



Oversight Board
Q2 2022
transparency report



Glossary of terms

Annual report – A report published by the Oversight Board that provides a summary of the cases it selects and reviews, as well as an overview of its operations.

Bylaws – These specify the Oversight Board’s operational procedures.

Case Management Tool (CMT) – The platform created by Meta and used by the Oversight Board to receive and review case submissions, and collect and store case files.

Case Selection Committee – A sub-committee of the Board, comprised of at least five Oversight Board Members, formed to address case selection.

Case Selection Team – A team within the Oversight Board Administration that assists the Case Selection Committee with identifying cases for panel review.

Longlist – An initial list of cases drawn up by the Case Selection Team. This is based on selection criteria set out by the Case Selection Committee.

Meta’s content policies – Facebook and Instagram’s content policies and procedures that govern content on the platforms (e.g., Community Standards or Community Guidelines).

Meta’s legal review – Step in case selection process where Meta may exclude cases from the shortlist which are ineligible for review by the Board in accordance with the Bylaws.

Meta-referred case – A case submitted to the Oversight Board by Meta. Meta has the ability to refer cases to the Board both on an ongoing basis and under emergency circumstances, with the latter being heard under the process for expedited review.

Oversight Board Administration – The full-time professional staff that support Board Members and the day-to-day operations of the Board.

Panel – Five Members of the Oversight Board assigned to review a case.

Policy advisory opinion – A text issued by the Oversight Board outlining how Meta should change its policies in a given area.

Policy advisory statement – A statement appended to an Oversight Board decision on a specific case that reflects policy considerations beyond the binding content decision.

Shortlist – A small number of cases chosen from the longlist by the Case Selection Committee to be considered for selection.

User appeal – An appeal submitted by a Facebook or Instagram user to the Oversight Board for review.



Transparency Report for second quarter of 2022

This transparency report for the second quarter of 2022 (April 1–June 30, 2022) sets out key statistics on cases selected by the Board, as well as the decisions and recommendations we made in this quarter.

In this quarter, the Board published three case decisions: “Sudan graphic video”; “Reclaiming Arabic words”; and “Knin cartoon.”



Key Findings – Second quarter 2022¹



~347,304 cases submitted to the Oversight Board, of which **eight** were submitted by Meta.

19 user-submitted cases longlisted.

10 user-submitted cases shortlisted.

Announced that **three cases** had been **selected**, two submitted by users and one referred by Meta.

Three case decisions were published in this period. These contained **10 recommendations** for Meta.

¹ Cases may pass through stages of the review process in multiple reporting periods. The data presented reflect the number of cases that are within each stage of the review process during the reporting period in question. Thus, a case submitted in Q2 but longlisted in Q3 would appear in “submitted cases” in Q2 and “longlisted cases” in Q3.



Q2 2022 Submitted User Cases

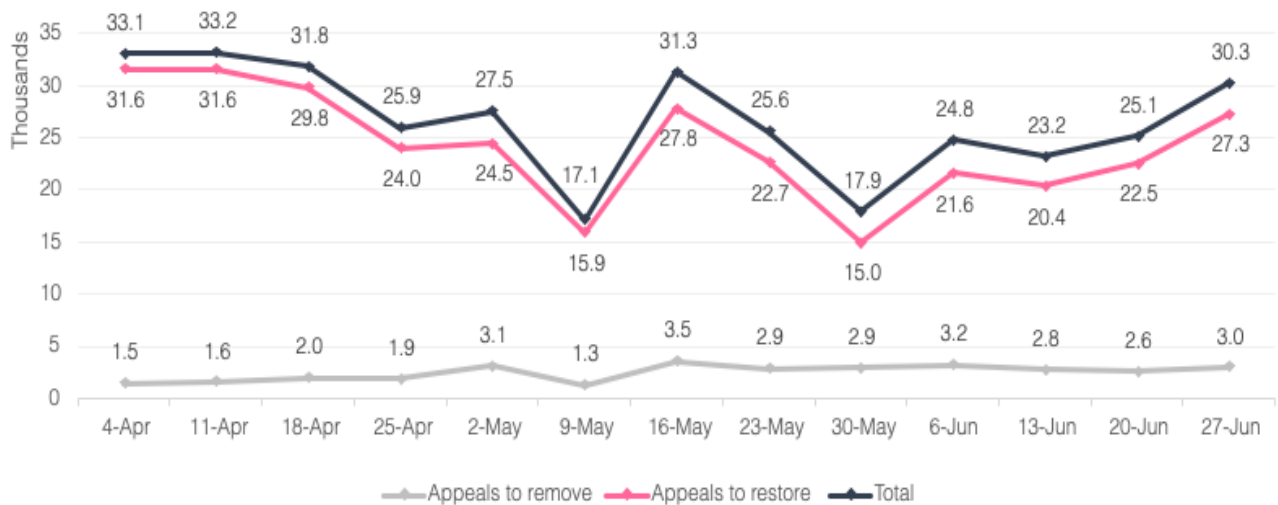
Where users have exhausted Meta’s appeals process, they can challenge the company’s decision by appealing eligible content to the Oversight Board.

In Q2 2022, **347,296 cases** were submitted by users. This is a reduction of 28% on Q1 2022, when the Board received more cases than any previous quarter. However, the number of user-submitted cases in Q2 2022 remains higher than in Q3 and Q4 2021.

In total, users submitted almost 2 million cases (1.98 million) to the Board from October 2020 to June 2022.

Estimated number of cases submitted to Oversight Board by week

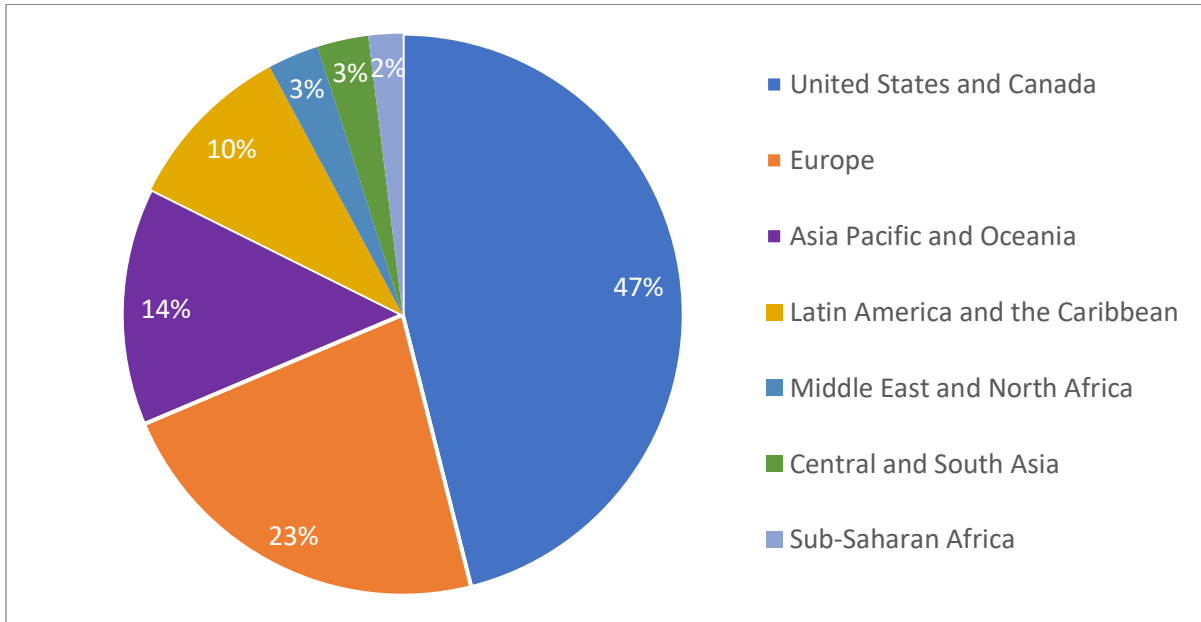
Number of cases, thousands



The Board received the highest number of appeals in Q2 in the week commencing 11 April, with users submitting 33,192 cases over seven days. As with previous quarters, the vast majority (91%) of cases submitted to the Board were appeals to restore content, while 9% of cases were appeals to remove other users’ content.

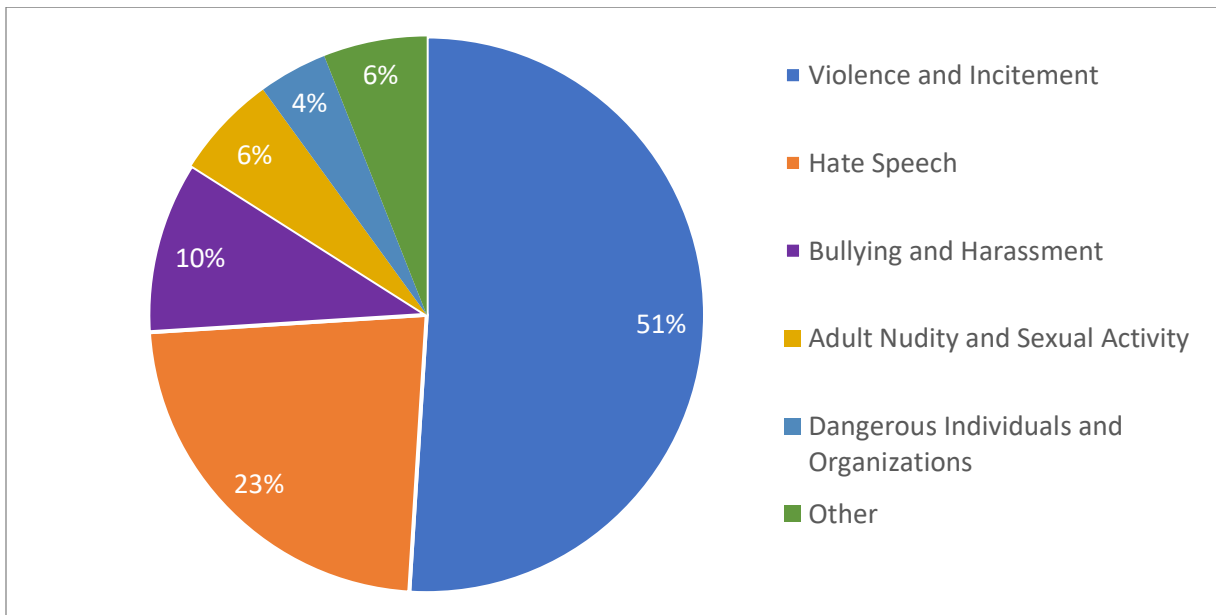


Estimated cases submitted by user-selected region Percent



Nearly half (47%) of submitted cases came from the United States & Canada, followed by 23% from Europe, 14% from Asia Pacific and Oceania, and 10% from Latin America and the Caribbean.

