



Oversight Board Q3 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 with membership rotating every three months, 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-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.

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. More detail about this stage can be found in the [Rulebook for Case Review and Policy Guidance](#) (page 8).

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 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.

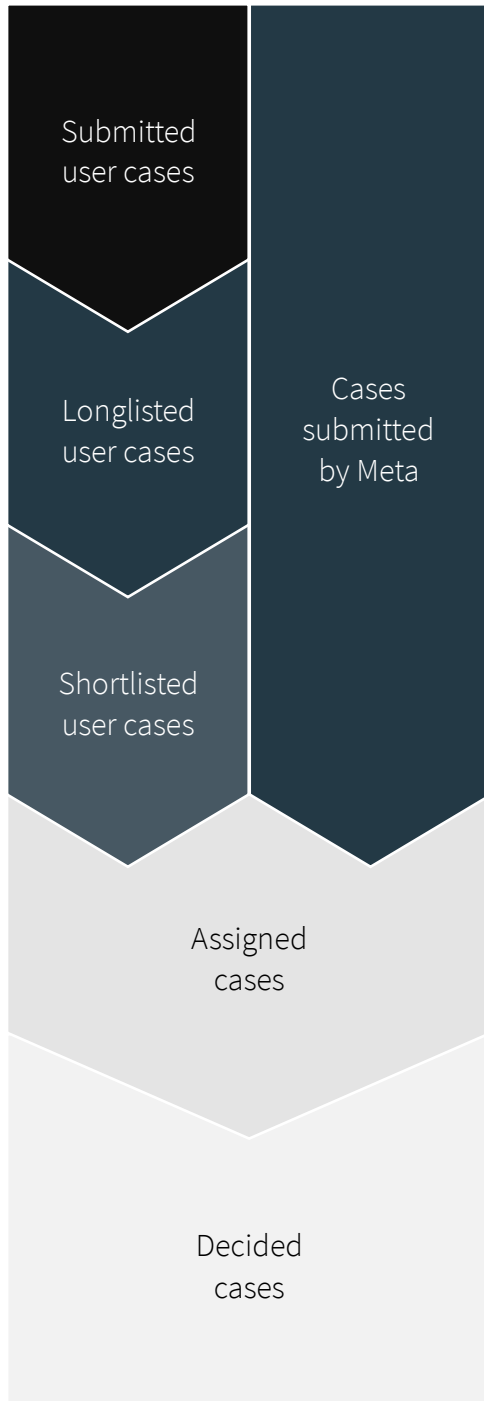


Transparency Report for third quarter of 2021

This transparency report for the third quarter of 2021 (July 1 – September 30, 2021) sets out key statistics on cases selected by the Board, as well as the decisions and recommendations we made in this quarter.



Key Findings – Third quarter 2021¹



~339,325 cases submitted to the Oversight Board, eight of which were submitted by Meta.

50 user-submitted cases longlisted.

28 user-submitted cases shortlisted.

Three cases assigned to panel, all of which were submitted by users.

Six cases decided, with 25 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. The next footnote provides further context on the numbers of submitted cases.



Q3 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.

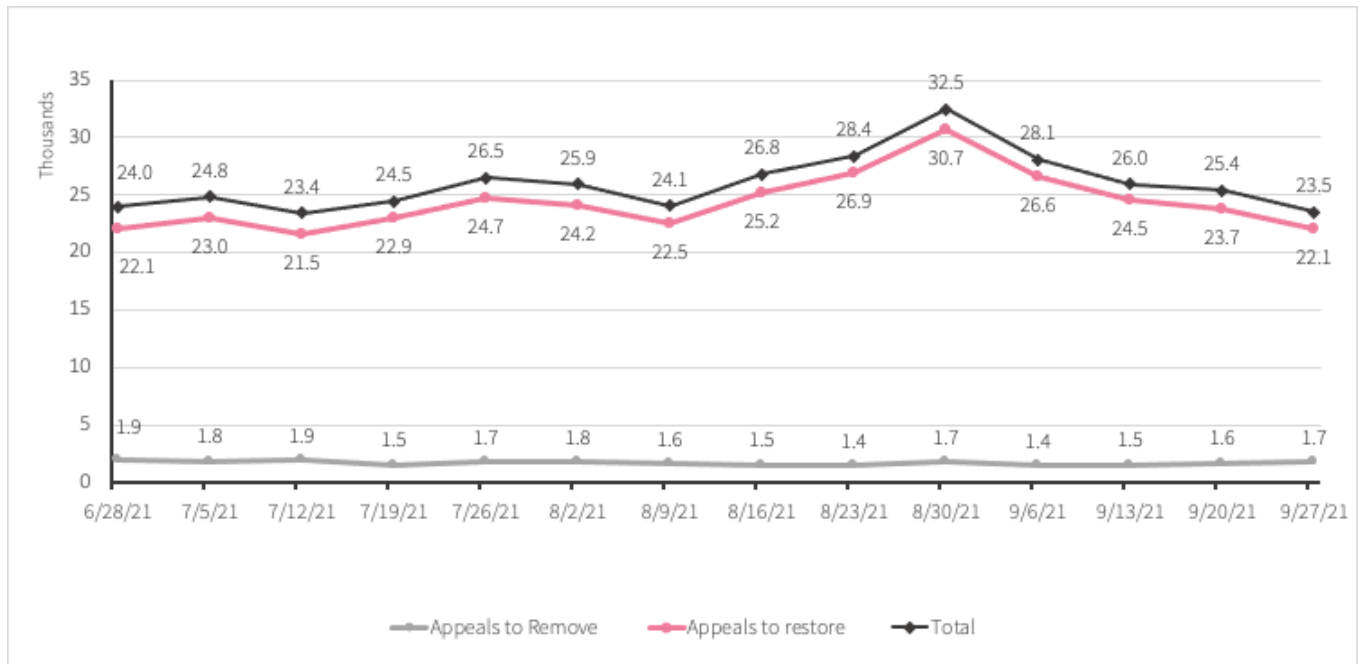
In this quarter, due to limitations in the functionality of the Case Management Tool (CMT), submitted cases have been counted manually by the Case Selection Team as they were seen in CMT at the time. As such, these numbers should be taken as an estimate and preliminary.

