



Public Comment Appendix for Sudan Graphic Video

Case number

Case description

On 21 December 2021, Meta referred a case to the Board concerning a graphic video depicting a civilian victim of violence in Sudan. The content was posted to the user's Facebook profile Page following a military coup in the country on 25 October 2021 and the start of protests against the military takeover of the government. The protests have been met with violence, with journalists and activists attacked and arrested by the security forces.

The video shows a person, possibly a minor, with a significant head wound lying next to a car. Voices can be heard saying in Arabic that someone has been beaten and left in the street. The post includes a caption, also in Arabic, calling on the people to stand together and not to trust the military, with numerous hashtags including #DocumentingMilitaryAbuses and #CivilDisobedience. The post was viewed fewer than 1,000 times and no users reported the content.

Meta's automated systems identified the content as potentially violating and, following review, removed the content for violating the [Violent and Graphic Content Community Standard](#). The user appealed Meta's decision to remove the post. Meta reviewed the post again and applied the newsworthiness allowance to restore the post. When Meta restored the post, it placed a warning screen on the video marking it as sensitive and requiring users to click through to view the content. The warning screen prohibits users under the age of 18 from viewing the video.

Under its Violent and Graphic Content policy, Meta states that it removes any content that "glorifies violence or celebrates suffering" but allows graphic content "to help people raise awareness". The policy prohibits posting "videos of people or dead bodies in non-medical settings if they depict dismemberment". According to its newsworthiness allowance, Meta allows violating content on its platforms "if keeping it visible is in the public interest".

In its referral, Meta states that the decision on this content is difficult because it highlights the tension between the public interest value of documenting human rights violations and the risk of harm associated with sharing such graphic content. Meta also highlighted the importance of allowing users to document human rights violations during a coup and when Internet access in the country has been shut down.

The Board has not received a statement from the user responsible for the content.

The Board would appreciate public comments that address:

- *Whether Meta's policies on violent and graphic content provide sufficient protection of users documenting or raising awareness of human rights violations.*
- *Meta's compliance with its human rights responsibilities around moderation of expression containing graphic and violent content, including whether the rights of all victims are equally protected and whether it sufficiently protects the rights of traumatised survivors and relatives or loved ones of depicted victims.*
- *Meta's moderation of violent and graphic content during periods of crisis, mass protests or Internet shutdowns and what factors Meta should consider when determining whether to remove or apply warning screens and age-gating to such content.*
- *How the use of a warning screen on graphic content, including to restrict access to minors, may affect the rights of Facebook users (e.g. to raise awareness and document abuses, right to privacy, physical integrity, physical and mental health)?*
- *How Meta's content moderation, including the use of automation, affects freedom of expression and documentation of human rights violations during a conflict, and how negative impacts may be prevented or mitigated.*

In its decisions, the Board can issue policy recommendations to Meta. While recommendations are not binding, Meta must respond to them within 60 days. As such, the Board welcomes public comments proposing recommendations that are relevant to this case.



Public Comment Appendix for Sudan Graphic Video

Case number

The Oversight Board is committed to bringing diverse perspectives from third parties into the case review process. To that end, the Oversight Board has established a public comment process.

Public comments respond to case descriptions based on the information provided to the Board by users and Facebook as part of the appeals process. These case descriptions are posted before panels begin deliberation to provide time for public comment. As such, case descriptions reflect neither the Board's assessment of a case, nor the full array of policy issues that a panel might consider to be implicated by each case.

To protect the privacy and security of commenters, comments are only viewed by the Oversight Board and as detailed in the [Operational Privacy Notice](#). All commenters included in this appendix gave consent to the Oversight Board to publish their comments. For commenters who did not consent to attribute their comments publicly, names have been redacted. To withdraw your comment, please email contact@osbadmin.com.

To reflect the wide range of views on cases, the Oversight Board has included all comments received except those clearly irrelevant, abusive or disrespectful of the human and fundamental rights of any person or group of persons and therefore violating the [Terms for Public Comment](#). Inclusion of a comment in this appendix is not an endorsement by the Oversight Board of the views expressed in the comment. The Oversight Board is committed to transparency and this appendix is meant to accurately reflect the input we received.



Public Comment Appendix for Sudan Graphic Video

Case number

5

Number of Comments

Regional Breakdown

0 Asia Pacific & Oceania	0 Central & South Asia	2 Europe	0 Latin America & Caribbean
0 Middle East and North Africa	1 Sub-Saharan Africa	2 United States & Canada	

Sudan graphic video PC-10392

Case number

Public comment number

Europe

Region

Tetiana

Commenter's first name

Avdieieva

Commenter's last name

English

Commenter's preferred language

Centre for Democracy and Rule of Law No

Organization

Response on behalf of organization

Yes

Consent to Attribute

Yes

Consent to Publish

Short summary provided by the commenter

Where survivors and victims of violence themselves apply to Meta with a request to remove such content, moderators shall delete content.