Estimated appeals to *restore* content to Facebook/Instagram by Community Standard Percent



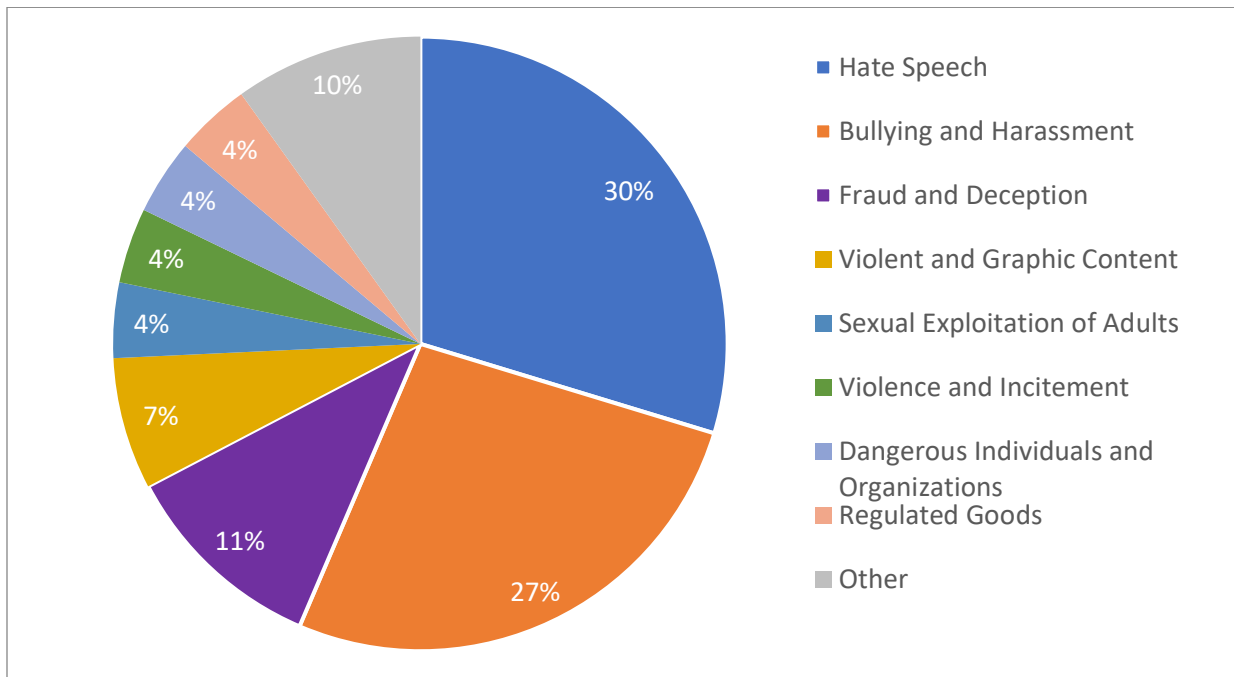


Users primarily submitted appeals to restore content which Meta removed for violating its policies on Violence and Incitement (51%), Hate Speech (23%) and Bullying and Harassment (10%).

The share of cases relating to content removed for supposedly violating Meta’s Violence and Incitement policy reached an all-time high, increasing from 44% in Q1 2022 to 51% in Q2 2022– an increase of seven percentage points. In Q2 2021, by contrast, only 18% of user appeals related to content Meta removed under its Violence and Incitement policy.

In Q2 2022, the share of cases relating to content removed for supposedly violating Meta’s Bullying and Harassment policy reached an all-time low of 10%. In Q2 2021, by contrast, cases related to Bullying and Harassment made up 29% of user appeals.

Estimated appeals to *remove* content from Facebook/Instagram by user-selected Community Standard
Percent



For user appeals to *remove* content from Facebook or Instagram, the Community Standard which has supposedly been violated is determined not by Meta (as with appeals to *restore* content) but by users themselves. When users appeal to the Board to remove someone else’s content from Facebook or Instagram, they can select one or more Community Standards.

As shown in the chart above, in Q2 2022 users submitted the most appeals to remove posts they thought violated Meta’s Hate Speech Community Standard (30%), followed by Bullying and Harassment (27%), and Fraud and Deception (11%).



Cases submitted by platform

Percent

Facebook	97.5%
Instagram	2.5%

The vast majority of cases submitted by users (97.5%) concerned content shared on Facebook, with only 2.5% of cases concerning content shared on Instagram.

Q2 2022 Longlisted User Cases

In this period, **19 user-submitted cases were longlisted** for the Case Selection Committee's consideration. These were chosen using the selection criteria set by the committee, including the 'overarching criteria for case selection' available [here](#). These cases covered 14 pieces of content on Facebook and five pieces of content on Instagram. 12 were cases to restore content and seven were cases to remove content.

Longlisted cases by region and affected country²

Number of cases

Asia Pacific and Oceania	3
Philippines	1
China	2
Europe	18
Ukraine	6
Russia	5
Romania	1
Moldova	1
Poland	1
Greece	1
Germany	1
Czech Republic	1
Sweden	1
United States & Canada	3
United States	3
Sub-Saharan Africa	1
Ethiopia	1
Central & South Asia	4
Sri Lanka	1
Myanmar	1
Pakistan	2
Total	29

² 'Countries Affected' is a user-selected field in the Board's appeal process and users can select multiple countries. For this reason, 19 user-submitted cases were longlisted but a tally of 29 is listed here. While the user selects the relevant country in the first instance, the Case Selection Team also have the ability to change the country to improve accuracy.



Of the 19 cases longlisted by the Board in this quarter, seven concerned content which had been left up on Facebook or Instagram and had not been deemed to have violated Meta’s rules. As such, they have not been assigned a Community Standard below. Of those cases where the content had been deemed by Meta to violate its rules, the most common Community Standards were Hate Speech (four cases) and Sexual Solicitation (four cases), followed by Violence and Incitement (three cases) and Dangerous Individuals and Organizations (one case).

Longlisted cases by Community Standard

Number of cases

Undefined ³	7
Sexual Solicitation	4
Hate Speech	4
Violence and Incitement	3
Dangerous Individuals and Organizations	1
Total	19

Q2 2022 Shortlisted User Cases

The Case Selection Committee identifies a shortlist of cases from the longlist to consider for selection, selecting cases by majority vote.

In this quarter, the Case Selection Committee reviewed and shortlisted cases on two occasions, **shortlisting 10 cases in total.**

The shortlist is sent to Meta’s legal team to review for eligibility as, per the Bylaws, certain cases are not eligible for review by the Board due to legal restrictions. While this eligibility review does not include a re-review on the merits, in practice, Meta has also assessed whether its original decision on a piece of content was correct or not. In many cases, it then decides to remove content which it incorrectly left up on its platforms or restore content which was incorrectly removed. This does not affect the eligibility of the content under review.

Of the 10 cases shortlisted in this period, all were confirmed as eligible by Meta’s legal team. Meta determined that its original decision on the piece of content was incorrect in six out of the 10 cases shortlisted by the Board. In all six cases, Meta found it had incorrectly removed content and restored it (a summary of each of these posts can be found in ‘Annex I’ at the end of this report).

While this is only a small sample, and the Board intentionally seeks out difficult and challenging cases, it is noted that Meta found its original decision to have been incorrect in 60% of cases the Board shortlisted in Q2 2022, and 70% in Q1 2022. The Board continues to raise with Meta the questions this poses for the accuracy of the company's content moderation and the appeals process the company applies before cases reach the Board.

³ For content that is still live on Facebook and Instagram and reported by users, the applicable Community Standard violated is undefined as no Community Standard is purportedly violated.



Cases where Meta identified that its original decision on content was incorrect

Number of cases

<i>Community Standard</i>	<i>Facebook</i>	<i>Instagram</i>	<i>Total</i>
Sexual Solicitation	0	4	4
Violence and Incitement	1	0	1
Hate Speech	1	0	1
Total	2	4	6



Q2 2022 Cases Submitted by Meta

In addition to appeals from users, Meta can also refer significant and difficult cases to the Board for consideration.

Meta submitted eight cases to the Oversight Board during this period. Meta removed the content in four cases, which were related to the Violence and Incitement Community Standard (two cases) and the Child Sexual Exploitation, Abuse and Nudity Community Standard (two cases). The remaining four cases concerned content that the company had left up on Facebook or Instagram.

Meta referrals submitted⁴

<i>Case ID</i>	<i>Name</i>	<i>Meta's decision</i>	<i>Platform</i>	<i>Language</i>	<i>Community Standard</i>	<i>Countries⁵</i>
2022-007-IG-MR	UK drill music	Take down	IG	English	Violence and Incitement	UK
N/A	N/A	Take down	FB	Malay	Violence and Incitement	Malaysia
N/A	N/A	Take down	FB	Polish	Child Sexual Exploitation, Abuse and Nudity	Israel
N/A	N/A	Keep up	IG	English	N/A	US
N/A	N/A	Keep up	FB	N/A (image)	N/A	Bulgaria
N/A	N/A	Take down	FB	Portuguese	Child Sexual Exploitation, Abuse and Nudity	Brazil
N/A	N/A	Keep up	FB	Hebrew	N/A	Israel
N/A	N/A	Keep up	IG	English	N/A	India

⁴ Cases which are not selected for assignment do not have a Case ID.

⁵ Countries listed do not necessarily align with countries assigned in longlisted cases above as a more thorough review is done at this stage of the appeals process to identify the principal countries concerned.



Policy advisory opinion on Russia’s ongoing war in Ukraine

In Q2 2022, Meta informed the Oversight Board that the company would be withdrawing an earlier request for policy guidance concerning content moderation issues related to Russia’s ongoing war in Ukraine. In taking this action, the company cited specific ongoing safety and security concerns.

While the Board understands these concerns, we believe that the request raised important issues and were disappointed by the company’s decision to withdraw it. The Board also noted that the withdrawal of this request does not diminish Meta’s responsibility to carefully consider the ongoing content moderation issues that have arisen from this war, which we continue to follow. Indeed, the importance for the company to defend freedom of expression has only increased.

Since Meta withdrew this policy advisory opinion, the Board [selected the “Russian poem” case \(2022-008-FB-UA\)](#), which is related to the Russian invasion of Ukraine and alleged crimes committed by the Russian army in the Ukrainian city of Bucha. The Board will publish its decision on this case later this year.

Q2 2022 Selected Cases

The Case Selection Committee selects cases for review, which are then announced publicly on the Board’s website.

Cases selected

<i>Case ID</i>	<i>Name</i>	<i>Date announced</i>	<i>Platform</i>	<i>Source</i>	<i>Community Standard</i>	<i>Countries</i>
2022-004-FB-UA	Colombia police cartoon	5/10/22	FB	User (appeal to restore)	Dangerous Individuals and Organizations; Violence and Incitement	Colombia
2022-005-FB-UA	Mention of the Taliban in news reporting	5/10/22	FB	User (appeal to restore)	Dangerous Individuals and Organizations	India & Afghanistan
2022-006-FB-MR ⁶	Tigray Communication Affairs Bureau	5/10/22	FB	Meta referral	Violence and Incitement	Ethiopia

⁶ This case was referred to the Board by Meta in Q1 2022



Q2 2022 Published Case Decisions

After being selected, the Board assigns cases to a five-member panel. Members of the panel include at least one member from the region implicated in the content and a mix of gender representation. The panel looks at whether Meta’s decision is consistent with its content policies and values, and its international human rights responsibilities. The Board’s decisions are binding, and Meta must implement them within seven days of publication.

In this period, the Oversight Board **decided three cases**. It overturned Meta’s decision in two of them, and upheld the company’s decision in the third.

Cases decided

<i>Case ID</i>	<i>Name</i>	<i>Platform</i>	<i>Source</i>	<i>Language of content</i>	<i>Community standard</i>	<i>Countries⁷</i>	<i>Outcome</i>
2022-001-FB-UA	Knin cartoon	FB	User appeal (to remove)	Croatian	Hate Speech	Croatia	Overturned
2022-002-FB-MR	Sudan graphic video	FB	Meta referral	Arabic	Violent and graphic content	Sudan	Upheld
2022-003-IG-UA	Reclaiming Arabic words	IG	User appeal (to restore)	Arabic and English	Hate Speech	Morocco, Egypt, Lebanon	Overturned

⁷ Countries listed do not necessarily align with countries identified in the metadata of longlisted cases above as a more thorough review is done at this stage of the appeals process.



