

Oversight Board
Q1 2023
transparency report

Transparency Report for first quarter of 2023

This transparency report for the first quarter of 2023 (January 1 – March 31, 2023) sets out key statistics on cases selected by the Board, as well as the decisions and recommendations we made in this quarter.

This report also includes new data on the impact of recommendations from our “Iran protest slogan” and “breast cancer symptoms and nudity” cases.

Other highlights from this report include:

- We published decisions on four cases in Q1 2023: “Iran protest slogan,” “gender identity and nudity” (which covered two cases), and “Sri Lanka pharmaceuticals.”
- We received nearly 300 public comments to the Board ahead of the deliberations on these cases.
- We announced four new cases for consideration, as well as a policy advisory opinion on the Arabic term “Shaheed.”
- Users submitted more than 140,000 cases to the Board in Q1 2023.
- Following our concerns about Meta’s opaque penalty system and user concerns about being placed in “Facebook jail,” the company changed its ‘strikes’ system to make it fairer and more transparent.

For the convenience of readers, a glossary of terms used in this report may be found at the end of this report.

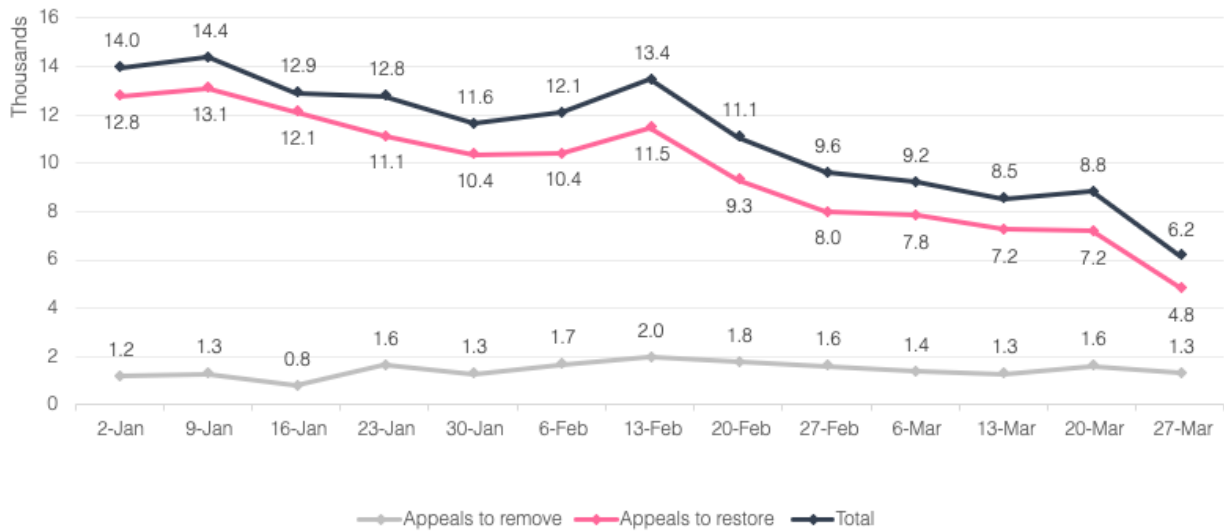
Q1 2023 Submitted User Cases

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

In Q1 2023, **142,759 cases** were submitted by users, a reduction of 26% on Q4 2022. In total, users submitted over 2.5 million cases to the Board since the Board started accepting cases in October 2020 to March 2023.

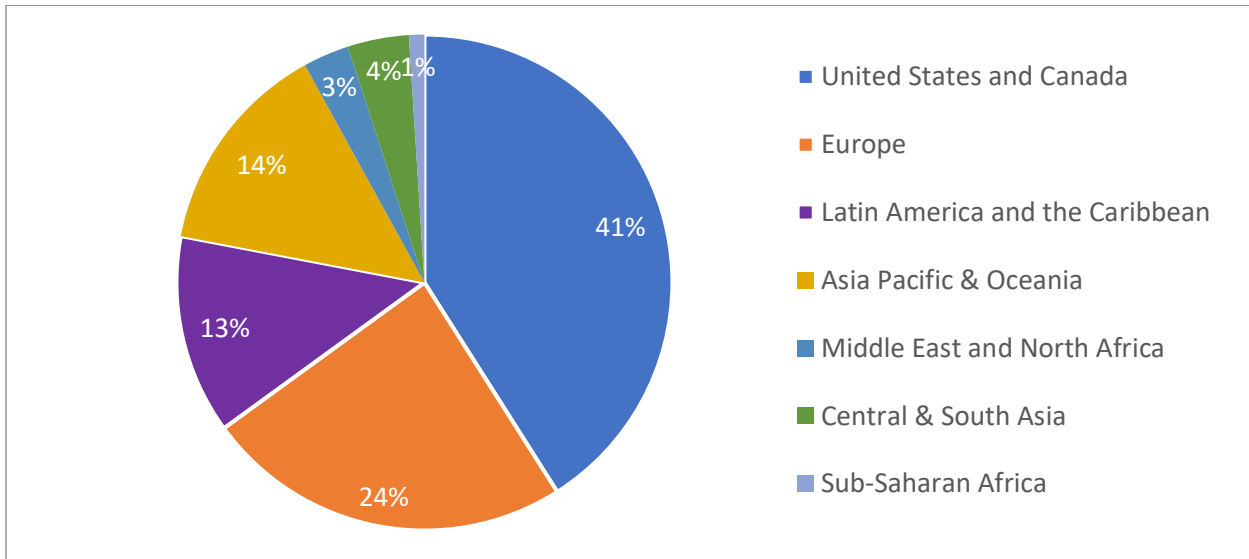
Estimated number of cases submitted to Oversight Board by week

