certain areas through Meta's platforms (see public comment from Civic Media Observatory, PC-24020).

Sudanese civilians and media rely on social media platforms, Facebook in particular, for acquiring [crucial information and updates](#) about social, political, military and humanitarian developments and spreading these beyond Sudan; finding [routes to safety](#) within the country or to flee Sudan; finding crucial information on military operations or violent outbreaks to learn about the [military actions](#) being taken in certain locations and to seek shelter or take refuge from those actions (see also public



comment from Civic Media Observatory, PC -24020); seeking [humanitarian and medical](#) help; and learning [about hostages and prisoners of war](#).

### **3. Oversight Board Authority and Scope**

The Board has authority to review Meta’s decision following an appeal from the person who previously reported content that was left up (Charter Article 2, Section 1; Bylaws Article 3, Section 1).

The Board may uphold or overturn Meta’s decision (Charter Article 3, Section 5), and this decision is binding on the company (Charter Article 4). Meta must also assess the feasibility of applying its decision in respect of identical content with parallel context (Charter Article 4). The Board’s decisions may include non-binding recommendations that Meta must respond to (Charter Article 3, Section 4; Article 4). When Meta commits to act on recommendations, the Board monitors their implementation.

### **4. Sources of Authority and Guidance**

The following standards and precedents informed the Board’s analysis in this case:

#### *I. Oversight Board Decisions*

- [Nazi Quote](#)
- [Tigray Communication Affairs Bureau](#)
- [Armenian Prisoners of War Video](#)
- [Punjabi Concern over the RSS in India](#)
- [Öcalan’s Isolation](#)
- [Holocaust Denial](#)
- [Weapons Post Linked to Sudan’s Conflict](#)
- [Referring to Designated Individuals as “Shaheed” policy advisory opinion](#)

#### *II. Meta’s Content Policies*

The Board’s analysis was informed by Meta’s [commitment to voice](#), which the company describes as “paramount,” and its values of safety, privacy and dignity.

After the Board identified this case for review, Meta removed the content for violating the Dangerous Organizations and Individuals policy for support of a designated entity. The content also violated the Coordinating Harm and Promoting Crime policy on depicting identifiable prisoners of war in an armed conflict. As Meta has informed the Board in previous cases, when the content violates several policies, the company enforces under the most severe violation. In this case, Meta considered the Dangerous Organizations and Individuals policy violation to be the most severe.

#### [Dangerous Organizations and Individuals](#)





According to the [Dangerous Organizations and Individuals](#) policy rationale, in an effort to prevent and disrupt real-world harm, Meta does not allow organizations or individuals that proclaim a violent mission or are engaged in violence to have a presence on its platforms. Meta assesses these entities based on their behavior both online and offline, most significantly, their ties to violence. At the time the content in this case was posted, the policy prohibited “praise, substantive support and representation” of designated entities.

“Substantive support” covered “channeling information or resources, including official communications, on behalf of a designated entity or event,” by “directly quoting a designated entity without [a] caption that condemns, neutrally discusses or is a part of news reporting.”

On December 29, 2023, Meta updated the policy line for “substantive support.” The updated version stipulates that Meta removes “glorification, support and representation of Tier 1 entities.” Meta added two sub-categories: “material support” and “other support.” The rule for “channeling” now appears under “other support.”

#### Coordinating Harm and Promoting Crime

According to the policy rationale, the [Coordinating Harm and Promoting Crime Community Standard](#) aims to “disrupt offline harm and copycat behaviour” by prohibiting people from “facilitating, organizing, promoting or admitting to certain criminal or harmful activities targeted at people, businesses, property or animals.” This Community Standard prohibits “outing: exposing the identity of a person and putting them at risk of harm.” Among the groups protected from “outing,” the policy lists “prisoners of war, in the context of an armed conflict.”

#### *III. Meta’s Human Rights Responsibilities*

The UN Guiding Principles on Business and Human Rights (UNGPs), endorsed by the UN Human Rights Council in 2011, establish a voluntary framework for the human rights responsibilities of private businesses. In 2021, Meta [announced](#) its [Corporate Human Rights Policy](#), in which it reaffirmed its commitment to respecting human rights in accordance with the UNGPs. Significantly, the UNGPs impose a heightened responsibility on businesses operating in a conflict setting (“Business, human rights and conflict-affected regions: towards heightened action,” [A/75/212](#)).