Human Rights standards referenced in decisions

In making its decisions the Oversight Board considers international human rights standards. The table below shows which human rights standards have been referenced in decisions published this quarter.

Human rights standards referenced

Source	2022-001-FB-UA	2022-002-FB-MR	2022-003-IG-UA
UN Treaties			
ICCPR:			
Non-Discrimination & Remedies (Art. 2)	✓	✓	✓
Privacy (Art. 17)		✓	
Expression (Art. 19)	✓	✓	✓
Prohibition on incitement (Art. 20)	✓		
ICERD			
Equality and non-discrimination (Arts 2 & 5)	✓		
Convention on the Rights of the Child (CRC)			
Freedom of expression (Art. 13)		✓	
Access to information (Art. 17)		✓	
UN Treaty Bodies: Guidance & Recommendations			
Human Rights Committee			
General Comment 31 on General Legal Obligation (CCPR/C/21/Rev.1/Add. 13)		✓	
General Comment 34 on the Freedoms of Opinion and Expression (CCPR/C/GC/34)	✓	✓	✓
Communication 488/1992 (Toonen v. Australia)			✓
Committee on Elimination of Racial Discrimination			
General Recommendation 35 on hate speech	✓		
Committee on the Rights of the Child			
General Comment 25 on children's rights in relation to the digital environment (CRC/C/GC/25)		✓	
Other UN Human Rights Standards			
Responsibilities of Businesses			
Guiding Principles on Business and Human Rights (UNGPs)	✓	✓	✓
Business, human rights and conflict-affected regions: towards heightened action report (A/75/212)	✓		
Incitement			



Rabat Plan of Action (A/HRC/22/17/Add.4)	✓		
UN SR Freedom of Expression Reports			
Online hate speech (A/74/486)	✓		✓
Online content regulation (A/HRC/38/35)	✓	✓	✓
Protection from discrimination and violence on the basis of sexual orientation and gender identity			
Human Rights Council Resolution 32/2 on protection against violence and discrimination based on sexual orientation and gender identity			✓
Report of the UN High Commissioner for Human Rights on discriminatory laws and practices and acts of violence against individuals based on their sexual orientation and gender identity, A/HRC/19/41			✓

Decision timelines

A previous version of the Board’s Bylaws specified that the Board must make its decision on a case a maximum of 90 days from the date the case is selected for review by the Case Selection Committee. In January 2022, the Board’s Bylaws were amended so that the 90-day timeline starts when the Board announces a new case on its website.

For the three decisions the Oversight Board published in this quarter, the average time from assignment of case to implementation of the Board’s decision has been 91 days.⁸

<i>Case ID</i>	<i>Name</i>	<i>Beginning of 90-day period</i>	<i>Board’s decision published</i>	<i>Number of days taken out of 90 days</i>
2022-001-FB-UA	Knin cartoon	3/15/22	6/17/22	94
2022-002-FB-MR	Sudan graphic video	3/15/22	6/13/22	90

⁸ The “Knin cartoon” case went over the 90-day deadline due to challenges with full panel availability for deliberations. Those scheduling issues were identified at the point of assignment to panel and were not possible to mitigate in this case.



2022-003-IG-UA	Reclaiming Arabic words	3/15/22	6/13/22	90
----------------	-------------------------	---------	---------	----

Questions for Meta

To assist with making its decisions, the Oversight Board sends questions to Meta. Of the 59 questions sent by the Oversight Board to Meta about decisions published in this quarter, Meta answered 51 questions (86%), partially answered seven questions (12%), and did not answer one question (2%).

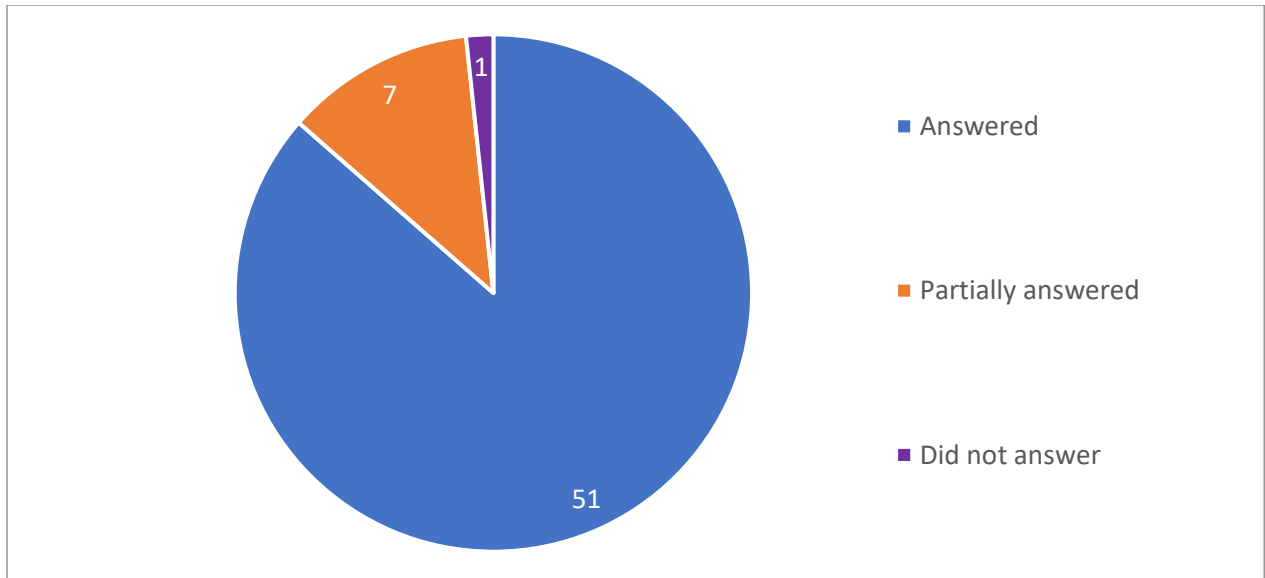
In a notable improvement, Meta gave constructive responses when not answering, or partially answering, questions from the Board. For example, the company explained why certain data-related questions would take a long time to answer and showed openness to finding ways to provide data insights in the future.

The question that Meta did not answer in this quarter was asked as part of the “Knin cartoon” case. It asked for the data Meta uses to predict the number of times a piece of content would be viewed by users, to help the Board assess its potential virality. Meta explained that generating the requested data would take at least six weeks, as it was not readily available in the form requested, and several teams would need to be involved.

Oversight Board questions answered by Meta

Number of questions

<i>Case ID</i>	<i>Name</i>	<i>Answered</i>	<i>Partially answered</i>	<i>Did not answer</i>	<i>Total</i>
2022-001-FB-UA	Knin cartoon	18	2	1	21
2022-002-FB-MR	Sudan graphic video	17	4	0	21
2022-003-IG-UA	Reclaiming Arabic words	16	1	0	17
Total		51	7	1	59



Information around wider context of Board’s decisions

In the Board’s first [quarterly transparency reports](#), published in October 2021 (page 11), in the wake of disclosures around its cross-check program, Meta agreed to provide information about the wider context which may be relevant to the Board’s case decisions. In this report we committed to provide further analysis in our transparency reporting on whether Meta is fulfilling this commitment.

This quarter, we have noted a welcome increase in the level of detail Meta has provided the Board in its initial submission on each case. This includes a more extensive explanation of its enforcement process and decisions, which improves the Board’s understanding of Meta’s internal systems.

However, in those user-appealed cases where the company found its original decision was an “enforcement error” (“Knin cartoon” and “Reclaiming Arabic words” cases), the Board observes that Meta continues to focus on defending the revised decision, rather than explaining the rationale for its original decision. The Board would encourage Meta to focus on explaining its original decision, and the trade-offs involved in improving enforcement in these types of cases.



Public comments

The Oversight Board conducts a public comment process to assist in its decision making.

For the three decisions published in Q2 2022, the Board received 21 comments, of which seven were published. The majority of public comments (67%) came from individuals, while a minority came from organizations (33%).

Public comments received by publication status

Number of comments

<i>Case ID</i>	<i>Name</i>	<i>Comments published</i>	<i>Comments not published (no consent)</i>	<i>Comments not published (violated terms)</i>	<i>Total</i>	<i>Comments unattributed⁹</i>
2022-001-FB-UA	Knin cartoon	2	0	11	13	1
2022-002-FB-MR	Sudan graphic video	2	0	3	5	0
2022-003-IG-UA	Reclaiming Arabic words	3	0	0	3	0
Total		7	0	14	21	1

Public comments received by commenter type

Number of comments

<i>Case ID</i>	<i>Name</i>	<i>Individual comments</i>	<i>Organizational comments</i>	<i>Total</i>
2022-001-FB-UA	Knin cartoon	10	3	13
2022-002-FB-MR	Sudan graphic video	3	2	5
2022-003-IG-UA	Reclaiming Arabic words	1	2	3
Total		14	7	21

⁹ Unattributed comments are published comments with the author's name redacted by request.



Public comments received by region

<i>Region</i>	Knin cartoon (2022-001-FB-UA)	Sudan graphic video (2022-002-FB-MR)	Reclaiming Arabic words (2022-003-IG-UA)	<i>Total</i>
United States & Canada	5	2	1	8
Europe	2	2	0	4
Asia Pacific & Oceania	4	0	0	4
Latin America & Caribbean	0	0	1	1
Sub-Saharan Africa	0	1	0	1
Middle East & North Africa	1	0	1	2
Central & South Asia	1	0	0	1
Total	13	5	3	21

More than a third (38%) of the public comments received for decisions published in this quarter came from the US & Canada, followed by Asia Pacific and Oceania (19%) and Europe (19%).



Recommendations

In addition to providing decisions on appealed content, the Oversight Board also made 10 policy recommendations to Meta, which the company responded to publicly within 60 days.

Of these 10 recommendations, four related to content policy (clarification of or changes to rules), three related to enforcement (clarification of or changes to how rules are applied), and three related to transparency (on disclosure of information to the public).

Oversight Board recommendations to Meta

Number of recommendations

<i>Case ID</i>	<i>Name</i>	<i>Content policy</i>	<i>Enforcement</i>	<i>Transparency</i>	<i>Total</i>
2022-001-FB-UA	Knin cartoon	1	1	0	2
2022-002-FB-MR	Sudan graphic video	3	1	0	4
2022-003-IG-UA	Reclaiming Arabic words	0	1	3	4
Total		4	3	3	10

Meta’s responses

Of the 10 recommendations the Board made in Q2 2022, Meta said it was implementing five “fully,” implementing one “in part,” and “assessing feasibility” on three. Meta said it would take “no further action” on one recommendation.

The Board’s assessment of Meta’s responses to our recommendations

Meta’s initial responses to the recommendations were almost entirely ‘comprehensive’ (nine recommendations), aside from one which was ‘somewhat comprehensive.’ This means that, for the sixth quarter in a row, the share of Meta’s responses falling into these two categories increased – rising from 91% in Q1 2022 to 100% in Q2 2022. This demonstrates that overall, Meta has increasingly been committing to concrete action under a set timeline, and has understood the core concern underlying the Board’s recommendation.