Number of cases



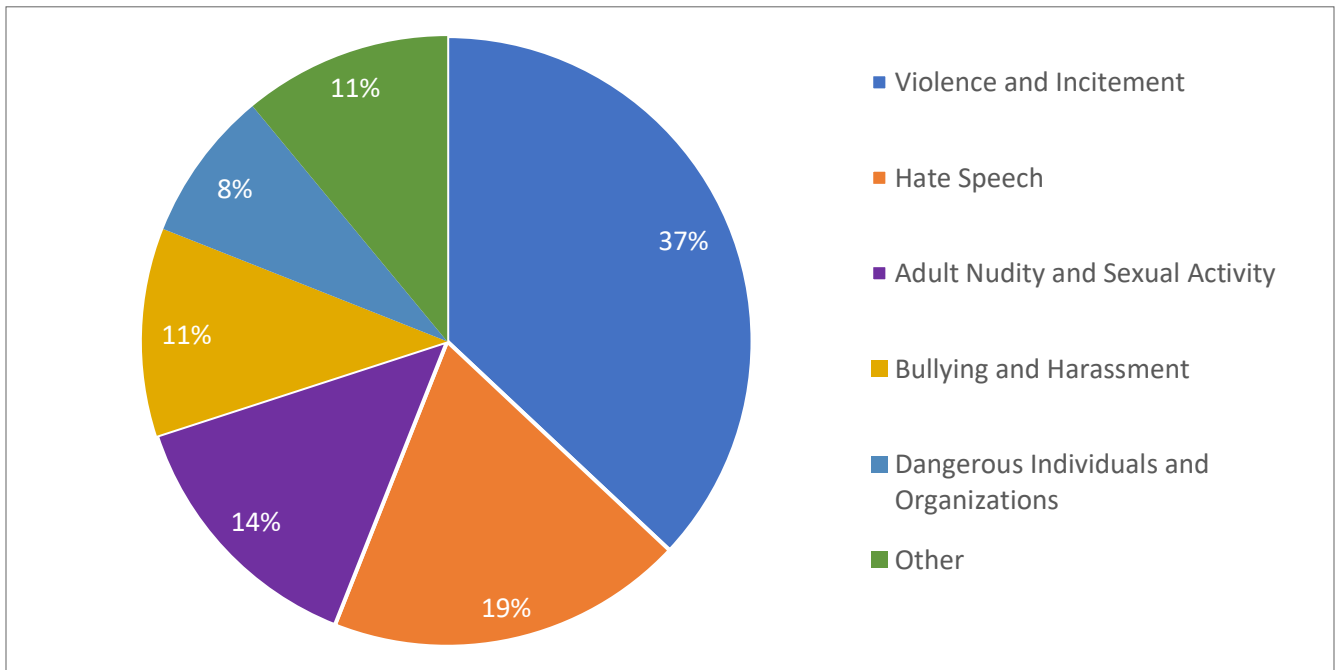
The week commencing January 9 saw the highest number of cases sent to the Board in Q1 2023, with around 14,400 cases. While the majority (87%) of cases submitted to the Board were appeals to restore content, 13% of cases were appeals to remove other users’ content.

Estimated cases submitted by user-selected region (Percent)



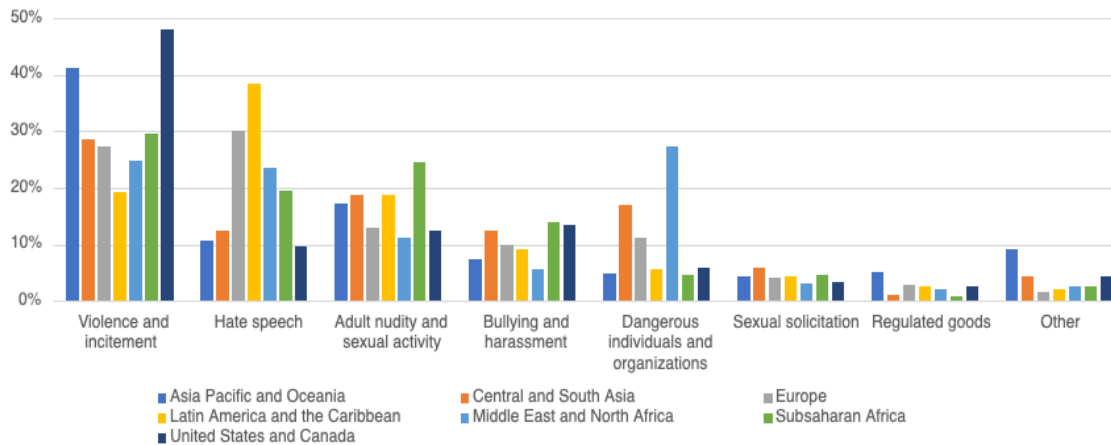
Around 41% of cases submitted to the Board in Q1 2023 came from the US & Canada. While still representing a significant share, this was the lowest share of appeals from this region in the history of the Board’s quarterly reporting (the previous lowest share for US & Canada was 44.8% in Q1 2021). Europe made up 24% of appeals, while 14% of appeals came from Asia Pacific and Oceania and 13% from Latin America and the Caribbean. Central and South Asia received 4% of appeals, the Middle East and North Africa received 3% and Sub-Saharan Africa received 1%.