The Board's analysis of Meta’s human rights responsibilities in this case was informed by the following international standards:

- The right to freedom of expression: Article 19, International Covenant on Civil and Political Rights ([ICCPR](#)), [General Comment No. 34](#), Human Rights Committee, 2011; UN Special Rapporteur on freedom of opinion and expression, reports: [A/HRC/38/35](#) (2018) and [A/74/486](#) (2019), UN Special Rapporteur’s report on Disinformation and freedom of opinion



and expression during armed conflicts, Report [A/77/288](#); OHCHR [report](#) on International legal protection of human rights in armed conflict (2011).

- [The right to life](#): Article 6, ICCPR.
- [The right to liberty and security of person](#): Article 9, ICCPR.
- [The right to be free from torture, inhuman and degrading treatment](#): Article 7, ICCPR.
- [The right to privacy](#): Article 17, ICCPR.
- [The right to be treated with humanity in detention](#): Article 10, ICCPR.
- [Protection of prisoners of war from degrading and humiliating treatment, including from insults and public curiosity](#): [Common Article 3 of the Geneva Conventions](#) and [Additional Protocol II](#); Article 13, para. 2, [Geneva Convention III, Commentary to Geneva Convention \(III\)](#), International Red Cross Committee (ICRC), 2020; [General Comment 31](#), Human Rights Committee, 2004.

## 5. User Submissions

The user who appealed the company's decision to keep the content up stated that the post includes misleading information and scenes and threats of violence by the RSF in Sudan's capital, Khartoum. The user asked that the content be removed because it poses a danger to people in Sudan.

## 6. Meta's Submissions

Meta told the Board that its initial decision to keep the content up was because its automated systems did not prioritize the content for human review. According to [Meta's Transparency Center](#), in general, reports are dynamically prioritized for review based on factors such as the severity of the predicted violation, the content's virality and the likelihood that the content will violate the Community Standards. Reports that are consistently ranked lower in priority than others in the queue will typically be closed after 48 hours. Shortly after the content in this case was posted, three Facebook users reported the content four times for "terrorism," "hate speech" and "violence." Due to a low severity and a low virality score, these reports were not prioritized for human review and the content was left on the platform. One of these users appealed Meta's decision to keep the content up. According to the company, that appeal was automatically closed due to COVID-19 automation policies, which Meta introduced at the beginning of the pandemic in 2020 to reduce the volume of reports being sent to human reviewers, while keeping open potentially "high-risk" reports (see the [Holocaust Denial](#) decision). When the report was auto-closed, the content was not escalated to policy or subject matter experts for additional review.

Meta explained that following the Board selecting this case, the company decided to remove the post because it violated the Dangerous Organizations and Individuals Community Standard. Meta concluded that by posting a video that shows a self-proclaimed member of the RSF speaking about the organization's activities, without a





caption that “condemns, neutrally discusses or is a part of news reporting,” the user violated the “substantive support” policy line by “channeling information” about a Tier 1 designated entity. Meta therefore removed it from the platform.

The Board asked Meta 13 questions in writing. Questions related to Meta’s enforcement measures for content related to Sudan’s conflict, automated systems and ranking models, the processes related to the designation of dangerous organizations and individuals, the rationale for designating the RSF a Tier 1 terrorist organization and the impact of this decision on access to information in Sudan. Meta answered all questions.

## **7. Public Comments**

The Oversight Board received 16 public comments that met the terms for submission. Eleven Ten of them were submitted from the Middle East and North Africa, three from Europe and one each from Latin America and the Caribbean, Sub-Saharan Africa and the United States and Canada. To read the public comments submitted with consent to publish, click [here](#).

The submissions covered the following themes: the RSF’s treatment of hostages, detainees and civilians; the RSF’s alleged abuses and instances of violence in the region; the risks of exposing identifiable hostages and detainees on social media; the RSF’s use of social media; the importance of social media for civilians in Sudan; the consequences of Meta’s designation of the RSF on the information environment in Sudan; and Meta’s prioritization of content for automated and human review in conflict situations.