In Q3 2021, an estimated **339,317 cases** were submitted by users. This represents an increase of around 64% on the 207,235 cases submitted by users in Q2.

It is possible that improvements to how users appeal through Facebook’s mobile application as well as the growing profile of the Board contributed to this significant increase.

Estimated number of cases submitted to Oversight Board by week

Number of cases, thousands

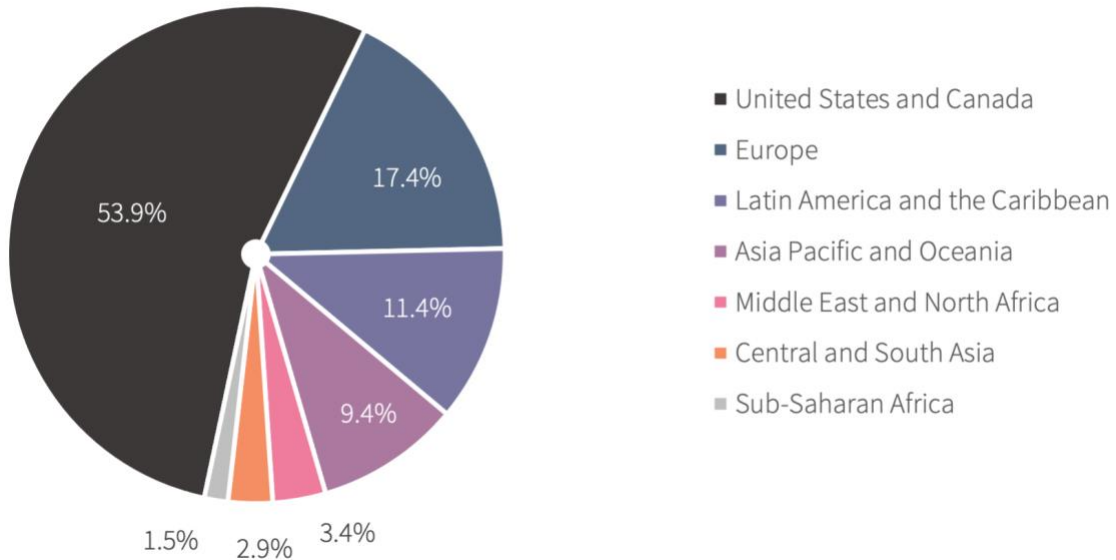




The majority of submitted cases were from the United States & Canada (53.9%) and Europe (17.4%).