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



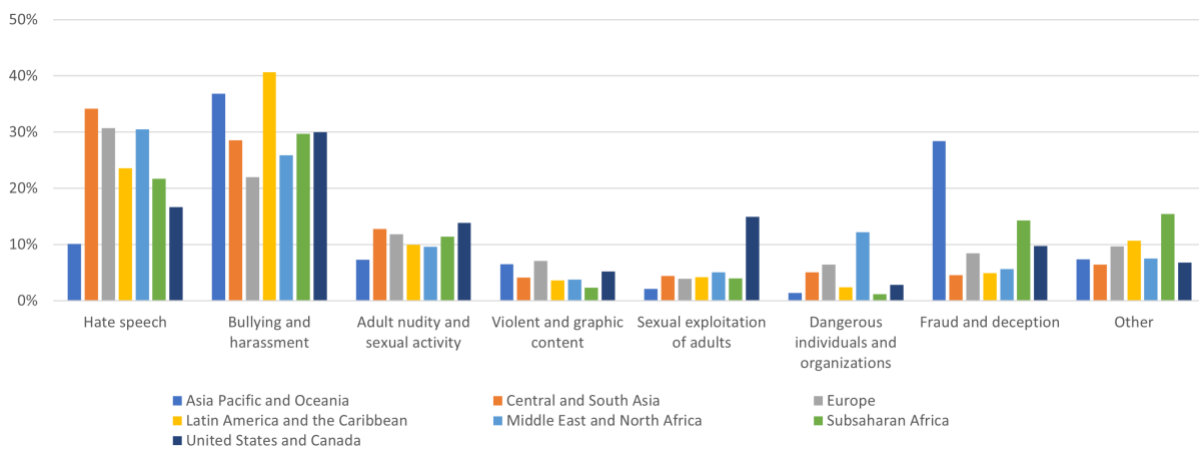
Users primarily submitted appeals to restore content which Meta removed for violating its policies on Violence and Incitement (37%), Hate Speech (19%), Adult Nudity and Sexual Activity (14%), and Bullying and Harassment (11%).

Proportion of all-user submitted appeals to restore content by Community Standard in each region (Percent)



In this quarter, Violence and Incitement was the Community Standard most frequently cited in user appeals to restore content in Asia Pacific and Oceania, Central and South Asia, Sub-Saharan Africa, and the United States and Canada. Hate Speech was the most frequently cited Community Standard for user appeals to restore content in Europe and Latin America and the Caribbean, while Dangerous Individuals and Organizations was the most cited in the Middle East and North Africa.

Proportion of all user-submitted appeals to *remove* content by Community Standard in each region (Percent)



In this quarter, Bullying and Harassment was the Community Standard most cited in user appeals to remove content in Asia Pacific and Oceania, Latin America and the Caribbean, Sub-Saharan Africa, and

the United States and Canada. Hate Speech was the Community Standard most cited in user appeals to remove content in Central and South Asia, Europe, and the Middle East and North Africa.

Percent	
Facebook	81%
Instagram	19%

The vast majority of cases submitted to the Board by users (81%) concerned content shared on Facebook, while 19% of cases concerning content shared on Instagram. This represents the highest-ever share of cases about content on Instagram, which was around 19 times higher than the 1% share it received in Q1 2022.

Q1 2023 Longlisted User Cases

In this period, **47 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 38 pieces of content on Facebook and 9 pieces of content on Instagram. 16 were cases to restore content and 31 were cases to remove content.

Longlisted cases by region and affected country¹

Number of cases

Asia Pacific and Oceania	7
Australia	3
Indonesia	1
New Zealand	1
Philippines	2
Europe	17
Bulgaria	1
Denmark	1
France	1
Germany	1
Hungary	1
Norway	1
Romania	1
Spain	1
Sweden	1
Turkey	4
Ukraine	1
United Kingdom	3
United States & Canada	11
United States	11
Sub-Saharan Africa	3
Ethiopia	1
Ghana	1
Guinea-Bissau	1
Central & South Asia	9
Afghanistan	1
Bangladesh	2
Iran	1
Myanmar	3
Pakistan	1

¹ 'Countries Affected' is a user-selected field in the Board's appeal process and users can select multiple countries. For this reason, 47 user-submitted cases were longlisted but a tally of 65 countries 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.

Thailand	1
Latin America and Caribbean	7
Argentina	1
Brazil	4
Guatemala	1
Peru	1
Middle East and North Africa	11
Egypt	1
Iran	1
Iraq	1
Israel	6
Palestine	2
Total	65

Of the 47 cases longlisted by the Board in this quarter, 31 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, and are included under “undefined.” Of those cases where the content had been deemed by Meta to violate its rules, the most common Community Standards were Hate Speech (six cases), followed by Dangerous Individuals and Organizations (five cases), Coordinating Harm and Publicizing Crime (three cases) and Violence and Incitement and Violence and Graphic Content (one case each).

