Meta Q1 2023 Quarterly Update on the Oversight Board

Meta

TABLE OF CONTENTS

Introduction

At Meta, we believe in empowering outside voices to weigh in on our policy decision-making process and giving people control over what they see on our platforms. That's why we've developed multiple tools, including the Oversight Board, that allow people to control their experience and have created mechanisms to hold us accountable. We are therefore committed to publishing regular updates[1] to give our community visibility into our responses to the Oversight Board's independent decisions about some of the most significant and difficult content decisions Meta makes. These Quarterly Updates provide regular check-ins on the progress of this long-term work and share more about how Meta approaches decisions and recommendations from the board. This update covers open recommendations from decisions issued by the board during and prior to Q4 2022, and it includes details of (1) Meta's content referrals and Policy Advisory Opinion requests to the board and (2) our progress on implementing the board's non-binding recommendations. This report is meant to strengthen transparency and hold us accountable to the board and to the public.

---

[1] We base these Quarterly Updates on best practices in human rights reporting principles, corporate disclosures, and goal-tracking reports. These include the Value Reporting Foundation's Integrated Reporting Framework and Sustainability Accounting Standards Board (SASB) Standards, Global Reporting Initiative (GRI) Reporting Principles, and the UN Guiding Principles for Business and Human Rights (UNGPs), among others.

I. Meta's Content Referrals & Requests for Policy Advisory Opinions

There are three ways cases can reach the Oversight Board for review: appeals by people, case referrals by Meta, and requests for Policy Advisory Opinions (PAOs). This means that people who use Facebook and Instagram may appeal our content enforcement decisions directly to the board, and that we also regularly and proactively seek input from the Oversight Board on some of the most significant and difficult content decisions, policies, and enforcement issues we face. We previously outlined how we prioritize cases for Meta content referrals in our [Newsroom](). Both Meta content referrals and PAOs generally involve issues that are significant, large-scale, and/or important for public discourse.

The Meta content referral process begins with an internal review of content decisions that are geographically diverse, cover a wide range of policies found in our Facebook [Community Standards]() and Instagram [Community Guidelines](), and represent both content we have taken down and content we have left up. Teams with expertise on our content policies, enforcement processes, and specific cultural nuances from regions around the world review the candidate cases for significance and difficulty. Finally, we refer the most challenging of these content decisions to the board. The board has sole discretion to agree or decline to review the content decisions referred through this process. The board's decisions on Meta content referrals and user appeals about Meta's content decisions are binding.

For PAO requests, we ask the board to advise us on our policies and content moderation systems more generally. Once the board issues the PAO, we consider and publicly respond to its recommendations within 60 days. While these recommendations are not binding, the board's guidance through the PAO process holds us publicly accountable for our policies, processes, and decisions. When we receive a recommendation from the board, it is integrated as an additional and important input and we make a substantial effort to consider and implement each recommendation.

Between January 1, 2023 and March 31, 2023, we submitted eight content referrals and one PAO request to the board:

1. A member of the Maharashtra Legislative Assembly in India posted photos on Facebook from a recent event. One of the photos depicts a Jain Saint who is nude, with visible genitalia. We removed the content for violating the Adult Nudity and Sexual Activity policy.

2. A Facebook page called "Azerbaijan War Crimes" posted a video reportedly showing Azerbaijani soldiers beating Armenian prisoners of war. The content violated Meta's Coordinating Harm and Promoting Crime policy, which prohibits identifying prisoners of war. However, given its value in raising awareness of human rights violations, we granted a newsworthy allowance, marked the content as sensitive, and allowed it to remain on the platform. The board selected this case in March.

3. A British art institution posted photos by a highly-recognized photographer, depicting a man running in the nude in violation of our Adult Nudity and Sexual Activity policy. After weighing the public interest value of the content against the risk of harm, we granted a newsworthiness allowance for the content, allowing it to remain on the platform (restricted to people aged 18+) with a warning screen.

4. The Prime Minister of Cambodia posted a video on his Facebook Page featuring a lengthy speech that included a threat, which violated our Violence and Incitement policy. However, we determined that the content's public interest value outweighed the risk of harm and therefore granted a newsworthy allowance. The board selected this case in March.

5. A Portuguese newspaper posted content on their Facebook page with a link to an article. The title of the article referenced indigenous people with mixed feelings after Pope Francis' visit. The thumbnail of the article contained an image which included a woman with uncovered nipples. Meta issued a newsworthiness allowance (and restricted the content to people aged 18+) given the non-sexualized nature of the nudity and the public interest value of the content.

6. The National Rifle Association of America Facebook page posted content quoting a county sheriff who advocated for the use of guns in self-defense in the context of a home invasion. We allowed the content to remain on the platform.

7. An Instagram user posted an image with a caption describing a positive experience with supervised use of ketamine to treat her anxiety and depression, a statement that reflects the medical community's increasing off-label use of ketamine to treat these mental health conditions. However, ketamine is also used illicitly for getting high. We allowed the content because it promoted ketamine as a pharmaceutical (rather than non-medical) drug. The board selected this case in May and it will be announced in the coming weeks.

8. A Brazilian actress and social media influencer reported a post that accused her of encouraging hedonism. We determined that the post did not violate our Bullying and Harassment policy because it did not contain a Tier 1 or Tier 3 attack and because the user is a public figure.

9. We [requested](#) a policy advisory opinion from the Oversight Board regarding the treatment of the term "shaheed" (شهيد) when used to refer to an individual designated under Meta's [Dangerous Individuals and Organizations](#) policy. The board [selected](#) this PAO in March.