Criteria for content removal: human rights record of the region, where the situation happens and possibility to cooperate with local authorities communicating the available evidence; degree of depicted violence (the less grave image – less restrictions); substance of the supplementary textual messages.

It is very important to preserve availability of images for relatives and friends, even if photos and videos might contain personal data of some individuals or be very harsh, since they might be the only source for determining person's fate.

Full Comment

Tetiana Avdieieva, Centre for Democracy and Rule of Law Public Comment for Sudan Graphic Video Case Meta's policies on violent and graphic content are formulated clearly and precisely, enabling users to predict the consequences of their conduct accordingly. A distinction, however, shall be drawn between different contexts in advance, not merely define the lawfulness of content based on the public interest notion. For example, in the regions subjected to armed conflicts, civil strife, mass atrocities or large-scale human rights violations the threshold for removal of content shall be higher than in the areas with no special geopolitical circumstances. I would advise this disclaimer to the policies themselves, stressing on a diversified approach based on the information environment, in which the content is published. As regards the equal protection of the rights of all victims, we shall understand that a protection in such circumstances can never be absolute, but rather based on the balancing exercise. For example, despite relatives and close friends of depicted victims might experience shock or anxiety observing the visual

materials, the prevailing interest in such circumstances still is to make materials available for further investigation. First and foremost, it is important to ensure the possibility of impartial investigation and bringing responsible persons to justice in order not to make outrageous crimes be bygones. Another situation might concern the traumatized survivors or individuals depicted on the visual materials in cases, where they themselves apply to Meta with a request to remove such content. In such situation, moderators might assume that all necessary evidence has already been passed to the respective authorities and thus a material does not perform its public interest function of preserving the proofs of violations. Pertaining to the criteria, which shall be considered when determining whether to remove or apply warning screens and age-gating to such content, attention shall be paid to the human rights record of the region, where the situation happens and possibility to cooperate with local authorities communicating the available evidence. If Meta understands that it deals with authoritarian regimes, which are aimed at suppressing the data on human rights violations, then the main task of social media is to serve a platform enabling the widest outreach of information regarding the violations. Also, warning screens might be applied in cases of blood present on the images, I would advise to abstain from placing such screens on imaginary with violent arrests or similar types of data, gravity of which has less direct influence on other users' mental health. Also, it is important to observe the supplementary textual messages. Where such messages are encouraging or stirring up violence, the publication itself might be blocked (with preserving of the material itself in the Meta archives for further communication to the responsible investigative authorities), but if accompanying messages are condemning human rights breaches or calling to stop massive violence (as in the Sudanese case), its availability still is the matter of public interest. The application of warning signs and blurring function significantly decreases the impact of violent content for mental health of Meta users. As regards restriction of access to minors, Meta moderators shall be attentive of the fact that children experiencing the periods of armed conflict, civil strife or mass human rights abuses shall be likewise notified of dangers taking place in societies. Moreover, location in the area of the conflict still puts minors under the threat of observing atrocities in the real life. Therefore, restricting access on social media may only deprive them of information regarding the locations, where violence takes place as well as individuals participating, which might only enhance worrying and feelings of anxiety (especially, where their relatives can be subjected to such treatment). As regards the privacy issues, I strongly appreciate the removal of facial recognition function since previously it significantly endangered activists depicted on the images with arrests, violence and persecutions. On the other hand, sometimes images depicting violence remain the only available source of information to determine one's fate. Respectively, it is very important to preserve availability of such images for relatives and friends, even if photos and videos might contain personal data of some individuals. This point relates to both identities of victims and offenders. AI and algorithmic moderation might impact right to know about mass atrocities and human right abuses on behalf of both local population and international community. Moreover, overly rapid removal of violent content might destroy the evidence, which sometimes remains the only available data

related to the committed crimes, especially where such violations are conducted by authoritarian governments. Probably, the best solution in cases where there are doubts in public interest of the shared materials is to disable a sharing function, preserving a material for future investigations and prosecutions of individuals responsible for mass violations of human rights.

Link to Attachment

[PC-10392](#)

Sudan graphic video PC-10407

Case number

Public comment number

Europe

Region

Jeffrey

Commenter's first name

Howard

Commenter's last name

English

Commenter's preferred language

University College London

Organization

No

Response on behalf of organization

Yes

Consent to Attribute

Yes

Consent to Publish

Short summary provided by the commenter

Meta's policies on violent and graphic content are broadly defensible. However, they would benefit from simplification and greater precision. Specifically, the newsworthiness exemption should be eliminated, and the distinction between violence used in aggression and in self-defense should be clarified.