Longlisted cases by Community Standard

Number of cases	
Coordinating Harm and Publicizing Crime	3
Dangerous Individuals and Organizations	5
Hate Speech	6
Violence and Incitement	1
Violent and Graphic Content	1
Undefined ²	31
Dangerous Individuals and Organizations	8

Q1 2023 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 four occasions, **shortlisting 27 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-

² 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.

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 27 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 17 out of the 27 cases shortlisted by the Board. In eight of these cases, Meta found it had incorrectly removed content and restored it. In nine cases, it found it had wrongly left up content and removed it.

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 63% of cases the Board shortlisted in Q1 2023. This is similar to Meta’s error rate in 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.

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>
Dangerous Individuals and Organizations	1	1	2
Hate Speech	4	1	5
Violence and Incitement	1	0	1
Undefined	8	1	9
Total	14	3	17

Q1 2023 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 one case, which was related to the Adult Nudity and Sexual Activity Community Standard. The remaining seven cases concerned content that the company had left up on Facebook.

Meta referrals submitted³

<i>Case ID</i>	<i>Name</i>	<i>Meta's decision</i>	<i>Platform</i>	<i>Language</i>	<i>Community Standard</i>	<i>Countries⁴</i>
2023-004-FB-MR	Armenian prisoners of war video	Leave up	FB	English & Turkish	N/A	Armenian/ Azerbaijan
N/A	N/A	Leave up	FB	English	N/A	US
N/A	N/A	Take down	FB	Marathi	Adult Nudity and Sexual Activity	India
N/A	N/A	Leave up	FB	Portuguese	N/A	Portugal
N/A	N/A	Leave up	IG	English	N/A	UK
2023-003-FB-MR	Cambodian prime minister	Leave up	FB	Khmer	N/A	Cambodia
N/A	N/A	Leave up	IG	Portuguese	N/A	Brazil
2023-010-IG-MR	Promoting Ketamine for non-FDA-approved treatments	Leave up	IG	English	N/A	US

Policy advisory opinion on "shaheed" and designated dangerous individuals (PAO 2023-01)

On March 9, the Oversight Board announced it had accepted a request from Meta for a policy advisory opinion on moderating the Arabic term "shaheed" when referring to individuals it classifies as "dangerous", including terrorists (PAO 2023-01). "Shaheed" has multiple meanings, but is often translated as "martyr," and accounts for more content removals under the Community Standards than any other single word or phrase on Meta's platforms. The company acknowledges that its current approach may

³ Cases which are not selected for assignment do not have a Case ID, Name, etc.

⁴ 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.

result in significant over-enforcement, particularly in Arabic-speaking countries, and has explored alternatives. However, it points to the difficulties and tensions in moderating use of the term at scale. Meta asked the Board to consider the following policy options: continue its current approach and remove content that uses "shaheed" to refer to an individual designated as dangerous under the Dangerous Individuals and Organisations policy; allow content that refers to a designated individual as "shaheed" when certain conditions are met; remove content that uses "shaheed" to refer to an individual designated as dangerous under Meta's Dangerous Individuals and Organisations policy only where there is additional praise, representation or support, or where there is a signal of violence. The Board's recommendations and policy advisory opinion may not be limited to these options. The Board invited public comments on the case between 9 March and 10 April 2023.

Q1 2023 Selected Cases

The Case Selection Committee selects cases for review, which are then announced publicly on the Board's website. In Q1, the Committee selected four cases.

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>
2023-001-FB-UA	Brazilian general's speech	9 March	FB	User appeal (to remove content)	N/A	Brazil
2023-002-IG-UA	Violence against women ⁵	9 March	IG	User appeal (to restore content)	Hate Speech	Sweden
2023-003-FB-MR	Cambodian prime minister	16 March	FB	User appeal (to remove content) and Meta referral	Violence and Incitement	Cambodia
2023-004-FB-MR	Armenian prisoners of war video	22 March	FB	Meta referral	Coordinating Harm and Promoting Crime	Armenia/ Azerbaijan
PAO 2023-01	"Shaheed" and designated dangerous individuals policy advisory opinion	9 March	N/A	Meta request	Dangerous individuals and organizations	Global

⁵ The Board originally announced the "violence against women testimony" case on March 9. However, as the Board later (April 27) added a second case (2023-005-IG-UA) to also be considered in the final decision text, the title was changed to "violence against women" to more accurately reflect both posts.

Q1 2023 Published Case Decisions

After cases are selected, the Board assigns them to a panel of Board Members. Members of the panel include at least one member from the region implicated in the content and a mix of gender representation. The panel examines whether Meta’s decision is consistent with the company’s content policies, values, and 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 issued decisions covering four cases. It upheld Meta’s original decision in one case and overturned the company’s original decision in three.

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>
2022-013-FB-UA	Iran protest slogan	FB	User appeal (to restore content)	Farsi	Violence and Incitement	Iran	Overturned
2022-009-IG-UA	Gender identity and nudity	IG	User appeal (to restore content)	English	Sexual Solicitation	US	Overturned
2022-010-IG-UA	Gender identity and nudity ⁷	IG	User appeal (to restore content)	English	Sexual Solicitation	US	Overturned
2022-014-FB-MR	Sri Lanka pharmaceuticals	FB	Meta request	English	Restricted Goods and Services	Sri Lanka	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.

