

December 2025

Meta Final Update: Israel and Palestine Human Rights Due Diligence

TABLE OF CONTENTS

Background	2
Meta Recommendations Implementation Update	4
Overview	4
Update to our response	6

Background

Meta¹ commissioned [Business for Social Responsibility \(BSR\)](#) to carry out an independent Human Rights Due Diligence (HRDD) Exercise on Israel and Palestine for the period of May 2021. This work was scoped as rapid human rights due diligence rather than a full Human Rights Impact Assessment in order to enable swift project launch and implementation. Although planning for rapid due diligence began prior to the Oversight Board's [September 2021 recommendation in the Al Jazeera case](#), the exercise seeks to be responsive to the Oversight Board's recommendation to:

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.

The HRDD was conducted by BSR in 2021 and 2022 in line with the United Nations Guiding Principles on Business and Human Rights. BSR is an organization of sustainable business experts that works with a global network of the world's leading companies to build a just and sustainable world. BSR's methodology and findings are published in the Insights and Recommendations Report of the Due Diligence Exercise.

¹ On October 28 2021, Facebook, Inc. changed its name to Meta Platforms, Inc. For consistency, this report uses "Meta" to refer to the company both before and after October 28 2021. References to "Facebook" apply only to the social media platform, not the company as a whole. Further, this response references steps taken, or plans to take steps, by Meta as a company regarding a specific entity. Such a statement is not intended to imply that Meta would, or will, take steps regarding all entities. No statement in this response is intended to create — or should be construed as creating — new obligations (legal or otherwise) regarding the application of a policy or procedure to other products or entities. For example (and in contrast to other Meta technologies), WhatsApp is an end-to-end encrypted messaging and calling application with unique human rights touchpoints. This response's discussion of content moderation and related issues on Facebook and Instagram do not apply to WhatsApp. Unless a policy or commitment is specified as applying to WhatsApp, it does not apply to WhatsApp.

The HRDD made 21 prioritized recommendations. In September 2022, Meta committed to implement 10 recommendations and partly implement four others. It committed to assess the feasibility of another six. Meta declined to take further action in relation to one recommendation.

In providing an update on the status of our response to the recommendations made in the HRDD, we used the following categories, which are the same as the ones we use for [reporting](#) on implementation of Oversight Board recommendations:

- **Complete**: We have completed full or partial implementation of our commitments in response to BSR's recommendation. We will not share future updates on these items.
- **In progress**: We are continuing to make progress, and will have further updates in the future.
- **No further updates**: We have already addressed the recommendation through an action we already do, or will not implement the recommendation. We will not share future updates on these items. updates

Our previous updates are available [here](#) and [here](#).

Meta Recommendations Implementation Update

Overview

This is our final update on Meta’s progress in implementing the 21 recommendations made by BSR. This update is part of our [broader commitment](#) to meaningful transparency about our human rights due diligence and about our integrity work.² In total, we fully implemented 16 out of 21 recommendations made by BSR, we are continuing implementation of one recommendation related to transparency reporting and not taking any further action on four recommendations.

At the end of the last reporting period (June 30, 2023 to June 30, 2024), implementation of 5 recommendations was still in progress. Since then, between June 30, 2024 and June 30, 2025, we completed implementation of two recommendations, including the launch of a system that detects and prioritizes content to moderators most likely to understand that particular Arabic dialect and the revamp of our Trusted Partner escalation channel leading to an improved rapid response to escalations. Implementation of one recommendation related to transparency reporting was still in progress and will be reported on as part of future Oversight Board Biannual Updates. We decided to take no further action on the remaining two recommendations, related to more granular content moderation decisions and hateful conduct prevalence tracking by protected characteristics.

Implementation status update as of June 2025			
BSR recommendations	Recommendations fully implemented ³	Implementation in progress	No further action
21	16	1	4

² Meta’s publication of this response should not be construed as an admission, agreement with, or acceptance of any of the findings, conclusions, opinions, or viewpoints identified by BSR, or the methodology employed to reach such findings, conclusions, opinions, or viewpoints. Likewise, while Meta references steps it has taken, or plans to take, that may correlate to points BSR raised or recommendations it made, these also should not be construed as an admission, agreement with, or acceptance of any findings, conclusions, opinions, or viewpoints.