Board’s analysis of Meta’s implementation of Q2 recommendations

Of the 10 recommendations published in Q2 2022:

- 40% of recommendations published in Q2 2022 have already been wholly or partially implemented, as verified via publicly available information. This indicates an implementation turnaround time of less than six months for these recommendations, which demonstrates a marked increase in responsiveness from Meta.



- Only one recommendation was declined – Reclaiming Arabic words recommendation no. 1, which recommended that Meta translate its internal Implementation Standard into Modern Arabic. This aligns with a general trend in Meta’s responses reflecting reluctance to translate guidance to content moderators into their native languages, instead having the Implementation Standard available only in English, and providing guidance based on language-based rules which may not transfer to other languages, structured differently. This is concerning to the Board, given the diversity of languages and language structures that may not receive the same level of contextual moderation as English-language content.
- The remaining recommendations were classified as “progress reported.” This means that Meta has committed to implementing these recommendations, but has not declared implementation to be complete, and the Board also has no evidence of implementation.

As of Meta’s Q2 2022 update, the Board has issued 118 recommendations. Meta has committed to implement or implemented the majority of all recommendations.

The Board has assessed that 21 out of 118 recommendations (18%) have been implemented, as demonstrated through published information. The Board assessed a further four recommendations (3%) as partially implemented, while for 43 recommendations (36%) Meta had reported progress towards implementation.

A full list of recommendations and the Board’s assessment of Meta’s response to and implementation of each of these can be found in ‘Annex II’ at the end of this document.

Implementation Category	No. of recommendations
Implementation demonstrated through published information: <i>Meta provided sufficient data for the Board to verify the recommendation has been implemented</i>	21
Partial implementation demonstrated through published information: <i>Meta has implemented a central component of the recommendation and has provided sufficient data to verify this to the Board.</i>	4
Progress reported: <i>Meta committed to implementation but has not yet completed all necessary actions.</i>	43
Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation: <i>Meta says it implemented the recommendation but has not provided sufficient evidence to verify this.</i>	27
Recommendation declined after feasibility assessment: <i>Meta engaged with the recommendation and then decided to decline its implementation after providing information on its decision.</i>	8
Recommendation omitted, declined, or reframed: <i>Meta will take no further action on the recommendation.</i>	15
Total Number of Recommendations	118

New implementation successes include:

- **Translation of the Community Standards into Hausa, Javanese, Kannada, Kinyarwanda, Malagasy, Malayalam, and Nepali.** After translating the Community Standards into Gujarati, Marathi, Punjabi, Tamil and Telugu in [Q4 2021](#), Meta stated that there would be no further



updates to Punjabi Concern Over RSS in India recommendation no. 1, in which the Board recommended that Meta translate the Community Standards into all languages widely spoken by its users. However, the Board is pleased to see that, in addition to translations into Assamese and Farsi announced in its [Q1 2022](#) report, Meta continues to translate the Community Standards into additional languages.

- **The launch of a [Crisis Policy Protocol](#)** codifying Meta’s policy response in crisis situations globally.
- **Systemic increases in user notification granularity**
 - Meta updated notifications to users relating to government takedown requests with an [expanded explanation of its takedown request process](#).
 - Meta is making progress on rolling out user notifications specifying policy sub-type violations, and is piloting notifications that specify whether enforcement occurred as a result of human or automated review. Meta also noted in its update on the recommendations that it has updated its “enforcement communication measurement framework” to include both the availability and the transparency level of communications related to each policy as an effort to uphold a “certain standard of granularity.” This signifies that the Board is advancing a systemic impact on Meta’s approach to user notifications, rather than only impacting one subset or type.
- **Increased transparency about Meta’s practices and policies**
 - Meta has provided more information about its approach to newsworthiness exceptions ([link here](#)), including examples of newsworthy content, and the number of newsworthy allowances documented in the last year with broken out statistics on those made for content posted by politicians.
 - It also provided more information about how it creates, enforces, and audits against lists of market-specific slurs ([link here](#)) - this set of recommendations from the Reclaiming Arabic words case was implemented or partially implemented within the 60-day response period, which has only happened for two other recommendations. This may indicate that increased transparency on this topic was already on Meta’s radar, but that the Board’s recommendation was able to help prioritize execution.
- In response to a Board recommendation in the Al Jazeera case, Meta released the findings of an independent due diligence report on the **impact of the company’s policies and processes in Israel and Palestine** during the May 2021 conflict.
 - According to the report, Meta’s actions in May 2021 appear to have had an adverse human rights impact on the rights of Palestinian users to freedom of expression, and on their ability to share information and insights about their experiences as they occurred.
 - Much of Meta’s bias identified in the BSR report relates to a lack of in-language proficiency and guidance. The Board highlighted this gap in several recommendations (Punjabi concern about the RSS in India recommendation no.1, Myanmar bot recommendation no. 1, Reclaiming Arabic words, recommendation no.1). Meta consistently declined these recommendations, stating that English-language implementation standards are sufficient as moderators are fluent in English. The BSR



report, however, shows that Meta’s lack of language capabilities and cultural context among moderators lead to over-enforcement of content in Palestinian Arabic, and under-enforcement of antisemitic content. As the Board has previously pointed out, English-only guidance may cause reviewers to miss context and nuance across varying languages and dialects. Additionally, because human review data is used to train classifiers, that bias is amplified across the platform.

- Other recommendations reiterated in the BSR report include clarity of policies, transparency on government requests, and review of Dangerous Individuals and Organizations enforcement.
- **Commitment from Meta to share confidential data with the Board**
 - In Meta’s response to Wampum belt recommendation no. 3, Meta said it will share confidential data with the Board on the experiments it ran related to the accuracy of its Hate Speech policy enforcement. This is the first time Meta has stated it will share “confidential data” and is a positive step for future data sharing.



Annex I - Summaries of cases where Meta identified its original decision on a piece of content was incorrect

1. Even though Meta acknowledged its original decision in this case as incorrect, the case was selected by the Oversight Board as the “Russian poem” case (2022-008-FB-UA). Read a description of the post [here](#).
2. Even though Meta acknowledged its original decision in this case as incorrect, the case was selected by the Oversight Board as part of the “Gender identity and nudity” cases (2022-009-IG-UA). Read a description of the post [here](#).
3. Even though Meta acknowledged its original decision in this case as incorrect, the case was selected by the Oversight Board as part of the “Gender identity and nudity” cases (2022-010-IG-UA). Read a description of the post [here](#).
4. The content is a post marking the first anniversary of the Burmese military’s shooting of civilians. The post refers to Burmese soldiers as “army dogs”, describes atrocities by the army, and calls for Burmese people to fight for themselves or donate to fighters. It also states that, without consent by the people, no “army dogs” shall be allowed to rule, and the people should strive to eradicate the “evil military dictatorship” from Burmese soil. The content was erroneously taken down under the Community Standard on Violence and Incitement. The Case Selection Committee referred this appeal to Meta, who restored the content. Without explanation from Meta on either the removal or the restoration, it seems that the post was made in the context of self-defense against documented violent crackdown by an authoritarian government. The Community Standard on Violence and Incitement currently only seems to allow self-defense as an exception to statements admitting to committing high- or mid-severity violence, which may not be the case here, as there is no apparent admission of violence.
5. The content contains a photo of someone’s thighs and groin area, the latter of which is covered by undergarment. The person in the photo puts their hand over the groin area, gesturing with their index and middle fingers. The rest of the content is text where the user questions why a male caregiver felt it was inappropriate for him to take care of or even look at a female student who had to get changed after wetting her pants. The post discussed how this reflected gender inequality in that, in Hong Kong’s society and laws, men are assumed to be able to take advantage of females sexually, while women are assumed incapable of doing so. The content was erroneously taken down under the Community Standard on Sexual Solicitation. Without explanation from Meta on either the removal or the restoration, it seems that neither the photo nor the text contains any sexual solicitation.
6. The content contains a photo of a topless woman, with her nipples blurred, tied to a cross with a white shawl/pants wrapped around her waist in front of a limestone quarrel in Gotland, Sweden. The content also contains text stating “STOP MINING BOOM” and quoting from Greta Thunberg that the recent decision (in March 2022) to allow mining on the native Sami people’s land in



Sapmi, Sweden is "racist" and "colonial". The content was erroneously removed under the Community Standard on Sexual Solicitation. Without explanation from Meta on either the removal or the restoration, it seems that neither the photo nor the text contains any sexual solicitation.



Annex II – The Board’s assessment of how Meta has responded to and implemented its recommendations

<p>Armenians in Azerbaijan #1</p> <p>Q12021</p>	<p>Ensure that users are always notified of the reasons for any enforcement of the Community Standards against them, including the specific rule Facebook is enforcing. Doing so would enable Facebook to encourage expression that complies with its Community Standards, rather than adopting an adversarial posture towards users. In this case, the user was informed that the post violated the Community Standard on hate speech but was not told that the post violated the standard because it included a slur targeting national origin. Facebook satisfied the principle of legality in this instance, but Facebook's lack of transparency left its decision susceptible to the mistaken belief that it had removed the post because the user was addressing a controversial subject or expressing a viewpoint Facebook disagreed with.</p>	<p>Enforcement</p>	<p>Somewhat Comprehensive</p>	<p>Progress reported</p>
<p>Breast cancer symptoms and nudity #1</p> <p>Q12021</p>	<p>Improve the automated detection of images with text-overlay to ensure that posts raising awareness of breast cancer symptoms are not wrongly flagged for review.</p>	<p>Enforcement</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>



Breast cancer symptoms and nudity #2

Q12021

Ensure that users are always notified of the reasons for the enforcement of content policies against them, providing the specific rule within the Community Standard Facebook based its decision on.

Enforcement

Not Comprehensive

Progress reported

Breast cancer symptoms and nudity #3

Q12021

Inform users when automation is used to take enforcement action against their content, including accessible descriptions of what this means.

Enforcement

Somewhat Comprehensive

Progress reported



<p>Breast cancer symptoms and nudity #4</p> <p>Q12021</p>	<p>Ensure users can appeal decisions taken by automated systems to human review when their content is found to have violated Facebook’s Community Standard on Adult Nudity and Sexual Activity. Where Facebook is seeking to prevent child sexual exploitation or the dissemination of non-consensual intimate images, it should enforce based on its Community Standards on Sexual Exploitation of Adults and Child Sexual Exploitation, Abuse and Nudity, rather than rely on over-enforcing policies on adult nudity. Appeals should still be available in these cases, so incorrect removals of permitted consensual adult nudity can be reversed.</p>	<p>Enforcement</p>	<p>Not Comprehensive</p>	<p>Recommendation declined after feasibility assessment</p>
<p>Breast cancer symptoms and nudity #5</p> <p>Q12021</p>	<p>Implement an internal audit procedure to continuously analyze a statistically representative sample of automated content removal decisions to reverse and learn from enforcement mistakes.</p>	<p>Enforcement</p>	<p>Not Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Breast cancer symptoms and nudity #6</p> <p>Q12021</p>	<p>Expand transparency reporting to disclose data on the number of automated removal decisions per Community Standard, and the proportion of those decisions subsequently reversed following human review.</p>	<p>Transparency</p>	<p>Somewhat Comprehensive</p>	<p>Progress reported</p>