## **8. Oversight Board Analysis**

The Board examined whether this content should be removed by analyzing Meta’s content policies, human rights responsibilities and values. The Board also assessed the implications of this case for Meta’s broader approach to content governance.

The Board selected this case because it offered the opportunity to explore how social media companies should respect access to information in countries such as Sudan where information can be vital during an ongoing conflict, especially for civilians, yet dangerous organizations can also use these platforms to further their violent mission and promote real-world harm.

Additionally, the case provides the Board with the opportunity to assess how Meta protects detainees in armed conflicts in line with international humanitarian law. The case primarily falls into the Board’s Crisis and Conflict Situations strategic priority, but also touches on Automated Enforcement of Policies and Curation of Content.



## **8.1 Compliance With Meta's Content Policies**

### *I. Content Rules*

#### Dangerous Organizations and Individuals Community Standard

The Board finds that the content in this case violates the Dangerous Organizations and Individuals policy because it supports a designated Tier 1 organization.

Meta informed the Board that it removed the content in this case because it contained “substantive support” for a Tier 1 terrorist organization. The company explained that substantive support includes “channeling information or resources, including official communications, on behalf of a designated entity” by “directly quoting a designated entity without [a] caption that condemns, neutrally discusses or is a part of news reporting.” In this case, the video shows a person who identifies himself as a member of the RSF, speaks of the RSF's activities and the actions that will be taken, and names the RSF commander.

Additionally, Meta's internal guidelines provide a non-exhaustive list of examples of written or visual elements that show substantive support. This includes posts where “the content features, or claims to feature, a leader, spokesperson, or a known or self-proclaimed member of a designated entity speaking about the organization or its cause.”

On December 29, 2023, Meta updated the Dangerous Organizations and Individuals policy line for “substantive support.” The updated version stipulates that Meta removes “glorification, support and representation of Tier 1 entities.” Although Meta has substituted “substantive support” with “support,” these changes do not impact the analysis in this case or how Meta would enforce against this content.

#### Coordinating Harm and Promoting Crime policy

The Board finds that the content also violates the Coordinating Harm and Promoting Crime Community Standard.

Meta's policy prohibits exposing the identity of a prisoner of war during an armed conflict. According to Meta, this policy is enforced on escalation only and does not include an exception for content raising awareness about prisoners of war or condemning their treatment. Meta defines a prisoner of war as “a member of the armed forces who has been captured or fallen into the hands of an opposing power during or immediately after an armed conflict.” The Board understands this rule to apply equally to international armed conflicts and non-international armed conflicts.

In this case, the Board finds that the video shows an identifiable individual described by the armed members of the RSF who have detained him as a “foreign captive”



associated with the SAF, which is the main opponent of the RSF in Sudan's ongoing conflict. Therefore, the content violates the policy and should be removed.

## **8.2 Compliance With Meta's Human Rights Responsibilities**

### *Freedom of Expression (Article 19 ICCPR)*

Article 19 of the ICCPR provides for broad protection of expression, including political expression. This right includes the “freedom to seek, receive and impart information and ideas of all kinds.” These rights are to be respected during active armed conflicts and should continue to inform Meta's human rights responsibilities, alongside the mutually reinforcing and complementary [rules of international humanitarian law](#) that apply during such conflicts (General Comment 31, Human Rights Committee, 2004, para. 11; Commentary to UNGPs, [Principle 12](#); see also UN Special Rapporteur's report on Disinformation and freedom of opinion and expression during armed conflicts, Report A/77/288, paras. 33-35 (2022); and OHCHR report on International legal protection of human rights in armed conflict (2011) at page 59).

The UN Special Rapporteur on freedom of expression has stated that “during armed conflict, people are at their most vulnerable and in the greatest need of accurate, trustworthy information to ensure their own safety and well-being. Yet, it is precisely in those situations that their freedom of opinion and expression, which includes ‘the freedom to seek, receive and impart information and ideas of all kinds,’ is most constrained by the circumstances of war and the actions of the parties to the conflict and other actors to manipulate and restrict information for political, military and strategic objectives,” (Report A/77/288, para. 1). The Board recognizes the importance of ensuring that people can freely share information about conflicts, especially when social media is the main source of information, while simultaneously ensuring content that is likely to fuel or incite further offline violence is removed.