⁷ In the “Gender identity and nudity” decisions, the Board considered two cases together for the first time, cases 2022-009-IG-UA and 2022-010-IG-UA.

Decision timelines

According to our Bylaws, the Board aims to make its decision on a case 90 days from the date it announces a new case on its website.

For the four decisions the Oversight Board published in this quarter, the average time from announcement of the case to publication of the Board’s decision has been 140 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</i>
2022-009-IG-UA	Gender identity and nudity	07/26/22	01/17/23	175
2022-010-IG-UA	Gender identity and nudity	07/26/22	01/17/23	175
2022-013-FB-UA	Iran protest slogan	10/04/22	01/09/23	97
2022-014-FB-MR	Sri Lanka pharmaceuticals	11/17/22	03/09/23	112

Questions for Meta

To assist with making its decisions, the Oversight Board sends questions to Meta. Of the 56 questions sent by the Oversight Board to Meta about decisions published in this quarter, Meta answered 53 questions (95%), it partially answered three questions (5%).

In the “Iran protest slogan” case, the partial responses related to: data comparing auto-closure of appeals for content in Farsi and English languages; the prevalence of several variations of the “death to Khamenei” slogan on Meta’s platforms; and the accuracy rates on the enforcement of the Violence and Incitement policy in Farsi.

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>
2022-009-IG-UA & 2022-010-IG-UA	Gender identity and nudity	18	0	0	18
2022-013-FB-UA	Iran protest slogan	26	3	0	29
2022-014-FB-MR	Sri Lanka pharmaceuticals	9	0	0	9
Total		53	3	0	56

⁸ All decisions published in this quarter exceeded the 90-day deadline due to delays in the timeline arising from Oversight Board staff taking leave during the December holiday period. Additionally, the “Sri Lanka pharmaceuticals” case was impacted by its translation turnaround. There were also challenges scheduling Board Member deliberations, alongside other significant operational challenges, impacting both the “Gender identity and nudity” cases.

Public comments

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

For the four cases reviewed by the Board in Q1 2023, the Board received 295 comments, of which 189 were published. 46 comments were submitted without consent to publish. 266 public comments (90%) came from individuals, while 29 comments (10%) came from organizations.

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>
2022-009-IG-UA & 2022-010-IG-UA	Gender identity and nudity	84	16	30	130	39
2022-013-FB-UA	Iran protest slogan	103	29	30	162	83
2022-014-FB-MR	Sri Lanka pharmaceuticals	2	1	0	3	2
Total		189	46	60	295	124

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>
2022-009-IG-UA & 2022-010-IG-UA	Gender identity and nudity	122	8	130
2022-013-FB-UA	Iran protest slogan	142	20	162
2022-014-FB-MR	Sri Lanka pharmaceuticals	2	1	3
Total		266	29	295

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

Public comments received by region

Number of comments

<i>Region</i>	2022-009-IG-UA & 2022-010-IG-UA	2022-013-FB-UA	2022-014-FB-MR	<i>Total</i>
United States & Canada	97	65	2	164
Europe	19	42	0	61
Asia Pacific & Oceania	10	13	0	23
Latin America & Caribbean	1	0	1	2
Sub-Saharan Africa	1	0	0	1
Middle East and North Africa	1	36	0	37
Central and South Asia	1	6	0	7
Total	130	162	3	295

56% of the public comments received for decisions published in this quarter came from the US and Canada. 21% came from Europe, 13% came from the Middle East and North Africa 8% came from Asia Pacific and Oceania, 2% came from Central and South Asia, and less than 1% came from Latin America and the Caribbean and Sub-Saharan Africa.

Recommendations

In addition to providing decisions on appealed content, the Oversight Board also made 14 policy recommendations to Meta.

Of these 14 recommendations, four related to content policy (clarification of or changes to rules), five related to enforcement (clarification of or changes to how rules are applied), and five related to transparency (on disclosure of information to the public).

The Board's recommendations seek to improve Meta's approach to content moderation, protect users, and increase transparency. Recommendations made in Q1 2023 include:

- In the "Gender identity and nudity" cases, the Board found that Meta's policies on adult nudity result in greater barriers to expression for women, trans and gender non-binary people on its platforms. Among other things, it recommended that Meta define clear, objective, rights-

respecting criteria to govern its Adult Nudity and Sexual Activity Community Standard, so that all people are treated in a manner consistent with international human rights standards, without discrimination on the basis of sex or gender. The Board said that Meta should first conduct a comprehensive human rights impact assessment on such a change, engaging diverse stakeholders and create a plan to address any harms identified.