Estimated cases submitted by user-selected region

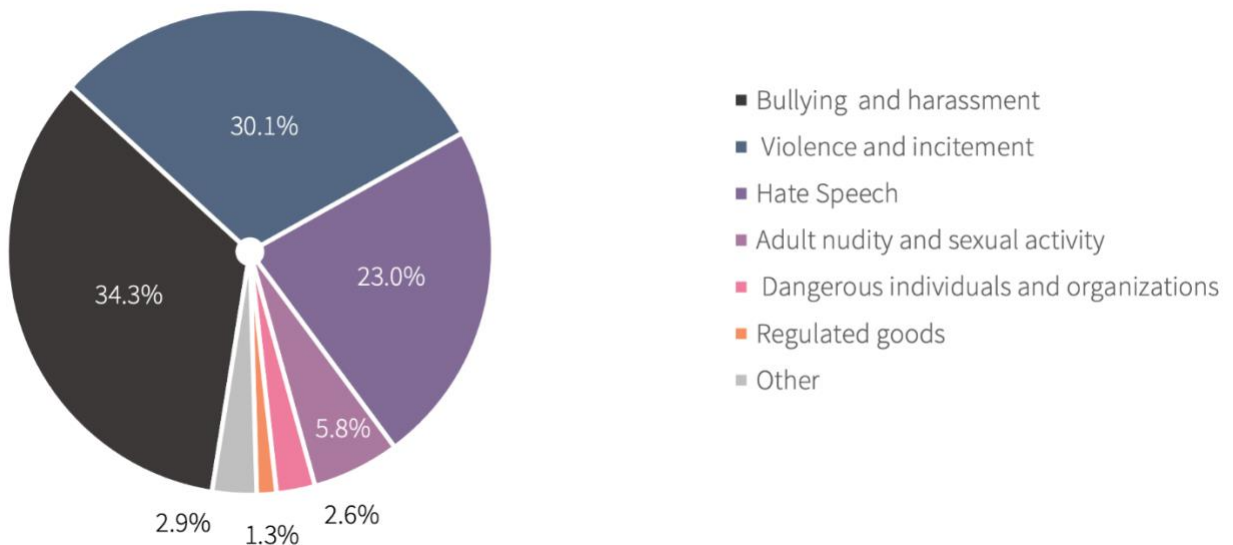
Percent



Users primarily submitted cases concerning Facebook’s Bullying and Harassment (34.3%), Violence and Incitement (30.1%) and Hate Speech (23.0%) policies. The share of cases related to Facebook’s rules on Violence and Incitement increased by two-thirds between Q2 and Q3, rising from around 18% in Q2 to around 30% in this quarter. The chart below only covers cases related to content that has been removed from Facebook and Instagram, and not content which is currently on the platforms as it supposedly does not violate a Community Standard.

Estimated cases submitted by Community Standard

Percent





Almost all cases submitted by users (99.3%) concerned content shared on Facebook, with only 0.7% of cases concerning content shared on Instagram.

Cases submitted by platform

Percent

Facebook	99.3%
Instagram	0.7%

Q3 Longlisted User Cases

In this period, **50 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 45 pieces of content on Facebook and five pieces of content on Instagram. 30 were cases to restore content and 20 were cases to remove content. The cases involved 50 posts and no comments. The United States, Afghanistan and China had the highest number of longlisted cases.

Longlisted cases by region and country²

Number of cases

Middle East & North Africa	3
Iran	1
Israel	2
Europe	6
Belarus	1
Belgium	2
France	1
Romania	1
Sweden	1
United States & Canada	14
Canada	2
United States	12
Sub-Saharan Africa	10
Ethiopia	2
Mauritania	1
Nigeria	1
Somalia	2
South Africa	1
Sudan	2

² 'Countries Affected' is a user-selected field in the Board's appeal process and users can select multiple countries. For this reason, while only 50 cases were longlisted in this quarter, the total for this table is 60. 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.



Tanzania	1
Central & South Asia	10
Afghanistan	8
India	2
Latin America & the Caribbean	4
Argentina	1
Brazil	1
Colombia	1
Mexico	1
Asia Pacific & Oceania	13
China	3
Japan	1
Malaysia	2
Singapore	2
South Korea	1
Taiwan	2
Thailand	1
Vietnam	1
Total	60

Two-fifths (40%) of longlisted cases did not have a Community Standard, as they concerned content which had been left up on Facebook or Instagram and had not been deemed to have violated Meta’s rules. A fifth (20%) of longlisted cases concerned Facebook’s Dangerous Individuals and Organizations policy, while around another fifth (18%) concerned Facebook’s rules on Hate Speech.

Longlisted cases by Community Standard

Number of cases

Undefined ³	20
Dangerous Individuals and Organizations	10
Hate Speech	9
Bullying and Harassment	3
Regulated Goods	3
Violence and Incitement	3
Violent and Graphic Content	2
Grand Total	50

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.



Q3 Shortlisted User Cases

The Case Selection Committee identifies a shortlist of cases from the longlist to consider for selection. Board Members on the Case Selection Committee rotate every three months, evaluating and selecting cases by majority vote.

In this quarter, the Case Selection Committee reviewed and shortlisted cases on five occasions. In this period, the committee **shortlisted 28 cases**.

The shortlist is sent to Meta’s legal team to review for legal obligations, as per the Bylaws. Such obligations may mean a case is 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, including deciding to restore content wrongly removed. This does not affect the eligibility of the content under review.

Of the 28 cases shortlisted in this period, all were confirmed as eligible by Meta’s legal team. Meta determined that in around half of these cases (13 out of 28 cases shortlisted by the Board in this quarter) its original decision on the piece of content was incorrect (See Annex). In all of these cases, Meta then reversed its original decision and restored the content.

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>
Bullying and Harassment	2	1	3
Dangerous Individuals and Organizations	3	0	3
Hate Speech	5	0	5
Regulated Goods	1	0	1
Violence and Incitement	1	0	1
Total	12	1	13

Q3 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. The Board also considered three referrals from Meta from the previous period. As such, 11 referrals from Meta were considered during this period.