Breast cancer symptoms and nudity #7	Q12021	Revise the “short” explanation of the Instagram Community Guidelines to clarify that the ban on adult nudity is not absolute.	Content Policy	Not Comprehensive	Recommendation omitted, declined, or reframed
Breast cancer symptoms and nudity #8	Q12021	Revise the “long” explanation of the Instagram Community Guidelines to clarify that visible female nipples can be shown to raise breast cancer awareness.	Content Policy	Comprehensive	Implementation demonstrated through published information
Breast cancer symptoms and nudity #9	Q12021	Clarify that the Instagram Community Guidelines are interpreted in line with the Facebook Community Standards, and where there are inconsistencies the latter take precedence.	Content Policy	Comprehensive	Progress reported
Nazi Quote #1	Q12021	Ensure that users are always notified of the reasons for any enforcement of the Community Standards against them, including the specific rule Facebook is enforcing (e.g. for support of a hate organization).	Enforcement	Somewhat Comprehensive	Progress reported
Nazi Quote #2	Q12021	Explain and provide examples of the application of key terms used in the Dangerous Individuals and Organizations policy, including the meanings of “praise,” “support” and “representation.” These should align with the definitions used in Facebook’s Internal Implementation Standards. The Community Standard should provide clearer guidance to users on how to make their intent apparent when discussing individuals or organizations designated as dangerous.	Content Policy	Comprehensive	Implementation demonstrated through published information



Nazi Quote #3	Q12021	Provide a public list of the organizations and individuals designated “dangerous” under the Dangerous Individuals and Organizations Community Standard. At a minimum, illustrative examples should be provided. This would help users to better understand the policy and conduct themselves accordingly.	Content Policy	Comprehensive	Recommendation declined after feasibility assessment
Claimed COVID Cure #1	Q12021	The Board recommends that Facebook set out a clear and accessible Community Standard on health misinformation, consolidating and clarifying existing rules in one place (including defining key terms such as misinformation). This rule-making should be accompanied with “detailed hypotheticals that illustrate the nuances of interpretation and application of [these] rules” to provide further clarity for users (See report A/HRC/38/35, para. 46 (2018)). Facebook should conduct a human rights impact assessment with relevant stakeholders as part of its process of rule modification (UNGPs, Principles 18-19).	Content Policy	Comprehensive	Implementation demonstrated through published information



Claimed
COVID Cure #2 Q12021

To ensure enforcement measures on health misinformation represent the least intrusive means of protecting public health, the Board recommends that Facebook: Clarify the particular harms it is seeking to prevent and provide transparency about how it will assess the potential harm of particular content; Conduct an assessment of its existing range of tools to deal with health misinformation; Consider the potential for development of further tools that are less intrusive than content removals; Publish its range of enforcement options within the Community Standards, ranking these options from most to least intrusive based on how they infringe freedom of expression; Explain what factors, including evidence-based criteria, the platform will use in selecting the least intrusive option when enforcing its Community Standards to protect public health; Make clear within the Community Standards what enforcement option applies to each rule

Enforcement

Comprehensive

Implementation demonstrated through published information



Claimed
COVID Cure #3 Q12021

In cases where users post information about COVID-19 treatments that contradicts the specific advice of health authorities and where a potential for physical harm is identified but is not imminent, the Board strongly recommends Facebook to adopt a range of less intrusive measures. This could include labelling which alerts users to the disputed nature of the post's content and provides links to the views of the World Health Organization and national health authorities. In certain situations it may be necessary to introduce additional friction to a post - for example, by preventing interactions or sharing, to reduce organic and algorithmically driven amplification. Downranking content, to prevent visibility in other users' newsfeeds, might also be considered. All enforcement measures, including labelling or other methods of introducing friction, should be clearly communicated to users, and subject to appeal.

Enforcement

Not
Comprehensive

Recommendation
omitted, declined,
or reframed



<p>Claimed COVID Cure #4</p> <p>Q12021</p>	<p>Publish a transparency report on how the Community Standards have been enforced during the COVID-19 global health crisis. This should include: data in absolute and percentage terms on the number of removals, as well as data on other enforcement measures, on the specific Community Standards enforced against, including on the proportion that relied entirely on automation; a breakdown by content type enforced against (including individual posts, accounts, and groups); a breakdown by the source of detection (including automation, user flagging, trusted partners, law enforcement authorities); a breakdown by region and language; metrics on the effectiveness of less intrusive measures (e.g., impact of labelling or downranking); data on the availability of appeals throughout the crisis, including the total number of cases where appeal was withdrawn entirely, and the percentage of automated appeals; conclusions and lessons learned, including information on any changes Facebook is making to ensure greater compliance with its human rights responsibilities going forward.</p>	<p>Transparency</p>	<p>Not Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Protest in India Against France #1</p> <p>Q12021</p>	<p>To ensure users have clarity regarding permissible content, the Board recommends that Facebook provide users with additional information regarding the scope and enforcement of this Community Standard. Enforcement criteria should be public and align with Facebook's Internal Implementation Standards. Specifically, Facebook's criteria should address intent, the identity of the user and audience, and context.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>



<p>Depiction of Zwarte Piet #1</p> <p>Q22021</p>	<p>Facebook should link the rule in the Hate Speech Community Standard prohibiting blackface to the company’s reasoning for the rule, including harms it seeks to prevent.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Depiction of Zwarte Piet #2</p> <p>Q22021</p>	<p>In line with the Board’s recommendation in case 2020-003-FB-UA, Facebook should “ensure that users are always notified of the reasons for any enforcement of the Community Standards against them, including the specific rule Facebook is enforcing.” In this case any notice to users should specify the rule on blackface, and also link to above mentioned resources that explain the harm this rule seeks to prevent. Facebook should provide a detailed update on its “feasibility assessment” of the Board’s prior recommendations on this topic, including the specific nature of any technical limitations and how these can be overcome.</p>	<p>Enforcement</p>	<p>Somewhat Comprehensive</p>	<p>Progress reported</p>
<p>Punjabi concern about RSS in India #1</p> <p>Q22021</p>	<p>Facebook should translate its Community Standards and Internal Implementation Standards into Punjabi. Facebook should aim to make its Community Standards accessible in all languages widely spoken by its users. This would allow a full understanding of the rules that users must abide by when using Facebook’s products. It would also make it simpler for users to engage with Facebook over content that may violate their rights.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>



<p>Punjabi concern about RSS in India #2</p> <p>Q22021</p>	<p>In line with the Board’s recommendation in case 2020-004-IG-UA, the company should restore human review and access to a human appeals process to pre-pandemic levels as soon as possible while fully protecting the health of Facebook’s staff and contractors.</p>	<p>Enforcement</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Punjabi concern about RSS in India #3</p> <p>Q22021</p>	<p>Facebook should improve its transparency reporting to increase public information on error rates by making this information viewable 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.</p>	<p>Transparency</p>	<p>Somewhat Comprehensive</p>	<p>Progress reported</p>



<p>Former President Trump's Suspension #1</p> <p>Q22021</p>	<p>The Board believes that it is not always useful to draw a firm distinction between political leaders and other influential users. It is important to recognise that other users with large audiences can also contribute to serious risks of harm. The same rules should apply to all users of the platform; but context matters when assessing issues of causality and the probability and imminence of harm. What is important is the degree of influence that a user has over other users [...] Facebook must assess posts by influential users in context according to the way they are likely to be understood, even if their incendiary message is couched in language designed to avoid responsibility, such as superficial encouragement to act peacefully or lawfully. Facebook used the six contextual factors in the Rabat Plan of Action in this case and the Board thinks that this is a useful way to assess the contextual risks of potentially harmful speech. The Board stresses that time is of the essence in such situations; taking action before influential users can cause significant harm should take priority over newsworthiness and other values of political communication.</p>	<p>Content Policy</p>	<p>Somewhat Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Former President Trump's Suspension #2</p> <p>Q22021</p>	<p>When posts by influential users pose a high probability of imminent harm, as assessed under international human rights standards, Facebook should take action to enforce its rules quickly.</p>	<p>Enforcement</p>	<p>Somewhat Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>



<p>Former President Trump's Suspension #3</p> <p>Q22021</p>	<p>While all users should be held to the same content policies, there are unique factors that must be considered in assessing the speech of political leaders. Heads of state and other high-ranking government officials can have a greater power to cause harm than other people. Facebook should recognize that posts by heads of state and other high officials of government can carry a heightened risk of encouraging, legitimizing, or inciting violence - either because their high position of trust imbues their words with greater force and credibility or because their followers may infer they can act with impunity. At the same time, it is important to protect the rights of people to hear political speech. Nonetheless, if the head of state or high government official has repeatedly posted messages that pose a risk of harm under international human rights norms, Facebook should suspend the account for a determinate period sufficient to protect against imminent harm.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Former President Trump's Suspension #4</p> <p>Q22021</p>	<p>Periods of suspension should be long enough to deter misconduct and may, in appropriate cases, include account or page deletion.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>



Former President Trump's Suspension #5 Q22021	Restrictions on speech are often imposed by or at the behest of powerful state actors against dissenting voices and members of political oppositions. Facebook must resist pressure from governments to silence their political opposition. When assessing potential risks, Facebook should be particularly careful to consider the relevant political context.	Enforcement	Somewhat Comprehensive	Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation
Former President Trump's Suspension #6 Q22021	In evaluating political speech from highly influential users, Facebook should rapidly escalate the content moderation process to specialized staff who are familiar with the linguistic and political context and insulated from political and economic interference and undue influence. This analysis should examine the conduct of highly influential users off the Facebook and Instagram platforms to adequately assess the full relevant context of potentially harmful speech. Further, Facebook should ensure that it dedicates adequate resourcing and expertise to assess risks of harm from influential accounts globally.	Enforcement	Not Comprehensive	Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation



<p>Former President Trump's Suspension #7</p> <p>Q22021</p>	<p>Facebook should publicly explain the rules that it uses when it imposes account-level sanctions against influential users. These rules should ensure that when Facebook imposes a time-limited suspension on the account of an influential user to reduce the risk of significant harm, it will assess whether the risk has receded before the suspension term expires. If Facebook identifies that the user poses a serious risk of inciting imminent violence, discrimination, or other lawless action at that time, another time-bound suspension should be imposed when such measures are necessary to protect public safety and proportionate to the risk.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Former President Trump's Suspension #8</p> <p>Q22021</p>	<p>When Facebook implements special procedures that apply to influential users, these should be well documented. It was unclear whether Facebook applied different standards in this case, and the Board heard many concerns about the potential application of the newsworthiness allowance. It is important that Facebook address this lack of transparency and the confusion it has caused.</p>	<p>Transparency</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>
<p>Former President Trump's Suspension #9</p> <p>Q22021</p>	<p>Facebook should produce more information to help users understand and evaluate the process and criteria for applying the newsworthiness allowance. Facebook should clearly explain how the newsworthiness allowance applies to influential accounts, including political leaders and other public figures.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>
<p>Former President Trump's Suspension #10</p> <p>Q22021</p>	<p>For cross check review, Facebook should clearly explain the rationale, standards, and processes of review, including the criteria to determine which pages and accounts are selected for inclusion.</p>	<p>Content Policy</p>	<p>Somewhat Comprehensive</p>	<p>Implementation demonstrated through published information</p>