- In the “Iran protest slogan” case, the Board raised concerns that in the context of Iranian protests, Meta must do more to respect freedom of expression. Among other things, the Board recommended that, pending changes to the Violence and Incitement Community Standard, Meta issue guidance to its reviewers that, in the context of protests in Iran, “marg bar Khamenei” statements do not violate the policy. The Board compared a dataset of posts using the phrase “marg bar Khamenei” from the same set of public pages, groups and Instagram accounts during the same time period before the case was decided and after Meta implemented the recommendation. We found that on Instagram, there was a 28.75% increase of posts using the phrase from the same pages, groups, and accounts within the same time period following implementation. We found this change to be statistically significant, meaning it was highly likely to be due to Meta implementing the recommendation. The Board’s full analysis can be found as an annex to this report.
- In the “Sri Lanka pharmaceuticals” case, the Board raised concerns about the “spirit of the policy” allowance being “secret and discretionary.” Among other things, it recommended that Meta publish information on the allowance in its Transparency Center, including the criteria Meta uses to decide whether to scale the allowance. It also recommended Meta explain in the Community Standards that allowances may be made when a policy's rationale, and Meta's values, demand a different outcome than a strict reading of the rules.

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>
2022-009-IG-UA & 2022-010-IG-UA	Gender identity and nudity	2	1	0	3
2022-013-FB-UA	Iran protest slogan	1	3	3	7
2022-014-FB-MR	Sri Lanka pharmaceuticals	1	1	2	4
Total		4	5	5	14

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

For the 14 recommendations made by the Board in Q1 2023, Meta's initial responses were 79% "comprehensive" (11 recommendations), and 21% "somewhat comprehensive" (3 recommendations). None were "not comprehensive."

Board's analysis of Meta's implementation of Q1 recommendations

Of the 14 recommendations the Board made in Q1 2023 that Meta had responded to at the time of writing:

- One was implemented by Meta, and the Board is awaiting evidence of implementation. This was "Iran protest slogan" recommendation no. 3.
- Ten 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.
- None had been wholly or partially implemented, as verified by publicly available information.
- One was declined by Meta following a feasibility assessment. This was "Iran protest slogan" recommendation no. 7.
- One was declined by Meta. This was "Sri Lanka pharmaceuticals" recommendation no. 2.
- One was reframed by Meta. This was "Iran protest slogan" recommendation no. 2.

Board analysis of other recommendations

We also received further information about the recommendation from our "breast cancer symptoms and nudity" decision urging Meta to, "improve the automated detection of images with text-overlay to ensure that posts raising awareness of breast cancer are not wrongly flagged for review."

Meta previously shared that improvements to its text-overlay detection, made following our recommendation, led to 2,500 pieces of content being sent to human review over 30 days in February and March 2023, when it would previously have been automatically removed.

In addition to these changes, Meta has now told us that it tested and deployed a new health content classifier to further enhance Instagram's techniques for identifying breast cancer context content deployed those in July 2021. Those enhancements have been in place since, and in the 28 days from March 21 to April 18, 2023, they contributed to an additional 1,000 pieces of content being sent for human review that would have previously been removed.

This new data further demonstrates our impact on how Meta moderates content raising awareness of breast cancer and shows the different ways in which Meta is responding to this important recommendation.

Meta has committed to implement or implemented the majority of the Board's recommendations. The Board has assessed that 29 out of the 214 recommendations it has made since January 2021 have been implemented fully, as demonstrated through published information. The Board assessed a further 15 recommendations as partially implemented, while for 81 recommendations, Meta had reported

progress towards implementation and we will continue to monitor implementation. On 35 recommendations, Meta has declined to implement the recommendation. For 31 recommendations, the company has reported implementation, or said it already does what the Board recommends, but has not published information to demonstrate this.

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>	29
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>	15
Progress reported: <i>Meta committed to implementation but has not yet completed all necessary actions.</i>	81
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>	31
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>	11
Recommendation omitted, declined, or reframed: <i>Meta will take no further action on the recommendation.</i>	24
Awaiting first response	23
Total Number of Recommendations	214

Alongside this report we are also publishing a full breakdown of our assessment of Meta's implementation of all 214 recommendations.

Annex I – list of shortlisted cases where Meta reversed its original decision in Q1 2023.

Below is a list of cases where Meta identified that its original decision on content was incorrect. As [announced on February 14, 2023](#), some of these cases will be published as summary decisions and they have been noted as such in the table.

#	Content	Platform	Meta’s Original Decision	Meta’s Final Decision after Shortlisting by the Board	Country	Policy
1	Assigned to panel as “Brazilian general’s speech” (2023-001-FB-UA)	FB	Keep up	Take down	Brazil	Violence & Incitement
2	Assigned to panel as “violence against women” (2023-002-IG-UA)	IG	Take down	Keep up	Sweden	Hate Speech
3	The Board intends to publish this case as a summary decision. The content featured a photo of a woman. The caption text above the photo wrote "For Sale! Used dumptruck" and went on to describe how the "woman" ("truck") required heavy "paint" to hide bumps and bruises, etc, and that "it" smelled a little funny and did not get "washed often."	FB	Keep up	Take down	United States	Hate Speech
4	The user reshared a post from a civil rights activist which contained screenshots of hate speech from Twitter and which had a caption that condemned the speech. When the user appealed to the Board, the original post remained live while the	FB	Take down	Keep up	United States	Hate Speech

	user's resharing of the post was removed.					
5	The Board intends to publish this case as a summary decision. This case addressed the removal of former Peruvian president Pedro Castillo from office. The post stated, "we're going to hang you [Castillo] like Mussolini, but don't be scared (we're not philosophers like you, it's metaphorical.)"	FB	Take down	Keep up	Peru	Violence & Incitement
6	The content contained a drawn image of a pair of legs hanging from the ceiling with a toppled stool on the ground. The caption accompanying the image was "my dream POV" (point of view).	FB	Keep up	Take down	Philippines	Suicide & Self-Injury
7	The content contained an image of Jewish children and their father, all of them wearing traditional Jewish skullcaps and having long curly sideburns. A boy at the front of the image had misshapen teeth. The caption above the image read, "Ain't no way in hell these are God's chosen people."	FB	Keep up	Take down	United States	Hate Speech
8	The Board intends to publish this case as a summary decision. The content contained a poem about Pan-African anticolonial revolutionary Amilcar Cabral.	FB	Take down	Keep up	Guinea-Bissau	Dangerous Individuals & Organizations