While two cases were related to Facebook’s Adult Nudity and Sexual Activity Community Standard, the remaining nine cases concerned content which had been left up on Facebook or Instagram and thus had not been deemed to violate a Community Standard.



Meta referrals considered⁴

<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>
2021-016-FB-FBR	Swedish journalist reporting sexual violence against minors	Take down	FB	Swedish	Adult Nudity and Sexual Activity	Sweden
N/A	N/A	Take down	FB	Greek	Adult Nudity and Sexual Activity	Greece
N/A	N/A	Keep up	FB	French	N/A	Chad/France
N/A	N/A	Keep up	FB	English	N/A	Cameroon
N/A	N/A	Keep up	IG	Spanish	N/A	Cuba
N/A	N/A	Keep up	FB	English	N/A	Australia
N/A	N/A	Keep up	FB	English	N/A	United States
N/A	N/A	Keep up	FB	Arabic	N/A	Jordan
N/A	N/A	Keep up	FB	Latvian	N/A	Latvia/ Belarus

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.



N/A	N/A	Keep up	FB	Arabic	N/A	Egypt
N/A	N/A	Keep up	FB	English	N/A	United States

Q3 Assigned Cases

The Case Selection Committee assigns cases to panels. In this period, the Case Selection Committee assigned three cases to panel.

Cases assigned

<i>Case ID</i>	<i>Name</i>	<i>Date</i>	<i>Platform</i>	<i>Source</i>	<i>Community Standard</i>	<i>Countries</i>
2021-012-FB-UA	Wampum belt	9/1/21	FB	User (appeal to restore)	Hate Speech	Canada
2021-013-IG-UA	Ayahuasca brew	9/1/21	IG	User (appeal to restore)	Regulated Goods	Brazil
2021-014-FB-UA	Alleged crimes in Raya Kobo	9/1/21	FB	User (appeal to restore)	Hate Speech	Ethiopia



Q3 Decided Cases

After being selected, the Board assigns cases to a five-member panel. Members of the panel are randomly chosen but 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 the company’s 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 six cases**. The Oversight Board overturned Meta’s decision four times and upheld Meta’s decision twice.

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>Countries⁶</i>	<i>Outcome</i>
2021-006-IG-UA	Ocalan’s isolation	IG	User	English	Dangerous Individuals and Organizations	Turkey/ United States	Overturned
2021-007-FB-UA	Myanmar bot	FB	User	Burmese	Hate Speech	Myanmar	Overturned
2021-008-FB-FBR	COVID lockdowns in Brazil	FB	Meta	Portuguese	N/A	Brazil	Upheld

⁶ 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.



2021-009-FB-UA	Shared Al Jazeera post	FB	User	Arabic	Dangerous Individuals and Organizations	Israel/Egypt	Overturned
2021-010-FB-UA	Colombia protests	FB	User	Spanish	Hate Speech	Colombia	Overturned
2021-011-FB-UA	South Africa slurs	FB	User	English	Hate Speech	South Africa	Upheld

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	2021-006-IG-UA	2021-007-FB-UA	2021-008-FB-FBR	2021-009-FB-UA	2021-010-FB-UA	2021-011-FB-UA
UN Treaties						
ICCPR⁷						
Non-Discrimination & Remedies (Art. 2)	✓			✓	✓	✓
Life (Art. 6)			✓	✓		
Liberty and security of person (Art. 9)				✓		
Expression (Art. 19)	✓	✓	✓	✓	✓	✓
Peaceful Assembly (Art. 21)					✓	
Equality (Art. 26)				✓	✓	✓
ICERD⁸						
Elimination of Discrimination (Art. 2)						✓
Freedom of Expression (Art. 5)						✓
ICESCR⁹						
Health (Art. 12)			✓			

⁷ International Covenant on Civil and Political Rights

⁸ International Convention on the Elimination of All Forms of Racial Discrimination

⁹ International Covenant on Economic, Social and Cultural Rights



UN Treaty Bodies: Guidance & Recommendations

Human Rights Committee

General Comment 18 on non-discrimination						✓
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)	✓	✓	✓	✓	✓	✓
General Comment 35 on the Right to Liberty & Security of Person (CCPR/C/GC/35)					✓	
General Comment 36 on the Right to Life (CCPR/C/GC/36)					✓	
General Comment 37 on the Right of Peaceful Assembly (CCPR/C/GC/37)						✓
Committee on Economic, Social & Cultural Rights						
General Comment 14 on the Right to the Highest Attainable Standard of Health				✓		
Committee on the Elimination of Racial Discrimination						
General Recommendation 35 on combating racist hate speech (CERD/C/GC/35)						✓