### *I. Legality (Clarity and Accessibility of the Rules)*

The principle of legality under international human rights law requires rules that limit expression to be clear and publicly accessible (General Comment No.34, para. 25). Restrictions on expression should be formulated with sufficient precision to enable individuals to regulate their conduct accordingly (*Ibid.*). As applied to Meta, the company should provide guidance to users as to what content is permitted on the platform and what is not. Additionally, rules restricting expression “may not confer unfettered discretion on those charged with [their] execution” and must “provide sufficient guidance to those charged with their execution to enable them to ascertain what sorts of expression are properly restricted and what sorts are not.” (A/HRC/38/35, para. 46).

### Dangerous Organizations and Individuals



The UN Special Rapporteur on freedom of expression has raised concerns with social media platforms' rules prohibiting "praise" and "support," finding the terms "excessively vague," (A/HRC/38/35, para. 26). The Board has previously criticized the Dangerous Organizations and Individuals policy's lack of clarity. Meta does not publicly share the list of entities that it designates under the policy. The company explains that it chooses to designate entities that have been designated by the United States government as [Foreign Terrorist Organizations](#) (FTOs) or [Specially Designated Global Terrorists](#) (SDGTs). While Meta's full list of Tier 1 terrorist designations is created by the company and extends beyond U.S. designations, the Board understands a substantial proportion of Meta's designated Tier 1 terrorist entities are on the FTOs and SDGTs lists. While the U.S. government lists are public, Meta's Community Standards only reference the FTOs and SDGTs frameworks and do not provide a link to these U.S. government lists. The Board therefore recommends Meta hyperlink the U.S. Foreign Terrorist Organizations and Specially Designated Global Terrorists lists in its Community Standards to improve transparency and clarity for users.

However, in this case, the RSF has not been designated by the United States government, meaning the public would not know that Meta had designated one of the parties to this conflict. This lack of transparency on designations means the public may not know whether their content could be potentially violating. In the [Nazi Quote](#) decision, the Board recommended that Meta "provide a public list of the organizations and individuals designated 'dangerous' under the Dangerous Organizations and Individuals Community Standard." Meta declined to implement this recommendation after a feasibility assessment.

The Board is concerned that given the situation in Sudan, the designation of the RSF as a Tier 1 terrorist organization, together with the lack of transparency around that designation, means that people in Sudan are not aware of the fact that one party to the conflict is prohibited from having a presence on the platform, which may lead to a disproportionate impact on the access to information in Sudan, with no notice to users of this fact. The Board believes that Meta should be more transparent when making decisions in regions affected by armed conflicts, with restricted civic space, where reliable sources of information available to civilians are limited, media freedom is under threat and civil society is fragile. Given the RSF's de facto influence and control over parts of the country (see section 2), and the reliance of Sudanese civilians on Facebook to access critical security and humanitarian information, including from the RSF's communications channels, the Board finds that the unpredictable consequences on the local population of Meta's lack of transparency on designating parties to the conflict will put their physical security at additional risk.

In the [Referring to Designated Individuals as "Shaheed"](#) policy advisory opinion, the Board addresses this issue of transparency around Meta's lists of designated entities. In recommendation no. 4, the Board urges Meta to "explain the procedure by which entities and events are designated" in more detail. It should also "publish aggregated information on the total number of entities within each tier of its designation list, as



well as how many were added and removed in the past year,” (see Referring to Designated Individuals as “Shaheed,” recommendation no. 4). The Board re-emphasizes this recommendation, urging more transparency.

For a minority of the Board, although Meta did not publicly announce the RSF’s designation, abuses committed by the RSF are widely known, including alleged war crimes and crimes against humanity, and extensively reported as the conflict escalated (see section 2). With the public’s awareness of these abuses, users could reasonably expect that sharing content recorded or disseminated by the RSF could breach Meta’s Community Standards.

Considering the specific context in Sudan, the Board concludes that the rule applied in this case – prohibiting “substantive support” by “channeling information or resources, including official communications, on behalf of a designated entity or event” – is clearly explained by Meta’s Dangerous Organizations and Individuals policy, is accessible to users and therefore meets the legality test. The Board reaches the same conclusion for the updated policy rules on “support,” published by Meta on December 29, 2023, which substituted “substantive support” with “support.” The Board notes that although Meta added two sub-categories, “material support” and “other support,” with the policy line for “channeling” now appearing under “other support,” the rule itself did not materially change.