9	Assigned to panel as “image of gender-based violence” (2023-006-FB-UA)	FB	Keep up	Take down	Iraq	Bullying & Harassment
10	The Board intends to publish this case as a summary decision. The content contained a video, and in the first part a celebrity said, "I like Hitler," and "he didn't kill 6 million Jews". In the second part, a speaker responded by saying, "my family was killed by Hitler and Nazis during the holocaust", "seventeen death in my family alone," and "how dare you say that?".	IG	Take down	Keep up	Turkey & United States	Dangerous Individuals & Organizations
11, 12, 13	Assigned to panel as “political dispute ahead of Turkish elections” (2023-007-FB-UA, 2023-008-FB-UA, 2023-009-IG-UA)	FB & IG	Take down	Keep up	Turkey	Hate Speech
14	The content contained a meme of Kermit the Frog sipping tea with lines asking why God could not open the gas chamber doors despite having split the sea in half. The content also contained a caption asking if God was tired of helping them (Jews).	FB	Keep up	Take down	Ghana	Hate Speech
15	The content was a video posted by a user who appeared to be inside a parked car near a Jewish community in New York. The footage captured Jewish people as they were walking down the streets, carrying babies, or waiting for a ride. At the same time the user, speaking in Bengali, was	FB	Keep up	Take down	United States	Bullying & Harassment

	mocking people featured in the video by saying that Jews "like to get free things", attacking women by stating that they "get fat" and claiming that the Yehudi people "give birth every month like this because they need to expand their clan."					
16	The content contained an image with the statement: "Kahane was Right. It's time to take out the garbage." Underneath the statement were multiple Israeli flags. The image was accompanied by the caption "enough is enough."	IG	Keep up	Keep up	Israel/ Palestine	Policy changed from Hate Speech to Dangerous Individuals & Organizations
17	The content contained a meme depicting two dogs: one dog was wearing a wig, while the other was wearing a yamaka and had a long, pointy nose. The two dogs were depicted as possessing human arms; the two dogs shook hands. The dogs were in front of what resembled a burning church, which had a Star of David affixed to it. The meme identified the dog wearing a wig as "women" and the dog with the yamaka as "small hat tribe". Below the image of the dogs was the line "lying about body count."	FB	Keep up	Take down	United States	Hate Speech

Annex II - Measuring the effect of Facebook implementation of Iran protest slogan recommendation no. 3.

Summary: In Q1 2023 Meta reported to have implemented the recommendation no. 3, to allow the use of the phrase "marg bar Khamenei" ("ای خامنه بر مرگ") in the context of ongoing protests in Iran. Using public data obtained from CrowdTangle, we ran a data analysis to investigate whether there was a statistical difference in the number of posts present on Facebook and Instagram following the implementation when compared with the same data pulled prior to the Board's case decision, derived from implementing the Oversight Board recommendation. We did not find statistically significant differences in the number of posts on Facebook groups and pages that were restored as a result of implementing the recommendation. However, a statistical difference was observed in the number of posts on Instagram that were restored as a result of the implementation of the recommendation.

Analysis: This study aims to investigate whether there is a statistical difference in the number of posts present on Facebook and Instagram for the same time period before the Board decided the "Iran protest slogan" case, compared with after Meta implemented the Board's Iran protest slogan recommendation no. 3. Our study focused on examining the effect of the recommendation being applied to different groups, pages, and accounts.

Methodology:

Data was obtained using CrowdTangle, a public insights tool owned and operated by Meta. Posts from Instagram and Facebook were collected for the period from July 17, 2022, to October 17, 2022, both by Memetica (a research consultancy that works with the Board) prior to the case decision, and by the Board's Implementation Team following the decision's implementation. We made sure that only the posts with the slogan "ای خامنه بر مرگ" were included in all analyzed datasets. Specifically, the datasets collected were:

Facebook Datasets:

- A dataset provided by Memetica, in October 2022, spanning July 17 - October 17, 2022, of Facebook posts that included the phrase "ای خامنه بر مرگ". The dataset contains a total of 504 Facebook posts from 114 groups or pages.
- A dataset pulled by the Data and Implementation Team in May 2022 of Facebook posts that mention the phrase "ای خامنه بر مرگ". The dataset spans the same time period as Memetica's original dataset and contains a total of 488 Facebook posts from 104 groups or pages.

Instagram Datasets:

- A dataset provided by Memetica in October 2022, spanning July 17 - October 17, 2022, of Facebook posts that included the phrase "ای خامنه بر مرگ". The dataset contains a total of 94 posts from 41 accounts.
- A dataset pulled by the Data and Implementation Team in May 2022 of Facebook posts that include the phrase "ای خامنه بر مرگ". The dataset spans the same time period as Memetica's original dataset and contains a total of 102 posts from 43 accounts.