Q1 Meta referred content decision case breakdown

● **TOTAL Q1 2023**   ● **TOTAL THROUGH Q1 2023 (3/31/23)**

9   77
Total number of referrals
(Meta content referrals + policy advisory opinions)

1	5	1	4	8	72	3	13
Total number of policy advisory opinion requests		Total number of policy advisory opinion requests selected		Total number of Meta content referrals[2]		Total number of Meta content referrals selected	

POLICY VIOLATION

Violence and Incitement
- 1
- 12

Violence and Graphic Content
- 0
- 0

Dangerous Individuals and Organizations
- 0
- 2

Hate Speech
- 0
- 4

Bullying and Harassment
- 0
- 2

Child Exploitation, Abuse, and Nudity
- 0
- 5

Adult Nudity and Sexual Activity
- 3
- 5

Restricted Goods and Services
- 0
- 1

Coordinating Harm
- 1
- 1

Not Applicable[3]
- 3
- 40

REGION[4]

Europe
- 2
- 20

United States and Canada
- 3
- 21

Middle East and North Africa
- 0
- 9

Asia Pacific and Oceania
- 1
- 8

Central and South Asia
- 1
- 4

Latin America
- 1
- 7

Sub Saharan Africa
- 0
- 3

[2] The numbers in the Policy Violation and Region sections of this table only apply to Meta Content Referrals, not PAOs.

[3] When we decide to leave content up – unless there is a newsworthiness or other noted allowance – there is, by definition, no policy violation. As a result, we categorize the policy violation as "not applicable" for referrals of content we left up on Facebook or Instagram.

[4] Meta defines "region" according to an analysis of several factors, including the location of the posting user, the language(s) the content includes, and the countries/regions referenced in the content.

II. Progress on Oversight Board Recommendations & Institutional Impact

As we kicked off a transformational year at Meta, the Oversight Board provided important oversight of our policies, operations, and products and continued to hold us accountable to our promises. The board's recommendations offer a crucial overlay of global human rights frameworks and diverse perspectives to our most significant and difficult decisions.

As described in previous Quarterly Updates, we expect the board to significantly increase its output and impact this year. With two new case types - expedited review and summary decisions - we expect the board will be able not only to increase the volume of cases it reviews and decisions it publishes, but also weigh in quickly on time-sensitive matters. We have already responded to 61 board recommendations from January 2023 to May 2023. These recommendations span a broad range of focus areas and scale the board's influence across our business. We respond to every Oversight Board recommendation publicly and have committed to implement or explore the feasibility of implementing 78% percent of recommendations to date.

In Q1 2023, because of the board's recommendations we:
- Expanded users' ability to appeal eligible content decisions, including those made on escalation, to the Oversight Board. Previously, content decisions that we made during an internal escalation process were often not appealable because of separate pathways used for specialized contextual review. We expect this product update to expand the volume and scope of content that users are able to appeal to the Oversight Board, further increasing the board's impact.
- Completed a feasibility assessment and finalized a plan to create a new crisis coordination team to provide dedicated 24/7 Operations oversight throughout the lifecycle of imminent and emerging crises, with dedicated regional support.
- Completed the global roll-out of new, more specific messaging to notify people of a change to the status of their content or content they have reported, because of an appeal to the Oversight Board, including providing secondary notifications should the status of that content change again.
- Added new information on our mistake prevention systems in the "[Detecting Violations](#)" section of our Transparency Center, following the board's decision on our cross-check systems.

- Established protocols for applying certain cross-check tags on entities included in the cross-check program for a short-term period in cases where the increase in risk is episodic or a shorter time period is otherwise more appropriate.
- Allowed the phrase "marg bar Khamenei" in the context of the ongoing protests in Iran and ran sweeps to identify previous enforcement actions and reverse strikes made on the basis of this type of content.
- Continued to publish Community Standards translations, including publishing the Community Standards in Armenian, Belarusian, and Kazakh, making the Community Standards available in a total of 80 translations.

Policy Advisory Opinions Driving Meaningful Change

In April 2023, the board published its third Policy Advisory Opinion (PAO) regarding Meta's treatment of harmful health misinformation in the context of COVID-19. This PAO, along with its two predecessors – concerning our rules around private residential information and our cross-check mistake prevention system, respectively – was an opportunity for the board to provide us with in-depth analyses and external guidance on our broader policies and systems, rather than individual pieces of content. These PAOs have taught us that it is not only the board's recommendations that create change – the process of responding to the board's wide-ranging inquiries is also an important exercise for our internal teams. We conduct multiple briefings, answer written questions, and provide historical and current documentation in order to provide the board with a complete picture of the systems or policies they are weighing in on. Developing these resources is a valuable opportunity for our teams to increase internal visibility across the full ecosystem of our platforms. The process brings Policy, Operations, Product and other teams together to ensure the board understands how every function interacts. As we saw in the Cross-Check PAO, many of the board's questions at the beginning of the process spurred immediate action, helping teams to identify areas of improvement proactively, even before the board's decision was finalized. For example, after the board selected the Cross-Check PAO in October 2021, we launched a new Early Response Secondary Review (ERSR) list-based prevention system in early 2022, and even began publishing the categories of entities we enroll in the ERSR system to a new Transparency Center page based on the focus of the board's initial inquiries. While we continue to pursue opportunities to better measure the impact of the Oversight Board, we recognize that the impact of the board is often bigger than its recommendations or decisions. The mere existence of the board and the oversight it provides Meta has prompted valuable self-reflection and new perspectives for us to consider.