³ In this update, we added the four recommendations which we had completed by September 2022 as actions we already did prior to the due diligence to the ‘recommendations fully implemented’ category for ease of understanding. We had previously counted it as part of the ‘No further action’ category.

Update to our response

Recommendation 5: Provide users with a more specific and granular policy rationale when strikes are applied. This should not just include the category of the violation, but how a post was violating, so that users can better understand the justification, submit an informed appeal, and be less likely to post violating content in the future.

Meta commitment: **Implementing in part**

Status update: **No further action**

In our [September 2024 Israel Palestine Human Rights Due Diligence update](#), we reiterated that we already provide specific granular reasoning when content is removed and strikes are applied in the vast majority of cases and that we were working to expand this to the limited number of areas where we do not yet provide this specificity. We are committed to regularly evaluating how we may improve the granularity of user messaging, including under our DOI policy. However, due to other higher priority workstreams, we are not prioritising this work stream at this time.

Recommendation 9: Continue establishing mechanisms to better route potentially violating Arabic content by dialect/region.

Meta commitment: **Assessing feasibility**

Status update: **Complete**

In our previous Israel Palestine Human Rights Due Diligence updates ([here](#) and [here](#)), we said that we were developing mechanisms to efficiently route content by Arabic dialect. We have determined following an extensive internal assessment that creating specialized routing for a variety of Arabic dialects across our systems would contribute to a greater precision in our Arabic content moderation for high severity content.

In November 2024, we launched a system that detects and prioritizes content to moderators most likely to understand that particular Arabic dialect. This new system is also designed to ensure that our content moderation responds dynamically to shifts in demand. This means that we can route content to reviewers who are best equipped to understand the specific dialect and cultural nuances of the content.

With these improvements, we expect our content moderation system for Arabic dialects to enforce more accurately, reduce errors, and be better positioned to respond to surges in violating content.

Recommendation 14: Continue plans to disclose the number of formal reports received from government entities (including the Israel State Attorney Office (ISAO) in Israel) about content that is not illegal but potentially violates Meta content policies. This should take place either quarterly (as part of the Community Standards Enforcement Report) or every six months (as part of the Content Restrictions Report).

Meta commitment: **Implementing**

Status update: **In progress**

Meta remains committed to implementing this recommendation. As we shared in our [September 2024 Israel/Palestine Human Rights Due Diligence update](#), we are still in the process of developing consistent and reliable systems for gathering metrics on the number of pieces of content removed under the Community Standards as a result of government requests. We continue to evaluate approaches to building the necessary internal data logging infrastructure to enable us to publicly report this information across the diversity of request formats we receive.

As noted previously, this is a complex, long-term project whose implementation will take time in part due to the need to assess existing reporting and other regulatory reporting requirements. We will endeavor to provide an update on the timeline for public reporting of these metrics in future [Oversight Board Biannual Updates](#).

Recommendation 16: Develop a mechanism to track the prevalence of content that attacks based on specific protected characteristics (e.g., antisemitic, Islamophobic, homophobic content). This might involve, for example, prompting users to mark relevant hate speech content with tags.

Meta commitment: **Assessing feasibility**

Status update: **No further action**

While we remain committed to improving our performance metrics (including prevalence), we are facing resourcing and capacity to pursue this specific work. Our evolving company landscape requires us to address and prioritize overall changes and updates across our policies, which currently takes precedence over granular assessments for specific protected characteristics.

We will continue to engage with representative groups and experts to inform our work as we take these issues seriously and will continue working towards creating a safe and inclusive online environment. However, we will not be taking any further action on this particular workstream.

Recommendation 18: Increase the capacity of Meta’s special escalation channels via more staff and more resources to enable sufficiently prompt response to escalations from Trusted Partners, governments, and other actors in both normal times and times of crisis.

Meta commitment: **Implementing in part**

Status update: **Complete**

At the end of 2023 we conducted a comprehensive review of our civil society escalations channels and consulted with external partners about how to improve the program. In line with the findings of our internal review and feedback received from external partners, we implemented key changes to the program, which include clarifying the scope of these escalation mechanisms and providing improved tools to facilitate partner reporting and increase the efficiency of review processes. These changes have already led to an improved rapid response to escalations in line with the spirit of the recommendation.

Trusted Partners continue to play a critical role in reporting content trends and helping us keep our communities safe.

We will continue onboarding Trusted Partners in the region as needed and in line with our onboarding criteria.