Assumptions:

It is important to note that since CrowdTangle was the only source of information we had available to obtain the data, we filtered our datasets on the basis of the following assumptions:

- We discarded the groups, pages and accounts that appeared in Memetica's datasets but did not appear in our datasets. This is because we are unable to account for what happened to the posts that were missing, as they could have been removed for different reasons (i.e. users could have taken them down, etc). Overall, we found 14 groups with 22 posts on Facebook and 10 accounts with 10 posts on Instagram that met this assumption and eliminated them from the original Memetica datasets. This means that both our dataset and Memetica's original datasets now contained the same groups, pages and accounts.
- We considered posts that appear in our dataset but do not appear in Memetica's datasets to have been restored by Meta as a result of implementing the recommendation. Even though this assumption has its limitations, it is the only way that we have to infer that certain posts were restored by Meta.

Following our assumptions for filtering out the datasets, we uncovered 38 posts that appeared in our data and not Memetica's, suggesting they were restored in the time since Memetica's October 2022 analysis. 20 of those posts were on Facebook, mostly stemming from one anti-Khamenei group with 55,000 members. The remaining 18 posts were from an array of Instagram pages and have earned a cumulative 111,000 interactions on the platform. These posts contain a mix of support for, coverage of, and commentary on protests against the Iranian government.

Furthermore, our dataset was missing 14 posts that were present in the Memetica dataset. We cannot ascertain whether these posts were removed in the time since Memetica's analysis and never reinstated by Meta, whether the users removed the posts themselves, or if there is another explanation for their absence. Nevertheless, we made the choice to incorporate them into our statistical analysis because they were initially part of the Memetica dataset. By including them, we ensure that our analysis captures the overall impact of the recommendation implementation, taking into account any potential changes in the post landscape since the Memetica analysis.

Facebook Data Analysis

This first analysis aims to investigate whether there was a statistically significant difference in the number of posts restored in the different Facebook groups and pages contained in the datasets derived from implementing the board's Iran protest slogan recommendation no. 3 (Facebook datasets).

To determine the statistical difference, we employed a Wilcoxon Signed-Rank Test, a nonparametric test suitable for analyzing paired data. This choice was made considering the small sample size and the non-normal distribution of the data. A significance level (α) of 0.05 (5%) was chosen to assess the results.

Finding: The Wilcoxon test analysis revealed that there was no statistically significant difference at the 5% level in the number of posts in the groups and pages in the Memetica dataset and our dataset on Facebook (p -value=0.9322, $W=0$, $\alpha=0.05$). Consequently, we cannot conclude that the observed

changes in the number of posts are directly attributed to Meta implementing the recommendation. The lack of statistical significance suggests that factors other than the recommendation implementation may be responsible for the observed variations in the post numbers.

Instagram Data Analysis

This second analysis aims to investigate whether there was a statistical difference in the number of posts restored in the different Instagram accounts, derived from implementing the board's Iran protest slogan recommendation no. 3 (Instagram datasets).

To determine the statistical difference, we employed a Wilcoxon Signed-Rank Test, a nonparametric test suitable for analyzing paired data. This choice was made considering the small sample size and the non-normal distribution of the data. A significance level (α) of 0.05 (5%) was chosen to assess the results.

Finding: The Wilcoxon test analysis revealed that there was a 28.75% increase between the number of posts across in the Instagram accounts in the Memetica Dataset and our dataset – a statistically significant difference (p -value= 0.001977, $W=13$, $\alpha = 0.05$). This indicates that the changes seen in the number of posts are unlikely to be solely attributed to random variation. Instead, they are likely to be a result of Meta implementing the Oversight Board recommendation. However, it is crucial to acknowledge that assumptions were made that in turn led to these results.

Conclusion. In this study, we ran a data analysis to investigate whether there was a statistical difference in the number of posts restored on Facebook and Instagram, derived from implementing the Oversight Board's recommendation in the "Iran protest slogan" case. Our analysis showed that the effects of implementing the recommendation by Meta are statistically significant for Instagram but not for Facebook. However, it is crucial to interpret the results cautiously, considering the limitations arising from the inadequate availability of suitable data and the assumptions made to be able to perform this analysis. For instance, it is important to note that we cannot confidently attribute the observed increase in the number of posts in our dataset solely to the restoration following the implementation of the recommendation. Furthermore, it should be noted that CrowdTangle only tracks public content (including pages and public groups and accounts). Consequently, it is not possible to evaluate the impact of Meta implementing the recommendations on private groups and users on Instagram and Facebook. This limitation restricts the extent to which our findings can be generalized to the overall user base affected by the recommendation. Additionally, as our only source of information was CrowdTangle, we could only make certain assumptions with the public data obtained, but we acknowledge that other factors may influence the observed differences.

Annex III - Glossary of terms

Annual report – A report published each year 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, which decides which cases the Board will review, out of thousands of user appeals and Meta referrals.

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

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

Meta’s legal review – Step in case selection process where Meta may exclude cases from the shortlist that 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).

Meta-referred case – A case submitted to the Oversight Board by Meta. Meta has the ability to expedite cases for 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.

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

Panel – 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.