#### Coordinating Harm and Promoting Crime Community Standard

The Board finds that Meta’s rule prohibiting “outing prisoners of war” is sufficiently clear and accessible to users, satisfying the legality principle.

#### *II. Legitimate Aim*

Any limitation on expression should pursue one of the legitimate aims listed in the ICCPR, which include national security, public order and respecting the rights of others.

According to its rationale, the [Dangerous Organizations and Individuals policy](#) aims to “prevent and disrupt real-world harm” and does not allow “organizations or individuals that proclaim a violent mission or are engaged in violence to have a presence on Meta.” The Board has previously recognized that the Dangerous Organizations and Individuals policy pursues the aim of protecting the rights of others, including the right to life, security of person, and equality and non-discrimination (Article 19, para. 3, ICCPR; see also the [Punjabi Concern Over the RSS in India](#) and [Nazi Quote](#) decisions). The Board has also previously found that the purpose of the Dangerous Organizations and Individuals policy of preventing offline harm is a legitimate aim (see the [Öcalan’s Isolation](#) decision).



The Coordinating Harm and Promoting Crime policy serves the legitimate aim of protecting the rights of others (Article 19, para. 3, ICCPR), including the right to life, privacy and protection from torture or cruel, inhuman or degrading treatment. In this case, the legitimacy of the aim underlying the prohibition on depicting identifiable prisoners of war is informed by rules of international humanitarian law that call for the protection of life, privacy and dignity of prisoners of war (Common Article 3, Geneva Conventions; also see [Armenian Prisoners of War Video](#)), and the fact that the hostilities in Sudan have been qualified as an armed conflict (see section 2).

### *III. Necessity and Proportionality*

Any restrictions on freedom of expression “must be appropriate to achieve their protective function; they must be the least intrusive instrument amongst those which might achieve their protective function; [and] they must be proportionate to the interest to be protected,” (General Comment 34, para. 34). Social media companies should consider a range of possible responses to problematic content beyond deletion to ensure restrictions are narrowly tailored (A/74/486, para. 51).

As the Board has previously highlighted in the [Tigray Communication Affairs Bureau](#) decision, the UNGPs impose a heightened responsibility on businesses operating in conflict settings. In the [Armenian Prisoners of War Video](#) decision, the Board found that “in a situation of armed conflict, the Board’s freedom of expression analysis is informed by the more precise rules in international humanitarian law.”

### Dangerous Organizations and Individuals

The Board finds that removing the content in this case is necessary and proportionate. Prohibiting content that directly quotes a self-proclaimed member of a designated organization, involved in widespread violence against civilians, when there is no caption condemning, neutrally discussing or indicating the post is part of news reporting, is necessary. In this case, the post shares a video showing a member of the RSF describing the activities and plans of the group, including threatening anyone who opposes or challenges them. Spreading this kind of information on behalf of a designated organization on Facebook, especially in the context of the armed conflict in Sudan, with the RSF implicated in widespread violence, war crimes and crimes against humanity (see section 2), could lead to a heightened risk of real-world harm. Under such circumstances, no measure short of content removal will address the risk of harm, and removal is the least restrictive means of protecting the rights of others.

The Board is concerned that Meta failed to remove this content immediately or shortly after it was posted, acting only when the Board selected this case, two months later. Meta informed the Board that despite being reported by multiple users, its automated systems gave this content a low score, which means it was not prioritized for human review. Meta’s automated systems use a variety of features when determining what action to take on a piece of content, including machine-learning classifiers that score content on the probability of a violation, severity of the potential violation and virality





of the content. If added to a queue for review, these features may also be used to prioritize or rank the order in which content may be reviewed. According to Meta, the content in this case was not prioritized due to the company's systems not detecting a violation in the video and predicting the content to have a low number of views, and so was automatically closed without review, once the 48-hour period had passed. Specifically in ranking this video, there were two cross-problem classifiers that generated predicted severity ranking scores and both classifiers generated low scores. The Board is concerned that Meta's automated detection errors in this case may indicate broader issues, in particular the classifiers failures to identify content supporting the RSF, a designated entity not allowed to have a presence on the platform, and depicting a member identifying himself as belonging to the designated entity – without a caption that condemns, neutrally discusses or is a part of news reporting. In response to the Board's question about what caused the failure of classifiers to detect the violation in this case, Meta noted it could not identify what exactly factored in this content receiving a low score.