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), UN Working Group on the issue of human rights and transnational corporations and other business enterprises					✓	
UN SR Reports						
Online Hate Speech (A/74/486)	✓		✓	✓	✓	✓
Online Content Regulation (A/HRC/38/35)	✓		✓		✓	✓
Disease pandemics and the freedom of opinion and expression (A/HRC/44/49)			✓			
Disinformation and freedom of opinion and expression (A/HRC/47/25)			✓			
Special Rapporteur on human rights and counter-terrorism (A/HRC/40/52)	✓					



Special Rapporteur on freedom of peaceful assembly and of association (A/HRC/41/41)		✓
Situation of human rights in the Palestinian territories occupied since 1967 (A/75/532)		✓
Joint Declarations and Statements of UN & Regional Freedom of Expression Mandates "Fake News," Disinformation & Propaganda	✓	
UN Independent Expert Reports on protection against violence and discrimination based on sexual orientation and gender identity (A/HRC/35/36)		✓
(A/HRC/38/43)		✓
UN Human Rights Council Resolution 32/2 on the protection against violence and discrimination based on sexual orientation and gender identity		✓
UN Declaration on Human Rights Defenders (Art. 6) (A/RES/53/144)	✓	
Resolution on the Safety of Journalists (A/HRC/RES/45/18)		✓
Prohibition on torture, cruel, inhuman or degrading treatment or punishment		
Rule 43, UN Standard Minimum Rules for the Treatment of Prisoners (the Nelson Mandela Rules) (A/RES/70/175)	✓	

Decision timelines

The Bylaws which applied to these cases specified that, apart from exceptional circumstances, decisions and implementation will occur a maximum of 90 days from the date the case is selected for review by the Case Selection Committee. For the six decisions the Oversight Board published in this quarter, the average number of days from assignment of case to implementation of the Board's decision was 87 days.

Case ID	Name	Beginning of 90-day period	Board's decision published	Meta implements decision	Number of days taken out of 90 days
2021-006-IG-UA	Ocalan's isolation	4/14/21	7/8/21	N/A (Meta already restored content)	85



2021-007-FB-UA	Myanmar bot	5/13/21	8/11/21	8/11/21	90
2021-008-FB-FBR	COVID lockdowns in Brazil	5/26/21	8/19/21	N/A (Upheld Meta's decision to leave post up)	85
2021-009-FB-UA	Shared Al Jazeera post	6/23/21	9/14/21	N/A (Meta already restored content)	83
2021-010-FB-UA	Colombia protests	6/29/21	9/27/21	9/27/21	90
2021-011-FB-UA	South Africa slurs	6/30/21	9/28/21	N/A (Upheld Meta's decision to remove post)	90

Questions for Meta

To assist with making its decisions, the Oversight Board sends questions to Meta. Of the 103 questions sent by the Oversight Board to Meta about decisions published in this quarter, Meta answered 89 questions, partially answered nine questions and did not answer five. More information on the five questions Meta did not answer is provided below.

In the [Ocalan's isolation case](#), the Board asked Meta if it could determine how many pieces of content were wrongly removed while a policy exception related to the case was not being enforced. The Board also asked how much content mentioning Abdullah Ocalan Meta had removed in the last five years. In both cases, Meta said it was not technically feasible to provide the information. The Board also asked Meta why its moderators concluded that the content in this case constituted a call to action to support Ocalan and/or the PKK. In response, Meta noted that it does not require its at-scale content reviewers to document their reasoning for each content decision.

In the [South Africa slurs case](#), the Board asked Meta to provide metrics on the prevalence on the platform of three terms mentioned in the post. In response, Meta said it was not technically feasible to provide this information. Responding to another question, Meta did not provide information on whether its public policy team based in South Africa includes people from South Africa.

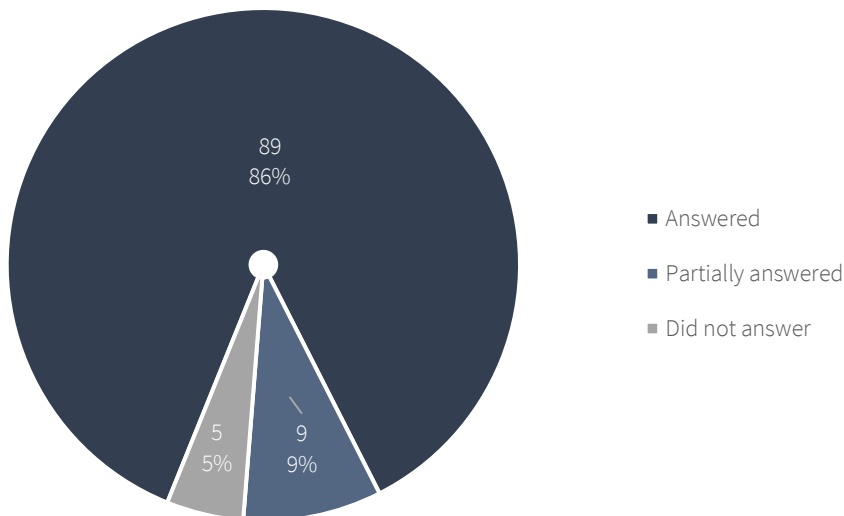


Oversight Board questions answered by Meta

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>
2021-006-IG-UA	Ocalan's isolation	15	1	3	19
2021-007-FB-UA	Myanmar bot	24	2	0	26
2021-008-FB-FBR	COVID lockdowns in Brazil	8	0	0	8
2021-009-FB-UA	Shared Al Jazeera post	23	4	0	27
2021-010-FB-UA	Colombia protests	8	0	0	8
2021-011-FB-UA	South Africa slurs	11	2	2	15
Total		89	9	5	103