As we shared in our [last Quarterly Update](#), the board's thoughtful Cross-Check PAO decision affirmed our belief in the value of its guidance on our policies and the complex technology and operations that help us uphold them at scale. Since our last update, we responded publicly to all 33 of the board's [cross-check recommendations](#), committing to implementing 82% either in part or in full and making significant changes to our systems and level of transparency. This includes:

- Structured and robust engagement with our internal Human and Civil Rights teams, our Trusted Partners, and other external civil society organizations to explore ways to inform the criteria we use to identify public interest entities for cross-check lists.
- Exploring a more formal cross-check list nomination process from global, regional, and local civil society groups.
- Investing in quality review and training resources as we move to staff all cross-check decisions with reviewers who speak the language and have regional expertise wherever possible.

Updates in Our Reporting and Approach to Data Sharing

In an effort to better align Meta's tracking of the Oversight Board's recommendations with the board's own public reporting of these same recommendations, we have re-assessed our definitions of the three categories into which they are divided. These are: 1) Policy 2) Transparency and 3) Enforcement. Previously, the Oversight Board and Meta's definitions for these categories diverged slightly, which resulted in substantially different counts for the number of recommendations included in each category. We felt that this created unnecessary confusion for our external stakeholders and re-categorized 169 recommendations included in this and past Quarterly Updates to fully align with the board's public reporting. We have also convened the Implementation Working Group, composed of implementation teams at the Oversight Board and Meta, to agree to shared definitions for these categories going forward. We hope this will help the public better track the types of impact the board has had on the way we do business. As shared in previous Quarterly Updates, the Implementation Working Group convenes regularly and, in addition to serving as an important opportunity to share progress updates, it also allows us to improve our communication and data-sharing with the Oversight Board to ensure that we can tie their recommendations to specific, measurable impacts.

For the first time since we began publishing Quarterly Updates, we've introduced a new category for recommendations: long-term recommendations. These are recommendations that we expect to take longer than two years to implement. Typically this is because those recommendations have been included on one of our team's long-term roadmaps–many of which outline a one- to

three-year vision with a series of dependencies that must be completed before the recommendation can be finalized. As the number of recommendations from the board continues to grow, we see this category as an important tool for ensuring that our Quarterly Update appendices don't become overly long and difficult to navigate. Importantly, this category will also allow our teams to focus on implementation progress without the distraction of reporting urgency. Our aim is not to inhibit transparency—we will always provide at least one standard Quarterly Update progress update before a recommendation is moved to the long-term category and we will always set a clear expectation for when we will provide the next meaningful update on our progress. Instead, this will ensure that the standard appendix updates are always recent, comprehensive, and substantial and allow the teams working on long-term efforts behind the scenes to focus their capacity on implementation.

As we examine our company-wide operations to ensure efficiency, we also recognize the importance of measuring the impact of the Oversight Board. As this and past Quarterly Updates emphasize, recommendations are the primary means of scaling the Oversight Board's influence on the company. In the past, we had no system for measuring that influence. This meant that we weren't able to quantify or demonstrate the true impact of the Oversight Board. To address this information gap, we have embarked on a pilot program focused on measuring the impact of the board's recommendations quantitatively. Because these recommendations are so varied, there is no one-size-fits-all approach to gathering data on their reach. Potential metrics for understanding each recommendation's impact must be scoped individually and data validation for each metric is complex and extensive. This is a time and resource-consuming endeavor, but it is an important one. This work will help us to better gauge the scale of the board's impact on the people who use our platforms. We shared our first recommendation implementation metric with the Oversight Board Implementation Committee in April 2023 and we look forward to sharing many more, with the board and with the public, as this program matures.

Expanding the Oversight Board's Scope & Impact

To better serve our growing audience, Meta hopes the Oversight Board will expand its scope and impact in 2023. At the start of the year, the Oversight Board [announced](#) plans to begin publishing two new types of cases: expedited review decisions and summary decisions.

Expedited review decisions provide the board with a new avenue for reaching case decisions on high-impact content with time-sensitive implications. Meta's internal teams will refer these cases to the board, at which point the Oversight Board will have the ability to accept or decline the case.

Should the board accept, the board and Meta will hold a short live Q&A briefing prior to the board issuing a binding decision on the content within 48 hours.

Summary decisions are binding non-precedential, short form decisions where Meta identified incorrect outcomes brought to our attention through the board's shortlisting process, we corrected those decisions, and the board agreed with our final enforcement decision. Summary decisions allow Meta to receive a decision from the board without conducting time and resource-intensive briefings, submitting lengthy case files, or decision rationales. They provide additional transparency around Meta's enforcement decisions and data points to enable Meta to avoid similar errors in the future.

As noted in our Q4 2022 update, in addition to expanding the board's influence, these changes are expected to enable the board to review more cases and to do so faster than before. In Q1 2023, we unfortunately did not see a material increase in case reviews and decisions as compared to previous quarters. We hope that the board will significantly increase its output next quarter and beyond.

This is a year of change for Meta, and we see the board's guidance as a crucial input in shaping our evolution. Because the people who use our products live all over the world, we need globally representative perspectives to inform our work. The board's grounding in international human rights standards and extensive engagement with diverse stakeholders ensures that Meta's decisions aren't made in a vacuum. We are grateful for the board's oversight and accountability and look forward to more in 2023.