The Board therefore concludes that Meta should take the necessary steps to enhance its automated detection and prioritization of content by auditing the training data used in its video content understanding classifier to evaluate whether it has sufficiently diverse examples of content supporting designated organizations in the context of armed conflicts, including different languages, dialects, regions and conflicts. Meta should ensure this change allows for more content to be lined up for human review. This will likely require increasing human review capacity to ensure Meta is able to effectively address an increase in volume of content necessitating review following the outbreak of a conflict. This adjustment will help the company calibrate how its automated systems respond to challenges related to armed conflicts, and better identify and address content involving dangerous organizations in these contexts, enhancing the effectiveness of its enforcement measures.

Additionally, the Board finds that Meta failed to establish a sustainable mechanism to adequately enforce its content policies during the war in Sudan. In the [Weapons Post Linked to Sudan's Conflict](#) case, Meta explained that it did not set up an Integrity Product Operations Center for Sudan, which is used to respond to threats in real-time, because the company was able to “handle the identified content risks through the current processes.” Meta reiterated a similar position in this case. Previously, the Board recommended that in order to “improve enforcement of its content policies during periods of armed conflict, Meta should assess the feasibility of establishing a sustained internal mechanism that provides the expertise, capacity and coordination required to review and respond to content effectively for the duration of a conflict (see [Tigray Communication Affairs Bureau](#) decision, recommendation no. 2).” In August 2023, Meta informed the Board that it set up “a team to address crisis coordination and provide dedicated operations oversight throughout the lifecycle of imminent and emerging crises. We have since fulfilled staffing requirements and are now in the process of ramping up this team for their operational execution responsibilities before, during, and after high risk events and elections. All operational logistics for the team have been established, and the team will be fully live across all regions in the coming



months. We will continue to improve its execution framework as we encounter conflict incidents and assess the effectiveness of this structure. We now consider this recommendation complete and will have no further updates.” However, in response to the Board’s question, Meta noted that it has not established such a mechanism for the conflict in Sudan, although the company considers the recommendation complete.

### Coordinating Harm and Promoting Crime Community Standard

The necessity and proportionality of removing this content under the policy on outing prisoners of war is informed by the more specific rules of international humanitarian law (see [Armenian Prisoners of War Video](#)). Common Article 3 to the Geneva Conventions prohibits “outrages upon personal dignity, in particular humiliating and degrading treatment” of detainees in international and non-international armed conflicts. Article 13 of the Geneva Convention (III) prohibits acts of violence or intimidation against prisoners of war as well as exposing them to insults and public curiosity. Only in some limited circumstances, international humanitarian law allows for the public disclosure of images of prisoners of war. As the ICRC notes in its guidance to media, if a “compelling public interest” or “the vital interest” of the prisoner requires it, images depicting prisoners of war may exceptionally be released so long as the dignity of the depicted prisoner is protected. When a prisoner is depicted in humiliating or degrading situations, the identity must be obscured “through appropriate methods, such as blurring, pixelating or otherwise obscuring faces and name tags,” (ICRC Commentary on Article 13 at p.1627). While the Board acknowledges that there are available online tools for users to anonymize sensitive prisoners of war content, Meta does not currently provide users with such means to blur or obscure the faces of prisoners of war in video content published on its platform.

These international humanitarian law prohibitions, and narrowly drawn exceptions, intend to protect detainees in conflict. As the Board previously held in the [Armenian Prisoners of War Video](#) case, prohibiting the sharing of images of prisoners of war “is consistent with goals embodied in international humanitarian law,” and “where content reveals the identity or location of prisoners of war, removal will generally be proportionate considering the severity of harms that can result from such content.”

In this case, removing the post was necessary given the rules of international humanitarian law and the risks present in the conflict in Sudan. As outlined in section 2 above, since the outbreak of the conflict, the RSF has detained thousands of civilians and members of the SAF’s forces or those suspected of providing them with support. There are reports of widespread violations of international humanitarian law, where detainees have been held in inhumane and degrading conditions, mistreated and even killed. Under such circumstances, and absent a compelling human rights reason for allowing this content to remain on the platform, removal is necessary and proportionate to ensure the dignity and safety of the prisoner.