Oversight Board questions answered by Meta





Public Comments

The Oversight Board conducts a public comment process to assist it in its decision making. In this period, the Oversight Board received 106 comments, 74 of which were published.

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 unattributed¹⁰</i>
2021-006-IG-UA	Ocalan's isolation	6	2	6	14	1
2021-007-FB-UA	Myanmar bot	10	1	1	12	4
2021-008-FB-FBR	COVID lockdowns in Brazil	25	1	4	30	9
2021-009-FB-UA	Shared Al Jazeera post	18	3	5	26	6
2021-010-FB-UA	Colombia protests	10	4	4	18	8
2021-011-FB-UA	South Africa slurs	5	1	1	6 ¹¹	2
Total		74	12	20¹²	106	30

The majority of public comments (61%) came from individuals, while a minority came from organizations (39%).

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>
2021-006-IG-UA	Ocalan's isolation	7	7	14
2021-007-FB-UA	Myanmar bot	7	5	12

¹⁰ Unattributed comments are published comments with the author's name redacted by request.

¹¹ One public comment received was not published as it both violated terms and it did not give consent to be published, for this reason it appears in both columns, but has only been counted once.

¹² One public comment received was not published as it both violated terms and it did not give consent to be published, for this reason it has only been counted once.



2021-008-FB-FBR	COVID lockdowns in Brazil	22	8	30
2021-009-FB-UA	Shared Al Jazeera post	12	14	26
2021-010-FB-UA	Colombia protests	12	6	18
2021-011-FB-UA	South Africa slurs	5	1	6
Total		65	41	106

Nearly half (48%) of the public comments received for decisions published in this quarter came from the US and Canada, followed by Latin America & Caribbean (16%) and Asia Pacific & Oceania (13%).

Public comments received by region

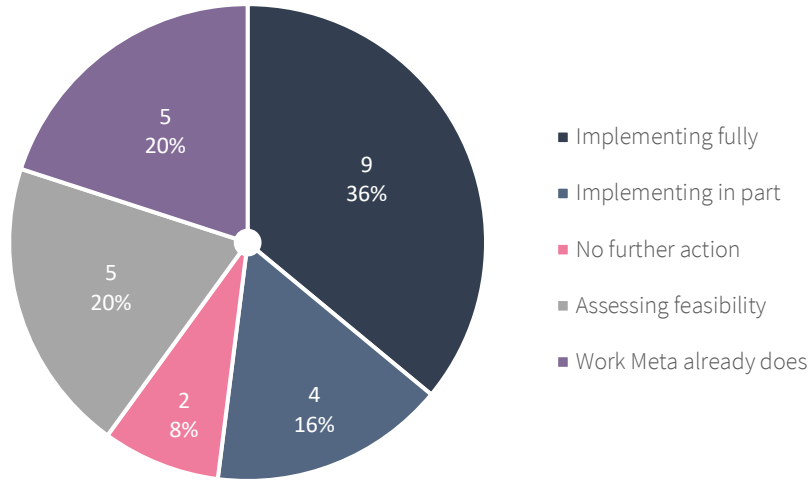
Number of comments

<i>Region</i>	<i>2021-006-IG-UA</i>	<i>2021-007-FB-UA</i>	<i>2021-008-FB-FBR</i>	<i>2021-009-FB-UA</i>	<i>2021-010-FB-UA</i>	<i>2021-011-FB-UA</i>	<i>Total</i>
United States & Canada	7	7	17	15	4	1	51
Europe	5	0	0	7	1	0	13
Asia Pacific & Oceania	0	5	3	0	5	1	14
Latin America & Caribbean	0	0	9	1	7	0	17
Sub-Saharan Africa	0	0	0	0	0	3	3
Middle East & North Africa	2	0	0	3	1	1	7
Central & South Asia	0	0	1	0	0	0	1
Total	14	12	30	26	18	6	106

Recommendations

In addition to providing decisions on appealed content, the Oversight Board also provided 25 policy recommendations to Meta, which the company responded to publicly within 30 days.

Of these recommendations, Meta said it was “implementing fully” or “implementing in part” more than half (13) of our recommendations. Meta said it was “assessing feasibility” on five recommendations, and claimed a further five recommendations represented “work Meta already does.” The company said it would take “no further action” on two recommendations.