Full Comment

What Meta's policies get right (and why): 1. There are legitimate and illegitimate reasons for posting violent and graphic content (VGC), which is why neither a categorical permission nor prohibition would be justified. 2. VGC posted for illegitimate purposes is rightly disallowed. These purposes involve celebrating, glorifying, threatening, inciting or otherwise promoting violent aggression. 3. VGC posted for legitimate purposes is rightly allowed (subject to qualification). Legitimate purposes involve documenting, condemning, and raising awareness about violent aggression in all its forms. Speakers and audiences have substantial interests in communicating and accessing such content. 4. Meta's first qualification on legitimate posting of VGC is best understood as a duty to respect privacy. This duty explains why Meta's policy of removing VGC at the request of family members is, in my view, justified. 5. Meta's second qualification is best understood as a duty to refrain from gratuitously disturbing one's audience. The best argument for this position is that some violent and graphic content must be permitted, so that speakers can convey adequately the seriousness of what is happening—but beyond a certain point the content becomes gratuitous, needlessly disturbing audiences. Where that point lies, however, is indeterminate; in my view, there is bound to be reasonable differences of opinion both across and within cultures on this point.

Accordingly, I think there is a range of reasonable policies that Meta could legitimately adopt to reflect this point (e.g., current bans on depictions of throat-slitting or visible internal organs are plausible but not obviously required). Provided such standards are clear and consistently enforced, they are defensible. 6. Since some users of Meta are minors, or otherwise prefer not to engage with VGC —e.g., because they are themselves traumatized—the current use of age-gating and interstitial warning screens can be easily justified. What Meta’s policies get wrong (and how to fix them): Re: Newsworthy content 7. The newsworthiness exemption for violent and graphic content should be eliminated, simplifying the policy. There are two reasons for drawing this conclusion. First, the policy is unnecessary. If content fulfils legitimate purposes (which include raising awareness on potentially newsworthy violence), it should be allowed, provided it respects privacy and non-gratuitousness constraints. If content pursues illegitimate purposes, it should be disallowed. These two principles can do all the work in explaining what should stay up and what should not. 8. Second, the newsworthiness exemption is bound to be applied in an ad hoc and contestable manner. What counts as newsworthy? Any speaker sharing a post with VGC will believe it is within the public interest; indeed they may be seeking to prompt a public discussion on precisely such an issue. Meta should be facilitating such a discussion, rather than taking a stand on it. Re: VGC involving defensive violence 9. According to current rules, it is a violation to use sadistic language alongside VGC “unless there is a self-defense context”. The self-defense exception needs to be specified, or else removed. The most plausible rationale for the self-defense exception is that people should be allowed to promote justified defensive violence—hence why I referred above to violent aggression rather than violence as such. (Note that this issue is not obviously relevant to the Sudanese case, but merits scrutiny as part of the wider review of the VGC Community Standard; indeed it is also relevant to the Violence & Incitement Community Standard.) 10. Meta needs to explain how its self-defense test works. Many aggressors will claim that their aggressive violence is, in fact, a response to some existing threat against which they are defending themselves. Surely speakers cannot immunize their posts from moderation simply by alleging (falsely) that the violence they are promoting is defensive in character; so what is the test? 11. One solution is for Meta to make its own determinations as to who the unjust aggressors (and justified defenders) are in any conflict situation where it engages in content moderation. In some cases, this will be straightforward (e.g., the Ukrainian invasion). But in other fraught conflict zones, the matter will be less straightforward. The difficulties are compounded by the question of whether Meta is referring to morally justified defensive action or legally justified defensive action. 12. It is clearly undesirable for Meta to make these ethical and/or legal judgements across manifold conflict zones, especially given how difficult they would be to automate at scale. An alternative solution, which avoids these judgments, is to eliminate the self-defense exception and simply ban all remarks promoting violence, whether defensive or aggressive. Such a policy would be overinclusive, restricting legitimate posts by, e.g., Ukrainians promoting their self-defense; but it would have the virtue of enabling Meta to prescind from making contestable ethical or legal classifications of different forms of violence. Either way, Meta needs to

clarify what precisely its self-defense exemption amounts to. Implications 13. Regarding the Sudanese case, the case description does not clarify what aspect of the video violated the VGC Community Standard, such that the newsworthiness exemption was required to rescue it. It merely refers to a “significant head wound”, but provided it was not a severed head, this would not qualify as “dismemberment”. If that is correct, then the right result would be to allow the Sudanese video, which raised awareness of violent aggression against civilians in a manner that was non-gratuitous, and so counted as a legitimate posting of VGC under current policy. 14. If, however, the video in fact violated the existing policy against gratuitousness, it was rightly restricted.

Link to Attachment

[PC-10407](#)