1. How to Read This Update

From January 2021 through March 2023, the board issued 193 non-binding recommendations. In our Q4 2022 Quarterly Update, we addressed 60 of these recommendations and indicated we would provide more information on 50 in our next update. In this update, we address those 50 recommendations, 33 recommendations the board included in its Q4 2022 Cross-Check PAO decision[2], and 14 new recommendations the board included in the decisions it issued in Q1 2022 for a total of 97 recommendations. We categorize our commitments to the board's recommendations as follows:

- **Implementing fully:** We agree with the recommendation and have or will implement it in full.
- **Implementing in part:** We agree with the overall aim of the recommendation and have or will implement work related to the board's guidance.
- **Assessing feasibility:** We are assessing the feasibility and impact of the recommendation.
- **No further action:** We will not implement the recommendation due to, for example, a lack of feasibility or disagreement about how to reach the desired outcome.
- **Work Meta already does:** We have addressed the recommendation through an action that we already do.
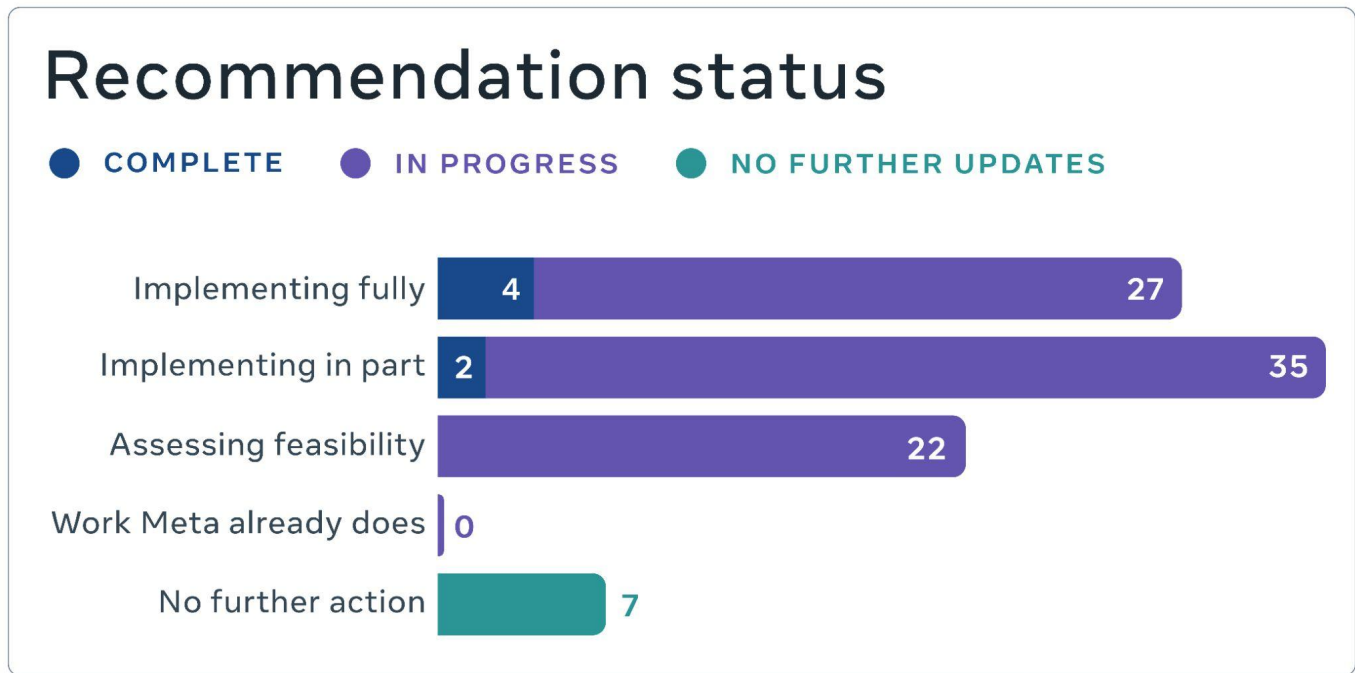
The current status for our responses to the board's recommendations are defined as:

- **Complete**: We have completed full or partial implementation in line with our response to the board's recommendation and will have no further updates on the recommendation.
- **In progress:** We are continuing to make progress on our response to the board's recommendation and will have further updates on the recommendation.
- **No further updates:** We will not implement the recommendation or have addressed the recommendation through an action that we already do and will have no further updates on the recommendation.

---

[2]As noted in our Q4 2022 Quarterly Update, Meta publicly responded to the board's 33 Cross-Check PAO recommendations on March 6, 2023. Therefore, they were omitted from the Q4 '22 update (published in February 2023) and are included in this update instead.

The below graph depicts the status of each of the current 97 recommendations:

Recommendation status

● COMPLETE   ● IN PROGRESS   ● NO FURTHER UPDATES

Implementing fully	4 — 27
Implementing in part	2 — 35
Assessing feasibility	22
Work Meta already does	0
No further action	7

Of the 33 recommendations issued in the Q4 2022 Cross-Check PAO decision and the 14 new recommendations issued in Q1 2023, we do not have an update for six. As explained in our 60-day response, we will either take no further action on these recommendations or the recommendations were addressed by work Meta already does.[3] We organize our substantive updates on the 91 open recommendations into three sections:

A. **Transparency (17 recommendations):** Helping people understand the rules on Facebook and Instagram, what violates them, and the consequences of violating them.

B. **Policy (28 recommendations):** Ensuring the Facebook Community Standards and Instagram Community Guidelines are clear and align with our values of voice, authenticity, safety, privacy, and dignity.

C. **Enforcement (41 recommendations):** Improving the quality and efficacy of our content moderation operations at scale.