In this quarter, the Board made recommendations on content policy (clarification or changes to rules), enforcement (clarification or changes to how rules are applied), and transparency (on disclosure of information to the public). Below, the Board lists all policy recommendations listed in case decisions during Q3. The Board has reproduced Meta’s initial commitments made within the 30-day response window. Meta continues to update their on-going progress to implement the recommendations. The Board will continue to monitor Meta’s implementation efforts and include information about this monitoring in future transparency reporting.

Oversight Board recommendations to Meta

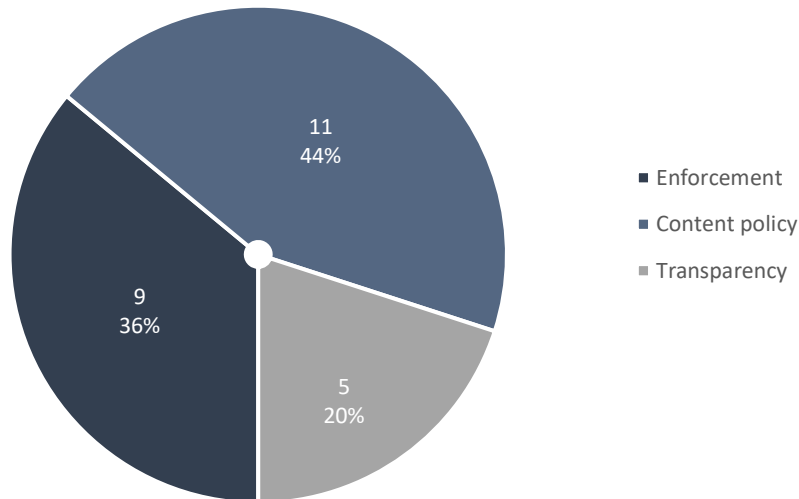
Number of recommendations

Case ID	Name	Content policy	Enforcement	Transparency	Total
2021-006-IG-UA	Ocalan’s isolation	5	4	3	12
2021-007-FB-UA	Myanmar bot	0	1	0	1
2021-008-FB-FBR	COVID lockdowns in Brazil	2	1	0	3
2021-009-FB-UA	Shared Al Jazeera post	2	0	2	4
2021-010-FB-UA	Colombia protests	2	2	0	4
2021-011-FB-UA	South Africa slurs	0	1	0	1
Total		11	9	5	25



Oversight Board recommendations to Meta

The recommendations are divided into three groups: content policy, enforcement, and transparency.



Content policy recommendations

Recommendations to clarify rules, for rules to be modified, or for new rules to be created.

[Ocalan's isolation \(2021-006-IG-UA\)](#)

- 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. **[Implementing fully]**
- 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. **[Implementing fully]**



- 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. **[Implementing in part]**
- 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. **[Work Meta already does]**
- 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. **[Implementing fully]**

[COVID lockdowns in Brazil \(2021-008-FB-FBR\)](#)

- 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. **[Work Meta already does]**
- 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. **[Implementing fully]**

[Shared Al Jazeera post \(2021-009-FB-UA\)](#)

- 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. **[Assessing feasibility]**
- Ensure swift translation of updates to the Community Standards into all available languages. **[Assessing feasibility]**

[Colombia protests \(2021-010-FB-UA\)](#)

- 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. **[Assessing feasibility]**
- Link the short explanation of the newsworthiness allowance provided in the introduction to the Community Standards to the more detailed explanation in the Facebook’s Transparency Center of how this policy applies. The company should supplement this explanation with illustrative



examples from a range of contexts, including reporting on large scale protests. **[Implementing fully]**

Enforcement recommendations

Recommendations to change content moderation enforcement processes, including about the use of automated or human review, notification of enforcement action to users, and access to appeals.

[Ocalan's isolation \(2021-006-IG-UA\)](#)

- 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. **[Implementing fully]**
- Evaluate automated moderation processes for enforcement of the Dangerous Individuals and Organizations policy and where necessary 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. **[No further action]**
- 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. **[Work Meta already does]**
- 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). **[Implementing fully]**

[Myanmar bot \(2021-007-FB-UA\)](#)

- 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. **[No further action]**

[COVID lockdowns in Brazil \(2021-008-FB-FBR\)](#)

- 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. **[Work Meta already does]**

[Colombia protests \(2021-010-FB-UA\)](#)

- 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. **[Work Meta already does]**
- 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. **[Assessing feasibility]**

[South Africa slurs \(2021-011-FB-UA\)](#)



- To ensure procedural fairness for users, Facebook should: 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. **[Implementing in part]**

Transparency recommendations

Recommendations to increase the transparency of Meta’s content moderation, including through its transparency reports.