Former President Trump's Suspension #11	Q22021	<p>Facebook should report on the relative error rates and thematic consistency of determinations made through the cross-check process compared with ordinary enforcement procedures.</p>	Transparency	Not Comprehensive	Recommendation omitted, declined, or reframed
Former President Trump's Suspension #12	Q22021	<p>When Facebook's platform has been abused by influential users in a way that results in serious adverse human rights impacts, it should conduct a thorough investigation into the incident. Facebook should assess what influence it had and assess what changes it could enact to identify, prevent, mitigate, and account for adverse impacts in future.</p>	Enforcement	Not Comprehensive	Recommendation omitted, declined, or reframed
Former President Trump's Suspension #13	Q22021	<p>Facebook should undertake a comprehensive review of its potential contribution to the narrative of electoral fraud and the exacerbated tensions that culminated in the violence in the United States on January 6, 2021. This should be an open reflection on the design and policy choices that Facebook has made that may enable its platform to be abused. Facebook should carry out this due diligence, implement a plan to act upon its findings, and communicate openly about how it addresses adverse human rights impacts it was involved with.</p>	Transparency	Not Comprehensive	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>



<p>Former President Trump's Suspension #14</p> <p>Q22021</p>	<p>Facebook 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. Facebook's corporate human rights policy should make clear the protocols the company has in place in this regard. The policy should also make clear how information previously public on the platform can be made available to researchers conducting investigations that conform with international standards and applicable data protection laws.</p>	<p>Transparency</p>	<p>Somewhat Comprehensive</p>	<p>Recommendation omitted, declined, or reframed</p>
<p>Former President Trump's Suspension #15</p> <p>Q22021</p>	<p>Facebook should explain in its Community Standards and Guidelines its strikes and penalties process for restricting profiles, pages, groups and accounts on Facebook and Instagram in a clear, comprehensive, and accessible manner. These policies should provide users with sufficient information to understand when strikes are imposed (including any applicable exceptions or allowances) and how penalties are calculated.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>
<p>Former President Trump's Suspension #16</p> <p>Q22021</p>	<p>Facebook should also provide users with accessible information on how many violations, strikes, and penalties have been assessed against them, as well as the consequences that will follow future violations.</p>	<p>Enforcement</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>



Former President Trump's Suspension #17	Q22021	In its transparency reporting, Facebook should include numbers of profile, page, and account restrictions, including the reason and manner in which enforcement action was taken, with information broken down by region and country.	Transparency	Not Comprehensive	Progress reported
Former President Trump's Suspension #18	Q22021	Facebook should develop and publish a policy that governs its response to crises or novel situations where its regular processes would not prevent or avoid imminent harm. While these situations cannot always be anticipated, Facebook's guidance should set appropriate parameters for such actions, including a requirement to review its decision within a fixed time.	Content Policy	Comprehensive	Partial Implementation demonstrated through published information
"Two Buttons" Meme #1	Q22021	Make technical arrangements to ensure that notice to users refers to the Community Standard enforced by the company. If Facebook determines that (i) the content does not violate the Community Standard notified to user, and (ii) that the content violates a different Community Standard, the user should be properly notified about it and given another opportunity to appeal. They should always have access to the correct information before coming to the Board.	Enforcement	Comprehensive	Recommendation declined after feasibility assessment
"Two Buttons" Meme #2	Q22021	Include the satire exception, which is currently not communicated to users, in the public language of the Hate Speech Community Standard.	Content Policy	Comprehensive	Implementation demonstrated through published information



"Two Buttons" Meme #3	Q22021	Make sure that it has adequate procedures in place to assess satirical content and relevant context properly. This includes providing content moderators with: (i) access to Facebook's local operation teams to gather relevant cultural and background information; and (ii) sufficient time to consult with Facebook's local operation teams and to make the assessment. Facebook should ensure that its policies for content moderators incentivize further investigation or escalation where a content moderator is not sure if a meme is satirical or not.	Enforcement	Comprehensive	Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation
"Two Buttons" Meme #4	Q22021	Let users indicate in their appeal that their content falls into one of the exceptions to the Hate Speech policy. This includes exceptions for satirical content and where users share hateful content to condemn it or raise awareness.	Enforcement	Comprehensive	Progress reported
"Two Buttons" Meme #5	Q22021	Ensure appeals based on policy exceptions are prioritized for human review.	Enforcement	Comprehensive	Progress reported
Pro-Navalny protests in Russia #1	Q22021	Facebook should amend and redraft the Bullying & Harassment Community Standard to explain the relationship between the Policy Rationale and the "Do not's" as well as the other rules restricting content that follow it.	Content Policy	Comprehensive	Implementation demonstrated through published information
Pro-Navalny protests in Russia #2	Q22021	Differentiate between bullying and harassment and provide definitions that distinguish the two acts. Further, the Community Standard should clearly explain to users how bullying and harassment differ from speech that only causes offense and may be protected under international human rights law.	Content Policy	Somewhat Comprehensive	Recommendation declined after feasibility assessment



Pro-Navalny protests in Russia #3	Q22021	Clearly define its approach to different target user categories and provide illustrative examples of each target category (i.e. who qualifies as a public figure). Format the Community Standard on Bullying and Harassment by user categories currently listed in the policy.	Content Policy	Comprehensive	Progress reported
Pro-Navalny protests in Russia #4	Q22021	Include illustrative examples of violating and non-violating content in the Bullying and Harassment Community Standard to clarify the policy lines drawn and how these distinctions can rest on the identity status of the target.	Content Policy	Comprehensive	Progress reported
Pro-Navalny protests in Russia #5	Q22021	When assessing content including a 'negative character claim' against a private adult, Facebook should amend the Community Standard to require an assessment of the social and political context of the content. Facebook should reconsider the enforcement of this rule in political or public debates where the removal of the content would stifle debate.	Content Policy	Not Comprehensive	Recommendation declined after feasibility assessment
Pro-Navalny protests in Russia #6	Q22021	Whenever Facebook removes content because of a negative character claim that is only a single word or phrase in a larger post, it should promptly notify the user of that fact, so that the user can repost the material without the negative character claim.	Enforcement	Comprehensive	Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation



<p>Öcalan's Isolation #1</p> <p>Q32021</p>	<p>Immediately restore the misplaced 2017 guidance to the Internal Implementation Standards and Known Questions (the internal guidance for content moderators), informing all content moderators that it exists and arranging immediate training on it.</p>	<p>Enforcement</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Öcalan's Isolation #2</p> <p>Q32021</p>	<p>Evaluate automated moderation processes for enforcement of the Dangerous Individuals and Organizations policy. Where necessary, Facebook should update classifiers to exclude training data from prior enforcement errors that resulted from failures to apply the 2017 guidance. New training data should be added that reflects the restoration of this guidance.</p>	<p>Enforcement</p>	<p>Not Comprehensive</p>	<p>Recommendation omitted, declined, or reframed</p>
<p>Öcalan's Isolation #3</p> <p>Q32021</p>	<p>Publish the results of the ongoing review process to determine if any other policies were lost, including descriptions of all lost policies, the period the policies were lost for, and steps taken to restore them.</p>	<p>Transparency</p>	<p>Somewhat Comprehensive</p>	<p>Partial Implementation demonstrated through published information</p>



Öcalan's
Isolation #4

Q32021

Reflect in the Dangerous Individuals and Organizations “policy rationale” that respect for human rights and freedom of expression, in particular open discussion about human rights violations and abuses that relate to terrorism and efforts to counter terrorism, can advance the value of “Safety,” and that it is important for the platform to provide a space for these discussions. While “Safety” and “Voice” may sometimes be in tension, the policy rationale should specify in greater detail the “real-world harms” the policy seeks to prevent and disrupt when “Voice” is suppressed.

Content Policy

Comprehensive

Implementation
demonstrated
through published
information



Öcalan's
Isolation #5

Q32021

Add to the Dangerous Individuals and Organizations policy a clear explanation of what “support” excludes. Users should be free to discuss alleged violations and abuses of the human rights of members of designated organizations. This should not be limited to detained individuals. It should include discussion of rights protected by the UN human rights conventions as cited in Facebook’s Corporate Human Rights Policy. This should allow, for example, discussions on allegations of torture or cruel, inhuman, or degrading treatment or punishment, violations of the right to a fair trial, as well as extrajudicial, summary, or arbitrary executions, enforced disappearance, extraordinary rendition and revocation of citizenship rendering a person stateless. Calls for accountability for human rights violations and abuses should also be protected. Content that incites acts of violence or recruits people to join or otherwise provide material support to Facebook-designated organizations should be excluded from protection even if the same content also discusses human rights concerns. The user’s intent, the broader context in which they post, and how other users understand their post, is key to determining the likelihood of real-world harm that may result from such posts.

Content Policy

Comprehensive

Implementation
demonstrated
through published
information



<p>Öcalan's Isolation #6</p> <p>Q32021</p>	<p>Explain in the Community Standards how users can make the intent behind their posts clear to Facebook. This would be assisted by implementing the Board's existing recommendation to publicly disclose the company's list of designated individuals and organizations (see: case 2020-005-FB-UA). Facebook should also provide illustrative examples to demonstrate the line between permitted and prohibited content, including in relation to the application of the rule clarifying what "support" excludes.</p>	<p>Content Policy</p>	<p>Somewhat Comprehensive</p>	<p>Partial Implementation demonstrated through published information</p>
<p>Öcalan's Isolation #7</p> <p>Q32021</p>	<p>Ensure meaningful stakeholder engagement on the proposed policy change through Facebook's Product Policy Forum, including through a public call for inputs. Facebook should conduct this engagement in multiple languages across regions, ensuring the effective participation of individuals most impacted by the harms this policy seeks to prevent. This engagement should also include human rights, civil society, and academic organizations with expert knowledge on those harms, as well as the harms that may result from over-enforcement of the existing policy.</p>	<p>Content Policy</p>	<p>Somewhat Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Öcalan's Isolation #8</p> <p>Q32021</p>	<p>Ensure internal guidance and training is provided to content moderators on any new policy. Content moderators should be provided adequate resources to be able to understand the new policy, and adequate time to make decisions when enforcing the policy.</p>	<p>Enforcement</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>



Öcalan's Isolation # 9	Q32021	Ensure that 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 violated (and the jurisdictional reach of any removal).	Enforcement	Comprehensive	Progress reported
Öcalan's Isolation #10	Q32021	Clarify to Instagram users that Facebook's Community Standards apply to Instagram in the same way they apply to Facebook, in line with the recommendation in case 2020-004-IG-UA.	Content Policy	Comprehensive	Progress reported
Öcalan's Isolation #11	Q32021	Include information in its transparency reporting 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.	Transparency	Somewhat Comprehensive	Progress reported
Öcalan's Isolation #12	Q32021	In transparency reporting, include more comprehensive information on error rates for enforcing rules on "praise" and "support" of dangerous individuals and organizations, broken down by region and language.	Transparency	Somewhat Comprehensive	Recommendation declined after feasibility assessment
Myanmar Bot #1	Q32021	Facebook should ensure that its Internal Implementation Standards are available in the language in which content moderators review content. If necessary to prioritize, Facebook should focus first on contexts where the risks to human rights are more severe.	Enforcement	Somewhat Comprehensive	Recommendation omitted, declined, or reframed



<p>COVID-19 Lockdowns in Brazil #1</p> <p>Q32021</p>	<p>Facebook should conduct a proportionality analysis to identify a range of less intrusive measures than removing the content. When necessary, the least intrusive measures should be used where content related to COVID-19 distorts the advice of international health authorities and where a potential for physical harm is identified but is not imminent. Recommended measures include: (a) labeling content to alert users to the disputed nature of a post's content and to provide links to the views of the World Health Organization and national health authorities; (b) introducing friction to posts to prevent interactions or sharing; and (c) down-ranking, to reduce visibility in other users' News Feeds. All these enforcement measures should be clearly communicated to all users, and subject to appeal.</p>	<p>Enforcement</p>	<p>Not Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>COVID-19 Lockdowns in Brazil #2</p> <p>Q32021</p>	<p>Given the context of the COVID-19 pandemic, Facebook should make technical arrangements to prioritize fact-checking of potential health misinformation shared by public authorities which comes to the company's attention, taking into consideration the local context.</p>	<p>Enforcement</p>	<p>Not Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>