Long-Term Recommendations Appendix

As shared in Section II, this Quarterly Update marks the launch of a new appendix for long-term recommendations. These are recommendations which, due to the complexity of their

---

[3] This applies to recommendations #5, #6, #12, #13, and #29 in the PAO on Meta's Cross-Check Policies, and recommendation #2 in the Donation of Pharmaceutical Drugs to Sri Lanka case.

implementation, placement on long-term roadmaps, and/or dependencies on completion of other recommendations or foundational work, we do not expect to have significant updates to share within the next two years. In this Quarterly Update, we are reporting on 5 long term recommendations:
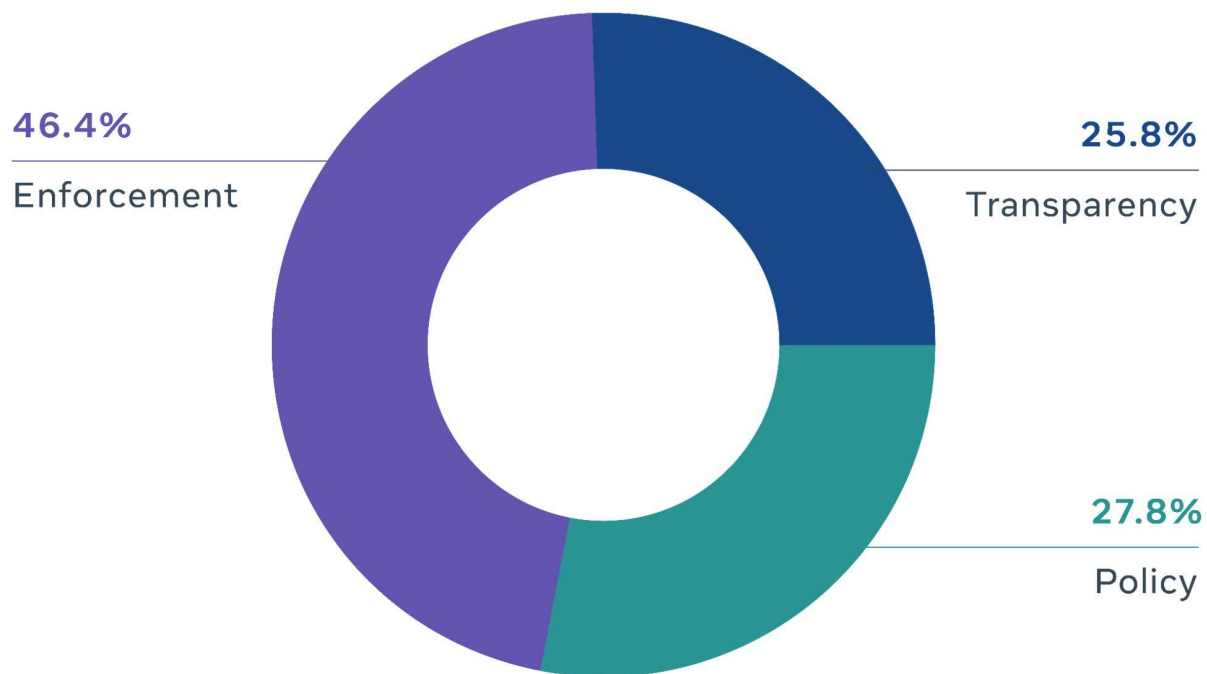
Long-Term Transparency (4 recommendations):

1. [Breast Cancer Symptoms and Nudity #6](#),
2. [Former President Trump's Suspension #18](#),
3. [Punjabi Concern Over the RSS in India #3](#), and
4. [Veiled Threat of Violence Based on Lyrics from a Drill Rap Song #7](#)

Long-Term Enforcement (1 recommendation):

1. [Post Requesting Advice on Pharmaceutical Drugs #3](#)

By structuring our updates this way, we aim to facilitate discussion about the progress made in these areas and improve the navigability of this document. In the body of each section, we provide a general overview of our progress. For further detail and the full text of each recommendation, please refer to the [Appendix](#).

Recommendations by Category

46.4%

Enforcement

25.8%

Transparency

27.8%

Policy

2. Oversight Board Recommendation Implementation Highlights

A.   Transparency

We are providing updates for our work on 17 board recommendations that address transparency and accountability.

We want to highlight our progress on:

1. **Continuing to translate our Community Standards to new languages:** In recommendation #1 in the [Punjabi Concern Over the RSS in India](#) case, which we implemented in [Q4 2021](#), the board asked us to set a goal of making our Community Standards accessible in all languages widely spoken by people who use Facebook and Instagram. By Q4 2021, we had progressed to publishing the Community Standards in 59 translations. Since then, we have continued this work and, in Q1 2023, published the Community Standards in Armenian, Belarusian, and Kazakh, making the Community Standards available in a total of 80 translations.

For a comprehensive list of all 17 recommendations in this category, see [Appendix A. Transparency](#).

B.   Policy

We are providing updates for our work on 28 board recommendations that address the Facebook Community Standards and Instagram Community Guidelines.

We want to highlight our progress on:

1. **Publishing new information about our mistake prevention systems to our Transparency Center:** In our initial response to the board's Cross-Check PAO, we shared that we had recently published information on how we detect and enforce violations, including via our cross-check systems, to our Transparency Center. We have now expanded upon this information by including further details around our mistake prevention systems in the "Detecting Violations" section of our Transparency Center.

2. **Ongoing Policy Development:** In response to a number of Oversight Board recommendations, we continue to pursue policy development across a number of policy areas. Recent work includes conducting policy forums on the definition of "praise" of

Dangerous Organizations and Individuals, Violent and Graphic content in human rights abuse contexts, and refining "functional identification". As with previous policy development processes, this work includes consulting with external experts, research, including the relevant human rights analysis, and gathering inputs and insights from internal teams across Meta. We have also invited the Oversight Board to attend policy forums related to the scope of their case decisions and recommendations, and will continue to do so.

For a comprehensive list of all 28 recommendations in this category, see [Appendix B. Policy](#).