[Ocalan’s isolation \(2021-006-IG-UA\)](#)

- 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. **[Implementing in part]**
- Include information on the number of requests Facebook receives for content removals from governments that are based on Community Standards violations (as opposed to violations of national law), and the outcome of those requests. **[Implementing fully]**
- 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. **[Assessing feasibility]**

[Shared Al Jazeera post \(2021-009-FB-UA\)](#)

- 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. **[Implementing fully]**
- 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. **[Implementing in part]**



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

1. Even though Meta identified this case as an enforcement error, it was assigned to panel by the Oversight Board as case 2021-012-FB-UA (“Wampum Belt”).
2. Even though Meta identified this case as an enforcement error, it was assigned to panel by the Oversight Board as 2021-014-FB-UA (“Alleged Crimes in Raya Kobo”).
3. Even though Meta identified this case as an enforcement error, it was assigned to panel by the Oversight Board as case 2021-015-FB-UA (“Asking for Adderall”).
4. Even though Meta identified this case as an enforcement error, it was assigned to panel by the Oversight Board as case 2021-017-FB-UA (“Afghan Journalist”).
5. The content is a post recounting the user’s travel experience in Afghanistan and giving his opinions on the recent takeover of the country by the Taliban. It expresses cautious optimism about the Taliban taking power, pointing to things like granting amnesty to Afghan government officials, and claims of respecting women’s rights and building a tolerant, multi-cultural nation. It also says that overall, aside from the unfortunate events at the Kabul airport, developments have been better than imagined, but more time is needed for powers like China, Russia and Iran to observe what is going to happen in Afghanistan. The content was erroneously taken down under the Community Standard on Dangerous Individuals and Organizations. 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 content does not rise to the level of praise, substantive support, or representation under the Community Standard.¹³
6. The post, written in Urdu, is sharing a news update concerning the ongoing fight in Panjshir and the Taliban’s control of the valley. The content was erroneously taken down under the Community Standard on Dangerous Individuals and Organizations. 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 content does not rise to the level of praise, substantive support, or representation under the Community Standard.
7. The content in question contains screenshots purportedly from a Belarusian state television news channel, where photos of opposition figures and celebrities can be seen next to nooses superimposed in television channel graphics. In the text of the post, the user does not express any stance on the screenshots. 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

¹³ In each of the short summaries in this annex, the sentence which includes the words “it seems that” represents an assessment by the Oversight Board Administration staff of why Meta may have reversed its original decision on this content, which may differ from the company’s actual reason for reversing its decision.



seems that based on context including the user statement, the content was intended to raise awareness and does not rise to the level of either a direct or an implicit threat.

8. A Mauritanian user wrote a post condemning the government and encouraging people to share videos and images of torture taking place in the country. The post said that people who remained silent were despicable and deserved the consequences of their inaction. The content was erroneously taken down under the Community Standard on Hate Speech. 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 while the word "despicable" on its own contains meanings of inferiority, in this post the word was used self-referentially to target people of the same nationality as the user, and based on context, use of "despicable" here does not appear to constitute a statement of inferiority under the Community Standard.
9. The content discusses issues around sexual assaults in Romania by referring to a news article and a video about a woman being taken advantage of in a club. The video in question is not shown in the content. The content was erroneously taken down under the Community Standard on Bullying and Harassment. 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 nothing in the content is related to bullying or harassment.
10. The content is a post with religious text about Islam. It seems to describe the conditions upon which Ayat Al Kursi, a famous Quranic verse, has been revealed and the benefits of reciting it. It also contains quotes from a hadith/Islamic literature. The content was erroneously taken down under the Community Standard on Hate Speech. 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 nothing in the content could constitute hate speech.
11. The content is about the remains of over 200 Indigenous children, as well as hundreds upon hundreds of other unmarked graves, being found at Canada's former residential schools across the country. The content describes how Canadians are demanding an apology from the Pope as the Roman Catholic Church operated about 70 percent of the schools, suggesting that in the apology the Pope should refer to himself and the Church as "We, the Catholic church, we priests and nuns... We criminals." The content was erroneously taken down under the Community Standard on Hate Speech. 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 while the content proposes that the Pope refer to himself and certain other Catholics as "criminals," this falls under an exception in the Community Standard on Hate Speech for subsets (in this case, Catholic priests and nuns) described as having committed violent crimes.
12. The content is a screenshot of what could be a chat on a dating app, or was made to look like one. One person says they support the Thai monarchy and the other person says, "let me block you" and uses "คนโง่", an offensive word in Thai meaning "idiot," with the literal translation in English being "buffalo." The content was erroneously taken down under the Community Standard on Bullying and Harassment. 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 while



the word is offensive, Meta possibly reversed its decision because the word is used in the context of opposing the monarchy in Thailand.

13. The content contains a screenshot shows a post from years ago by another user who criticized someone who shot an activist in Sudan. That user said "your day will come" and referred to the killer as "you despicable." Recently it appears that the killer was sentenced and the user in this appeal shared this post to celebrate the sentencing, quoting the other user's words including "you despicable." The content was erroneously taken down under the Community Standard on Bullying and Harassment. 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 content could fall under an exception under the Community Standard on Bullying and Harassment for negative character or ability claims, and expressions of contempt/disgust, in the context of criminal allegations.