<p>COVID-19 Lockdowns in Brazil #3</p> <p>Q32021</p>	<p>Facebook should provide more transparency within the False News Community Standard regarding when content is eligible for fact-checking, including whether public institutions' accounts are subject to fact-checking.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Shared Al Jazeera Post #1</p> <p>Q32021</p>	<p>Add criteria and illustrative examples to its Dangerous Individuals and Organizations policy to increase understanding of the exceptions for neutral discussion, condemnation and news reporting.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Progress reported</p>
<p>Shared Al Jazeera Post #2</p> <p>Q32021</p>	<p>Ensure swift translation of updates to the Community Standards into all available languages.</p>	<p>Content Policy</p>	<p>Somewhat Comprehensive</p>	<p>Recommendation declined after feasibility assessment</p>



<p>Shared AI Jazeera Post #3</p> <p>Q32021</p>	<p>Engage an independent entity not associated with either side of the Israeli-Palestinian conflict to conduct a thorough examination to determine whether Facebook’s content moderation in Arabic and Hebrew, including its use of automation, have been applied without bias. This examination should review not only the treatment of Palestinian or pro-Palestinian content, but also content that incites violence against any potential targets, no matter their nationality, ethnicity, religion or belief, or political opinion. The review should look at content posted by Facebook users located in and outside of Israel and the Palestinian Occupied Territories. The report and its conclusions should be made public.</p>	<p>Transparency</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>
<p>Shared AI Jazeera Post #4</p> <p>Q32021</p>	<p>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.</p>	<p>Transparency</p>	<p>Comprehensive</p>	<p>Progress reported</p>
<p>Colombia Protests #1</p> <p>Q32021</p>	<p>Publish illustrative examples from the list of slurs designated as violating under its Hate Speech Community Standard, including borderline cases with words which may be harmful in some contexts but not others.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Recommendation declined after feasibility assessment</p>



<p>Colombia Protests #2</p> <p>Q32021</p>	<p>Link the short explanation of the newsworthiness allowance provided in the introduction to the Community Standards to the more detailed Transparency Center explanation of how this policy applies. The company should supplement this explanation with illustrative examples from a variety of contexts, including reporting on large scale protests.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>
<p>Colombia Protests #3</p> <p>Q32021</p>	<p>Develop and publicize clear criteria for content reviewers for escalating for additional review public interest content that potentially violates the Community Standards but may be eligible for the newsworthiness allowance. These criteria should cover content depicting large protests on political issues.</p>	<p>Enforcement</p>	<p>Not Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>
<p>Colombia Protests #4</p> <p>Q32021</p>	<p>Notify all users who reported content which was assessed as violating but left on the platform for public interest reasons that the newsworthiness allowance was applied to the post. The notice should link to the Transparency Center explanation of the newsworthiness allowance.</p>	<p>Enforcement</p>	<p>Comprehensive</p>	<p>Progress reported</p>



South Africa Slurs #1 Q32021	Notify users of the specific rule within the Hate Speech Community Standard that has been violated in the language in which they use Facebook, as recommended in case decision 2020-003-FB-UA (Armenians in Azerbaijan) and case decision 2021-002-FB-UA (Depiction of Zwarte Piet). In this case, for example, the user should have been notified they violated the slurs prohibition. The Board has noted Facebook's response to Recommendation No. 2 in case decision 2021-002-FB-UA, which describes a new classifier that should be able to notify English-language Facebook users their content has violated the slur rule. The Board looks forward to Facebook providing information that confirms implementation for English-language users and information about the timeframe for implementation for other language users.	Enforcement	Comprehensive Progress reported
Wampum Belt #1 Q42021	Provide users with timely and accurate notice of any company action being taken on the content their appeal relates to. Where applicable, including in enforcement error cases like this one, the notice to the user should acknowledge that the action was a result of the Oversight Board's review process. Meta should share the user messaging sent when Board actions impact content decisions appealed by users, to demonstrate it has complied with this recommendation. These actions should be taken with respect to all cases that are corrected at the eligibility stage of the Board's process.	Enforcement	Comprehensive Progress reported



Wampum
Belt #2

Q42021

Study the impacts of modified approaches to secondary review on reviewer accuracy and throughput. In particular, the Board requests an evaluation of accuracy rates when content moderators are informed that they are engaged in secondary review, so they know the initial determination was contested. This experiment should ideally include an opportunity for users to provide relevant context that may help reviewers evaluate their content, in line with the Board's previous recommendations. Meta should share the results of these accuracy assessments with the Board and summarize the results in its quarterly Board transparency report to demonstrate it has complied with this recommendation.

Enforcement

Comprehensive

Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation



Wampum
Belt #3

Q42021

Conduct accuracy assessments focused on Hate Speech policy allowances that cover artistic expression and expression about human rights violations (e.g., condemnation, awareness raising, self-referential use, empowering use). This assessment should also specifically investigate how the location of a reviewer impacts the ability of moderators to accurately assess hate speech and counter speech from the same or different regions. The Board understands this analysis likely requires the development of appropriate and accurately labelled samples of relevant content. Meta should share the results of this assessment with the Board, including how these results will inform improvements to enforcement operations and policy development and whether it plans to run regular reviewer accuracy assessments on these allowances, and summarize the results in its quarterly Board transparency report to demonstrate it has complied with this recommendation.

Enforcement

Somewhat
Comprehensive

Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation



Ayahuasca Brew #1	Q42021	The Board reiterates its recommendation from case decision 2020-004-IG-UA and case decision 2021-006-IG-UA that Meta should explain to users that it enforces the Facebook Community Standards on Instagram, with several specific exceptions. The Board notes Meta’s response to these recommendations. While Meta may be taking other actions to comply with the recommendations, the Board recommends Meta update the introduction to the Instagram Community Guidelines (“The Short” Community Guidelines) within 90 days to inform users that if content is considered violating on Facebook, it is also considered violating on Instagram, as stated in the company’s Transparency Center, with some exceptions.	Enforcement	Somewhat Comprehensive	Progress reported
Ayahuasca Brew #2	Q42021	The Board reiterates its recommendation from case decision 2021-005-FB-UA and case decision 2020-005-FB-UA that Meta should explain to users precisely what rule in a content policy they have violated.	Enforcement	Somewhat Comprehensive	Progress reported
Ayahuasca Brew #3	Q42021	To respect diverse traditional and religious expressions and practices, the Board recommends that Meta modify the Instagram Community Guidelines and Facebook Regulated Goods Community Standard to allow positive discussion of traditional and religious uses of non-medical drugs where there is historic evidence of such use. The Board also recommends that Meta make public all allowances, including existing allowances.	Content Policy	Comprehensive	Progress reported



<p>Alleged Crimes in Raya Kobo #1</p> <p>Q42021</p>	<p>Meta should rewrite Meta’s value of “Safety” to reflect that online speech may pose risk to the physical security of persons and the right to life, in addition to the risks of intimidation, exclusion and silencing.</p>	<p>Content Policy</p>	<p>Comprehensive</p>	<p>Implementation demonstrated through published information</p>
<p>Alleged Crimes in Raya Kobo #2</p> <p>Q42021</p>	<p>Facebook’s Community Standards should reflect that in the contexts 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.</p>	<p>Content Policy</p>	<p>Not Comprehensive</p>	<p>Recommendation omitted, declined, or reframed</p>
<p>Alleged Crimes in Raya Kobo #3</p> <p>Q42021</p>	<p>Meta should 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. The assessment should cover a period from June 1, 2020, to the present. The company should complete the assessment within six months from the moment it responds to these recommendations. The assessment should be published in full.</p>	<p>Transparency</p>	<p>Somewhat Comprehensive</p>	<p>Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation</p>



Asking for Adderall #1	Q12022	Meta should publish its internal definitions for “non-medical drugs” and “pharmaceutical drugs” in the Facebook Community Standard on Restricted Goods and Services. The published definitions should: (a) make clear that certain substances may fall under either “non-medical drugs” or “pharmaceutical drugs” and (b) explain the circumstances under which a substance would fall into each of these categories. The Board will consider this recommendation implemented when these changes are made in the Community Standard.	Content Policy	Comprehensive	Progress reported
Asking for Adderall #2	Q12022	Meta should study the consequences and trade-offs of implementing a dynamic prioritization system that orders appeals for human review, and consider whether the fact that an enforcement decision resulted in an account restriction should be a criterion within this system. The Board will consider this recommendation implemented when Meta shares the results of these investigations with the Board and in its quarterly Board transparency report.	Enforcement	Comprehensive	Progress reported
Asking for Adderall #3	Q12022	Meta should conduct regular assessments on reviewer accuracy rates focused on the Restricted Goods and Services policy. The Board will consider this recommendation implemented when Meta shares the results of these assessments with the Board, including how these results will inform improvements to enforcement operations and policy development, and summarize the results in its quarterly Board transparency reports . Meta may consider if these assessments should be extended to reviewer accuracy rates under other Community Standards.	Enforcement	Not Comprehensive	Recommendation omitted, declined, or reframed



Swedish Journalist Reporting Sexual Violence Against Minors #1	Q12022 Meta should define graphic depiction and sexualization in the Child Sexual Exploitation, Nudity and Abuse Community Standard. Meta should make clear that not all explicit language constitutes graphic depiction or sexualization and explain the difference between legal, clinical or medical terms and graphic content. Meta should also provide a clarification for distinguishing child sexual exploitation and reporting on child sexual exploitation. The Board will consider the recommendation implemented when language defining key terms and the distinction has been added to the Community Standard.	Content Policy	Comprehensive	Progress reported
Swedish Journalist Reporting Sexual Violence Against Minors #2	Q12022 Meta should undergo a policy development process, including as a discussion in the Policy Forum, to determine whether and how to incorporate a prohibition on functional identification of child victims of sexual violence in its Community Standards. This process should include stakeholder and expert engagement on functional identification and the rights of the child. The Board will consider this recommendation implemented when Meta publishes the minutes of the Product Policy Forum where this is discussed.	Content Policy	Comprehensive	Progress reported



PAO on Sharing Private Residential Information #1

Q12022

removing the “publicly available” policy exception. Meta should remove the exception that allows the sharing of private residential information (both images that currently fulfill the Privacy Violations policy’s criteria for takedown and addresses) when considered “publicly available”. This means Meta would no longer allow otherwise violating content on Facebook and on Instagram if “published by at least five news outlets” or if it contains residential addresses or imagery from financial records or statements of an organization, court records, professional and business licenses, sex offender registries or press releases from government agencies, or law enforcement. The Board will consider this implemented when Meta modifies its Internal Implementation Standards and its content policies.

Content policy

Comprehensive

Progress reported

PAO on Sharing Private Residential Information #2

Q12022

Making sure the newsworthiness exception is consistently applied. Meta should develop and publicize clear criteria for content reviewers to escalate for additional review of public interest content that potentially violates the Community Standards but may be eligible for the newsworthiness exception, as previously recommended in case decision 2021-010-FB-UA. These criteria should cover, in addition to large protests as highlighted in Recommendation No. 3 from case decision 2021-010-FB-UA, content that shares information on investigations of corruption, tax evasion, money laundering and other criminal activities, in particular where property has been attained through such activities. The Board will consider this implemented when Meta publicly shares these escalation criteria.