The Board is concerned, given the gravity of the potential harms and the heightened risks in an armed conflict, that this content was not identified and removed for violating Meta’s rule against outing (revealing the identity of) prisoners of war. The lack of enforcement is likely because the rule that prohibits outing prisoners of war in an armed conflict is currently enforced on escalation only, meaning at-scale content moderators cannot enforce the policy. In the [Armenian Prisoners of War Video](#), the Board held that the “rule requiring additional context to enforce, and thus requiring escalation to internal teams before it can be enforced, is necessary, because determining whether a person depicted is an identifiable prisoner of war in the context of an armed conflict requires expert consideration.” However, since that decision was published, the Board learned that Meta’s at-scale moderators are not instructed or empowered to identify content that violates the company’s escalations-only policies, like the rule at issue in this case. In other words, the rule can only be enforced if content is brought to the attention of Meta’s escalations-only teams by some other means, e.g., through Trusted Partners or significant press coverage.

In practice, this means that significant amounts of content identifying prisoners of war is likely left on the platform. This raises additional concerns about the accuracy of Meta’s automated detection enforcement as escalations-only policies most likely do not produce enough human decisions to train an automated classifier.

Therefore, while the Board finds that the rule prohibiting the outing of prisoners of war in an armed conflict is necessary, it also finds that Meta’s enforcement of the policy is not adequate to meet the company’s responsibility to respect the rights of prisoners of war. To ensure effective protection of the rights of detainees under international humanitarian law, the company should develop a scalable solution to enforce the policy. Meta should establish a specialized process or protocol to proactively identify such content during an armed conflict.

### *Access to Remedy*

Meta informed the Board that the appeal in this case was automatically closed due to Meta’s COVID-19 automation policies, which meant the content was left on the platform. In the [Holocaust Denial](#) case, the Board recommended Meta to “publicly confirm whether it has fully ended all COVID-19 automation policies put in place during the COVID-19 pandemic.” The Board is concerned that Meta’s COVID-19 automation policies, justified by the temporary reduction in human review capacity due to the pandemic, are still in place and reiterates its recommendation, urging Meta to publicly explain when it will no longer have reduced human reviewer capacity.

## **9. Oversight Board Decision**

The Oversight Board overturns Meta’s original decision to leave up the content.



## 10. Recommendations

### Enforcement

1. To ensure effective protection of detainees under international humanitarian law, Meta should develop a scalable solution to enforce the Coordinating Harm and Promoting Crimes policy that prohibits outing prisoners of war within the context of armed conflict. Meta should set up a protocol for the duration of a conflict that establishes a specialized team to prioritize and proactively identify content outing prisoners of war.

The Board will consider this implemented when Meta shares with the Board data on the effectiveness of this protocol in identifying content outing prisoners of war in armed conflict settings and provides updates on the effectiveness of this protocol every six months.

2. To enhance its automated detection and prioritization of content potentially violating the Dangerous Organizations and Individuals policy for human review, Meta should audit the training data used in its video content understanding classifier to evaluate whether it has sufficiently diverse examples of content supporting designated organizations in the context of armed conflicts, including different languages, dialects, regions and conflicts.

The Board will consider this recommendation implemented when Meta provides the Board with detailed results of its audit and the necessary improvements that the company will implement as a result.

3. To provide more clarity to users, Meta should hyperlink the U.S. Foreign Terrorist Organizations and Specially Designated Global Terrorists lists in its Community Standards where these lists are mentioned.

The Board will consider this recommendation implemented when Meta makes these changes to the Community Standards.

### **Procedural Note:**

The Oversight Board's decisions are prepared by panels of five Members and approved by the majority of the Board. Board decisions do not necessarily represent the personal views of all Members.

For this case decision, independent research was commissioned on behalf of the Board. The Board was assisted by Duco Advisors, an advisory firm focusing on the intersection of geopolitics, trust and safety, and technology. Memetica, an organization that engages in



open-source research on social media trends, also provided analysis. Linguistic expertise was provided by Lionbridge Technologies, LLC, whose specialists are fluent in more than 350 languages and work from 5,000 cities across the world.