C.   Enforcement

We are providing updates for our work on 41 board recommendations that address our enforcement systems.

We want to highlight our progress on:

1. **Setting up a new crisis coordination team to provide dedicated Operations oversight before, during, and after global crises:** Following an assessment of the feasibility of setting up a new crisis coordination team to oversee all tactical execution of Operations support efforts preceding, for the duration of, and in the aftermath of crises, we are preparing to launch this model. This team will manage crisis coordination and provide dedicated Operations oversight throughout the lifecycle of imminent and emerging crises. It will be responsible for operational elements of critical events including high risk events and elections and will include crisis leads in our three main regions (APAC, EMEA and North America) to provide 24/7 coverage to crisis events.

2. **Expanding the types of content people are able to appeal directly to the Oversight Board:** Expanded users' ability to appeal eligible content decisions, including those made on escalation, to the Oversight Board. Previously, content decisions that we made during an internal escalation process were often not appealable because of separate pathways used for specialized contextual review. We expect this product update to significantly expand the volume and scope of content that users are able to appeal to the Oversight Board, and give the board the ability to take some of the most complex content moderation decisions that we have to contend with, further increasing the board's impact.

3. **Increasing transparency to people about the impact of the Oversight Board on content they have posted, reported, or appealed:** Following a series of two related recommendations, we completed the global roll-out of new, more specific messaging to

notify people of a change to the status of their content or content they have reported, because of an appeal to the Oversight Board. This includes providing secondary notifications should the status of that content change again. Because content that is the Oversight Board shortlists for review—whether it officially selects that content for a decision or not—is automatically double-checked by our review teams, the board's involvement often changes an enforcement decision even when it is not the subject of a full board case decision. This update will provide new visibility to people about how the board influences enforcement across our platforms.

4. **Establishing new protocols for time-bound inclusion on cross-check entity lists**: Following the board's recommendations around the importance of regular audits for entities included in our cross-check programs, we established protocols for applying certain cross-check tags for a short-term period in cases where the increase in risk is episodic or a shorter time period is otherwise more appropriate. This is particularly helpful for human rights and related entities who may become eligible for additional review during time-bound but significant world events.

5. **Implementing the board's recommendation to allow a phrase in the context of political protests in Iran and run sweeps to reverse strikes made previously because of the use of this phrase:** We have fully implemented the board's recommendation to allow the phrase "marg bar Khamenei" in the context of the ongoing protests in Iran. We have also run sweeps to identify previous enforcement actions and reverse strikes made on the basis of this type of content and will continue to pursue further reversals where they are applicable and feasible.

For a comprehensive list of all 41 recommendations in this category, see [Appendix C. Enforcement](#).

III. Appendix

How to Read This Appendix

The board recommendations in this appendix are categorized by implementation commitment level and current status of implementation, as detailed in page 13 under [How to Read This Update](#). For each recommendation, we include the following information:

- **Oversight Board Recommendation**: The board recommendation we are sharing a progress update on, along with any accompanying recommendations. As noted previously, when the board has issued similar recommendations across multiple cases, we combine these recommendations in progress updates.
- **Previous Commitment**: The implementation commitment level indicated in the last Quarterly Update, or the 60-day response to the board, whichever was more recent.
- **Updated Commitment**: The current implementation commitment level based on updated assessments or additional input from the board in the form of relevant recommendations, content decisions, or guidance from the Implementation Working Group.
- **Current Status**: The current status of our implementation work.
- **May 2023 Update**: An overview of our recent progress, challenges, considerations, and next steps for our work for each recommendation.

Appendix A. Transparency

Transparency Recommendations
Oversight Board Recommendation: *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.* (*Support of Abdullah Öcalan, Founder of the PKK #11* (along with *Al Jazeera Post on Tensions Between Israel and Palestine #4* and *Veiled Threat of Violence Based on Lyrics from a Drill Rap Song #6[4]*))

Previous Commitment	Implementing Fully
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	In our last Quarterly Update, we affirmed our commitment that we are working on increasing transparency around government requests. As we indicated previously, this includes government requests containing content that we review and may remove under our Community Standards. The objective is to produce government takedown request metrics in the most efficient manner given ongoing challenges including confidentiality obligations, data logging, and taxonomy gaps from internal systems. At this time, this continues to be a large, complex project requiring infrastructure and process investments as we want to provide accurate and robust details where possible. We will provide updates on this work in future Quarterly Updates.

Oversight Board Recommendation: *Meta should provide information about its implementation work in its quarterly reports on the Board. Additionally, Meta should convene a biannual meeting of high-level responsible officials to brief the Board on its work to implement the policy advisory opinion recommendations.* (*PAO on Meta's Cross-Check Policies #0*)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As committed in our initial response, we will continue to report Meta's implementation work towards cross-check in the quarterly report, such as we have done in this report covering all 32 recommendations. We will convene a meeting of key stakeholders to brief the board on the recommendation implementation status for this policy advisory opinion (PAO). The first of these will be scheduled in Q3 2023 and subsequently on a bi-annual cadence onwards. We will provide an update on the meeting in a future Quarterly Update.

[4] The board issued a similar recommendation in recommendation #4 in the Al Jazeera Post on Tensions Between Israel and Palestine case and recommendation #6 in the Veiled Threat of Violence Based on Lyrics from a Drill Rap Song case. We are tracking the progress of our work in response to these recommendations as part of our response to recommendation #11 in the Support of Abdullah Öcalan case.