Content Policy

Not Comprehensive

Recommendation omitted, declined, or reframed



PAO on Sharing Private Residential Information #3

Q12022

keeping public interest information in the form of images on the platform [this does not apply to addresses]. Meta should allow the sharing of “imagery that displays the external view of private residences” when the property depicted is the focus of the news story, even when the following conditions listed in the Privacy Violations Community Standard are met (“the residence is a single-family home, or the resident's unit number is identified in the image/caption”; “the city/neighborhood or GPS pin” is identified; “the content identifies the resident(s)”; and “that same resident objects to the exposure of their private residence”). However, Meta should not allow the sharing of such information when there is a “context of organizing protests against the resident,” that is, an attempt to organize protests in the future, and not news reporting on protests that have already taken place. The Board will consider this implemented when Meta modifies its content policies.

PAO on Sharing Private Residential Information #4

Q12022

Allowing the organization of protests at publicly owned official residences. Meta should allow the publication of addresses and imagery of official residences provided to high-ranking government officials, such as heads of state, heads of federal or local government, ambassadors and consuls. The Board will consider this implemented when Meta modifies its content policies.

Content Policy

Comprehensive

Progress reported

Content Policy

Somewhat Comprehensive

Progress reported



PAO on Sharing Private Residential Information #5	Q12022	Strengthening the role of user consent. Meta should allow the resharing of private residential addresses when posted by the affected user themselves or when the user consented to its publication. Users should not be presumed to consent to private information posted by others. The Board will consider this implemented when Meta modifies its content policies.	Content Policy	Somewhat Comprehensive	Recommendation omitted, declined, or reframed
PAO on Sharing Private Residential Information #6	Q12022	Strengthening the role of user consent. Users should have a quick and effective mechanism to request the removal of private information posted by others. We will consider this implemented when Meta demonstrates in its transparency reports that user requests to remove their information are consistently and promptly actioned. This recommendation is not applicable to official residences of high-ranking government officials.	Content Policy	Comprehensive	Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation
PAO on Sharing Private Residential Information #7	Q12022	Improving the clarity of the rules. Meta should better explain, in the text of Facebook's Privacy Violations policy, when disclosing the city where a residence is located will suffice for the content to be removed, and when disclosing its neighborhood would be required for the same matter (e.g., by specifically referencing the population threshold at which sharing only the city as part of the content will no longer be considered violating). The Board will consider this implemented when Meta modifies its content policies.	Content Policy	Somewhat Comprehensive	Progress reported



PAO on Sharing Private Residential Information #8	Q12022	Improving the clarity of the rules. Meta should explain, in the text of Facebook’s Privacy Violations policy, its criteria for assessing whether the resident is sufficiently identified in the content. The Board will consider this implemented when Meta modifies its content policies.	Content Policy	Comprehensive	Progress reported
PAO on Sharing Private Residential Information #9	Q12022	Improving the clarity of the rules. The Board reiterates Recommendation No. 1 from case decision 2021-013-IG-UA that Meta should explain to users that it enforces the Facebook Community Standards on Instagram, with several specific exceptions. The Board notes Meta’s response to these recommendations. While Meta may be taking other actions to comply with the recommendations, the Board recommends Meta update the introduction to the Instagram Community Guidelines (“The Short” Community Guidelines) within 90 days to inform users that if content is considered violating on Facebook, it is also considered violating on Instagram, as stated in the company’s Transparency Center, with some exceptions. Meta should also provide a link to the Privacy Violations Community Standard in the language of the Community Guidelines. The Board will consider this implemented when Meta modifies its content policies.	Content Policy	Comprehensive	Progress reported
PAO on Sharing Private Residential Information #10	Q12022	Responding more quickly to situations of increased risk. Meta should let users reporting content that may violate the Privacy Violations policy provide additional context about their claim. The Board will consider this implemented when Meta publishes information about its appeal processes that demonstrate users may provide this context in appeals.	Enforcement	Somewhat Comprehensive	Progress reported



<p>PAO on Sharing Private Residential Information #11</p> <p>Q12022</p>	<p>Responding more quickly to situations of increased risk. Meta should create a specific channel of communications for victims of doxing (available both for users and non-users) that may be easily accessed, allow the victim to explain in detail their situation and risks the content creates for them, and prompt swift action from the company. Additionally, Meta could provide financial support to organizations that already have hotlines in place. Meta should prioritize action when the impacted person references belonging to a group facing heightened risk to safety in the region where the private residence is located. The Board will consider this implemented when Meta creates the channel and publicly announces how to use it.</p>	<p>Enforcement</p>	<p>Somewhat Comprehensive</p>	<p>Recommendation omitted, declined, or reframed</p>
<p>PAO on Sharing Private Residential Information #12</p> <p>Q12022</p>	<p>Responding more quickly to situations of increased risk. Meta should consider the violation of its Privacy Violations policy as “severe,” prompting temporary account suspension, in cases where the sharing of private residential information is clearly related to malicious action that created a risk of violence or harassment. The Board will consider this implemented when Meta updates its Transparency Center description of the strikes system to make clear that some Privacy Violations are severe and may result in account suspension.</p>	<p>Enforcement</p>	<p>Somewhat Comprehensive</p>	<p>Progress reported</p>



PAO on Sharing Private Residential Information #13	Q12022 <p>Reversing enforcement errors. Meta should give users an opportunity to remove or edit private information within their content following a removal for violation of the Privacy Violations policy (i.e., a threat of harm is present, but the story could be allowed on the platform if the image of residence or other information, such as the address, the city/neighborhood, GPS pin, or the name and picture of the resident was removed). The company could issue a notification of the violation and give the user a short deadline for them to act on the content, during which the content should be made temporarily unavailable. If the user removes/edits the private residential information out of the content within the deadline, the temporary block on the content would be lifted. If not, then the content would remain blocked. The Board will consider this implemented when Meta publishes information about its enforcement processes that demonstrates users are notified of specific policy violations when content is removed and granted a remedial window before the content is permanently deleted.</p>	Enforcement	Comprehensive	Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation
PAO on Sharing Private Residential Information #14	Q12022 <p>Reversing enforcement errors. Meta should let users indicate in their appeals against content removal that their content falls into one of the exceptions to the Privacy Violations policy. The Board will consider this implemented when Meta publishes information about its appeal processes that demonstrates users may provide this information in appeals.</p>	Enforcement	Somewhat Comprehensive	Progress reported



PAO on Sharing Private Residential Information #15	Q12022	Improving enforcement transparency. Meta should publish quantitative data on the enforcement of the Privacy Violations policy in the company's Community Standards Enforcement Report. The Board will consider this implemented when Meta's transparency report includes Privacy Violations enforcement data.	Transparency	Comprehensive	Recommendation omitted, declined, or reframed
PAO on Sharing Private Residential Information #16	Q12022	Improving enforcement transparency. Meta should break down data in its transparency reports to indicate the amount of content removed following privacy-related government requests, even if taken down under the Privacy Violations policy and not under local privacy laws. The Board will consider this implemented when Meta's transparency reporting includes all government requests that result in content removal for violating the Privacy Violations policy as a separate category.	Transparency	Comprehensive	Recommendation omitted, declined, or reframed
PAO on Sharing Private Residential Information #17	Q12022	Improving enforcement transparency. Meta should provide users with more detail on the specific policy of the Privacy Violations Community Standard that their content was found to violate and implement it across all working languages of the company's platforms. The Board will consider this implemented when Meta publishes information and data about user notifications.	Transparency	Comprehensive	Progress reported



Sudan Graphic Video #1	Q22022	Meta should amend the Violent and Graphic Content Community Standard to allow videos of people or 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 will consider this recommendation implemented when Meta updates the Community Standard	Content Policy	Comprehensive	Progress reported
Sudan Graphic Video #2	Q22022	Meta should 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. The Board will consider this recommendation implemented when Meta publishes the findings of the policy development process, including information on the process and criteria for identifying this content at scale.	Content Policy	Comprehensive	Progress reported
Sudan Graphic Video #3	Q22022	Meta should make explicit in its description of the newsworthiness allowance all the actions it may take (for example, restoration with a warning screen) based on this policy. The Board will consider this recommendation implemented when Meta updates the policy.	Content Policy	Comprehensive	Implementation demonstrated through published information



Sudan Graphic Video #4	Q22022	<p>To ensure users understand the rules, Meta should notify users when it takes action on their content based on the newsworthiness allowance including the restoration of content or application of a warning screen. The user notification may link to the Transparency Center explanation of the newsworthiness allowance. The Board will consider this implemented when Meta rolls out this updated notification to users in all markets and demonstrates that users are receiving this notification through enforcement data.</p>	Enforcement	Comprehensive	Progress reported
Reclaiming Arabic Words #1	Q22022	<p>Meta should translate the Internal Implementation Standards and Known Questions to Modern Standard Arabic. Doing so could reduce over-enforcement in Arabic-speaking regions by helping moderators better assess when exceptions for content containing slurs are warranted. The Board notes that Meta has taken no further action in response to the recommendation in the "Myanmar Bot" case (2021-007-FB-UA) that Meta should ensure that its Internal Implementation Standards are available in the language in which content moderators review content. The Board will consider this recommendation implemented when Meta informs the Board that translation to Modern Standard Arabic is complete.</p>	Enforcement	Comprehensive	<p>Recommendation omitted, declined, or reframed</p>
Reclaiming Arabic Words #2	Q22022	<p>Meta should publish a clear explanation on how it creates its market-specific slur lists. This explanation should include the processes and criteria for designating which slurs and countries are assigned to each market-specific list. The Board will consider this implemented when the information is published in the Transparency Center.</p>	Transparency	Comprehensive	<p>Partial Implementation demonstrated through published information</p>



Reclaiming Arabic Words #3	Q22022	Meta should publish a clear explanation of how it enforces its market-specific slur lists. This explanation should include the processes and criteria for determining precisely when and where the slurs prohibition will be enforced, whether in respect to posts originating geographically from the region in question, originating outside but relating to the region in question, and/or in relation to all users in the region in question, regardless of the geographic origin of the post. The Board will consider this recommendation implemented when the information is published in Meta's Transparency Center.	Transparency	Comprehensive	Implementation demonstrated through published information
Reclaiming Arabic Words #4	Q22022	Meta should publish a clear explanation on how it audits its market-specific slur lists. This explanation should include the processes and criteria for removing slurs from or keeping slurs on Meta's market-specific lists. The Board will consider this recommendation implemented when the information is published in Meta's Transparency Center.	Transparency	Comprehensive	Implementation demonstrated through published information
Knin Cartoon #1	Q22022	Meta should clarify the Hate Speech Community Standard and the guidance provided to reviewers, explaining that even implicit references to protected groups are prohibited by the policy when the reference would reasonably be understood. The Board will consider this recommendation implemented when Meta updates its Community Standards and Internal Implementation Standards to content reviewers to incorporate this revision.	Content Policy	Comprehensive	Progress reported



Knin
Cartoon #2

Q22022

In line with Meta's commitment following the "Wampum belt" case (2021-012-FB-UA), the Board recommends that Meta notify all users who have reported content when, on subsequent review, it changes its initial determination. Meta should also disclose the results of any experiments assessing the feasibility of introducing this change with the public. The Board will consider this recommendation implemented when Meta shares information regarding relevant experiments and, ultimately, the updated notification with the Board and confirms it is in use in all languages.

Enforcement

Somewhat
Comprehensive

Progress reported