Oversight Board Recommendation: *Meta should strengthen its engagement with civil society for the purposes of list creation and nomination. Users and trusted civil society organizations should be able to nominate others that meet the criteria. This is particularly urgent in countries where the company's limited presence does not allow it to identify candidates for inclusion independently.*

([PAO on Meta's Cross-Check Policies #7](#))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	In our March 6th, 2023 [response](#), we reported that we plan to engage a variety of organizations across the company & external stakeholders (i.e. Human Rights, Civil Rights, Trusted Partners program, external civil society organizations) to inform the criteria we use to identify public interest entities for cross-check lists. While we have not made substantial progress on this effort due to the recency of the recommendation, we are in the process of identifying ways to strengthen the relationship between the Trusted Partner program & the Early Response Secondary Review governance program. As we work to lay the foundation in this space, we foresee deeper collaboration between the Trusted Partner and ERSR governance program being pivotal to ensuring a more equitable & holistic cross-check program.

Oversight Board Recommendation: *Meta should use specialized teams, independent from political or economic influence, including from Meta's public policy teams, to evaluate entities for list inclusion. To ensure criteria are met, specialized staff, with the benefit of local input, should ensure objective application of inclusion criteria.*

([PAO on Meta's Cross-Check Policies #8](#))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	In our March 6th, 2023 [response](#), we shared that Governance responsibilities for the Early Response Secondary Review (ERSR) list currently sit within our Global Operations (GO) organization. We also explained that these responsibilities include assessing eligibility for inclusion in three of the six categories of ERSR lists and conducting list audits. As the primary auditing team within GO continues to mature the ERSR auditing process, we have begun partnering with our specialized regional expert teams within the GO organization in an effort to expand the ERSR auditing workflow to them. We will continue to report on our progress in future Quarterly Updates.

Oversight Board Recommendation: *Meta should require that more than one employee be involved in the final process of adding new entities to any lists for false positive mistake-prevention systems. These people should work on different but related teams.*

([PAO on Meta's Cross-Check Policies #9](#))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	We are committed to exploring a host of efforts designated toward improving the integrity of our mistake prevention lists, such as through regular list audits and quality checks. However, doing so requires substantially maturing the governance processes for the Early Response Secondary Review (ERSR) list, and we are still in the early stages of this work. Once we fully onboard our specialized regional teams into the program, we aim to incorporate comprehensive quality checks in the ERSR auditing process – as explained in our response to PAO on Meta's Cross-Check Policies #8. While this remains a long-term goal, we will continue to provide updates on this recommendation in future Quarterly Updates.

Oversight Board Recommendation: Meta should establish clear criteria for removal. One criterion should be the amount of violating content posted by the entity. Disqualifications should be based on a transparent strike system, in which users are warned that continued violation may lead to removal from the system and or Meta's platforms. Users should have the opportunity to appeal such strikes through a fair and easily accessible process. (*PAO on Meta's Cross-Check Policies #10*)	

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	As stated in our March 6th, 2023 response, all entities on Facebook and Instagram, including those whose content is removed through cross-check, are subject to a strike threshold and are disabled once that threshold is reached. Additionally, having previous violations does not negate the potential for cases that are at higher risk for mistakes or where the potential impact of a mistake is especially severe for the community. For instance, there may be entities penalized for posting borderline violating content or content that would be permissible on the platform given certain contexts. Such signals prompt us to review our existing Community Standards and make necessary changes, to reduce overenforcement at scale. Thus, we do not feel the number of violations alone is enough reason to remove an entity from ERSR. With that in mind, we are committed to using the number of violations an entity has incurred as a signal that an entity should be prioritized for audit more quickly, and re-assessing whether they still meet our eligibility criteria. An example of this would be our current review model of cross-checked Groups or Pages. When such entities hit the standard strike limit, they are enqueued to the team that reviews ERSR content, and are quickly reviewed for potential removal from the platform, as opposed to being holistically reviewed only during regularly scheduled audits. We are also working to expand appeals availability to cross-checked entities as part of a broader company-wide appeals effort, in line with our previous commitments to the Oversight Board. Our timeline for this expansion will align with Meta's broader compliance

	requirements for upcoming regulations (e.g. EU DSA). We hope to have a substantive update on this by the end of Q2 2023, and will share updates on our progress in future Quarterly Updates.
Oversight Board Recommendation: Meta should conduct periodic audits to ensure that entities benefitting from automatic bars to enforcement ('technical corrections') meet all criteria for inclusion. At least two teams with separate reporting structures should participate in these audits to provide for cross-team vetting. (*PAO on Meta's Cross-Check Policies #22*)	
Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	In line with our established governance for automatic bars to enforcement ('technical corrections'), we are continuing to establish manual and automated audit processes for entities included in these lists. We have already established both strict criteria for inclusion in the list and a strict renewal process that requires an audit to take place through requesting teams, either manually or through automated processes. We currently require Operations and Legal/Policy teams with different reporting structures to approve and renew a technical correction. As shared in our initial response, we will continue to improve our automated and manual auditing processes, and will report on our progress in a future Quarterly Update.
Oversight Board Recommendation: Meta should conduct periodic multi-team audits to proactively and periodically search for unexpected or unintentional bars to enforcement that may result from system error. (*PAO on Meta's Cross-Check Policies #23*)	
Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As stated in our initial response, we are committed to conducting periodic audits and publishing a report on these efforts in the future. We are still in the early stages of conducting these audits, and in the interim we have a series of structures and processes that we are using to remove unexpected exemptions and control existing ones. General Secondary Review (GSR) is a machine learning ranking based system that escalates potential enforcement mistakes for secondary review for all 3 billion users across Meta's platforms. We also have stringent oversight of our guaranteed review segment "Early Response Secondary Review (ERSR) as well as the automatic bars to enforcement system known as "technical corrections". As we continue to plan these audits, we will also continue to identify and deprecate systems that cause unexpected bars to enforcement. We will provide an update on the progress of this work in a future Quarterly Update.

Oversight Board Recommendation: *Meta should conduct periodic multi-team audits to proactively and periodically search for unexpected or unintentional bars to enforcement that may result from system error.* *([PAO on Meta's Cross-Check Policies #28](#))*	
Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In progress
May 2023 Update	As mentioned in our March 6th, 2023 [response](#), implementing this recommendation in full is a long-term goal. Our ability to fully implement this recommendation is dependent on other efforts to improve measurements and further mature the review system overall. At this time, there is no substantial update other than the groundwork we're laying in expanding and enhancing our metrics reporting, such as a more nuanced understanding of overturn rates, but we intend to share updates on this recommendation in future Quarterly Updates.

Oversight Board Recommendation: *Meta should publish regular transparency reporting focused specifically on delayed enforcement of false-positive prevention systems. Reports should contain data that permits users and the public to understand how these programs function and what their consequences on public discourse may be. At minimum, the Board recommends Meta include:* A. *Overturn rates for false positive mistake-prevention systems, disaggregated according to different factors.* B. *The total number and percentage of escalation-only policies applied due to false positive mistake-prevention programs relative to total enforcement decisions.* C. *Average and median time to final decision for content subject to false-positive mistake prevention programs, disaggregated by country and language.* D. *Aggregate data regarding any lists used for mistake-prevention programs, including the type of entity and region.* E. *Rate of erroneous removals (false positives) versus all reviewed content, including the total amount of harm generated by these false positives measured as the predicted total views on the content (i.e., overenforcement)* F. *Rate of erroneous keep-up decisions (false negatives) on content, including the total amount of harm generated by these false positives, measured as the sum of views the content accrued (i.e., underenforcement)* *([PAO on Meta's Cross-Check Policies #30](#))*	
Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress
May 2023 Update	In our March 6th, 2023 [response](#), we shared that we will begin tracking and determining what information can be shared publicly in an annual report aimed to increase transparency around our cross-check program. Since our initial response, we have assembled a

	cross-functional taskforce to align on key impact metrics that can be shared publicly. This team has begun the work of scoping out the workstream, collecting initial data, and identifying whether there are any additional metrics to measure. We hope to complete this foundational work by Q3 2023, at which point we will begin the long-term work of compiling said metrics and publishing them as soon as readily available. As explained in our response to PAO on Meta's Cross-Check Policies #28, our ability to implement this recommendation is dependent on multi-part efforts to improve measurements and further mature the review system overall, as well as work to expand our transparency efforts more broadly – consistent with regulatory requirements and existing transparency roadmaps. This remains a long-term effort, but we will report on our progress in future Quarterly Updates.

Oversight Board Recommendation: *Meta should institute a pathway for external researchers to gain access to non-public data about false-positive mistake-prevention programs that would allow them to understand the program more fully through public-interest investigations and provide their own recommendations for improvement. The Board understands that data privacy concerns should require stringent vetting and data aggregation.*

(PAO on Meta's Cross-Check Policies #32)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In progress
May 2023 Update	As shared in our initial March 6th response, we are currently evaluating possible solutions to expand external research initiatives to include false-positive mistake prevention programs. Since our initial response, the team has begun project scoping to understand the work to align on how best to approach implementation. This recommendation is complex, and we will be prioritizing user privacy and security throughout. We are continuing to work on this effort in tandem with our regulatory obligations, and will provide an update on our progress in a future Quarterly Update.

Oversight Board Recommendation: *Meta should announce all scaled allowances that it issues, their duration and notice of their expiry, in order to give people who use its platforms notice of policy changes allowing certain expression, alongside comprehensive data on the number of "scaled" and "narrow" allowances granted. The Board will consider this recommendation implemented when Meta demonstrates regular and comprehensive disclosures to the Board.*

(A Cartoon Depicting Iran's Supreme Leader Ayatollah Ali Khamenei #5)

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress
May 2023 Update	In our 60-day response and more recently in our initial response to A Call for the Donation of Pharmaceutical Drugs to Sri Lanka #4, we committed to publishing the total number of

	scaled allowances issued on a yearly basis. We expect to share this number by the end of 2023, and will provide an update on our progress in the next Quarterly Update.

*Oversight Board Recommendation:*The public explanation of the newsworthiness allowance in the Transparency Centre should (i) explain that newsworthiness allowances can either be scaled or narrow; and (ii) provide the criteria that Meta uses to determine when to scale newsworthiness allowances. The Board will consider this recommendation to be implemented when Meta updates the publicly available explanation of newsworthiness and issues Transparency Reports that include sufficiently detailed information about all applied allowances.

(A Cartoon Depicting Iran's Supreme Leader Ayatollah Ali Khamenei #6)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In progress
May 2023 Update	We are in the process of drafting language to include as part of the Approach to Newsworthy Content page on the Transparency Center that details the scaled and narrow allowances as well as provides more details about how scaled newsworthy allowance decisions are made. We expect to complete this work by Q4 2023, and we will report on our progress in a future Quarterly Update.

Oversight Board Recommendation: Meta should provide a public explanation of the automatic prioritization and closure of appeals, including the criteria for both prioritization and closure. The Board will consider this recommendation implemented when Meta publishes this information in the Transparency Centre.

(A Cartoon Depicting Iran's Supreme Leader Ayatollah Ali Khamenei #7)

Previous Commitment	Assessing Feasibility
Updated Commitment	No Further Action
Current Status	No Further Updates
May 2023 Update	As explained in our initial response, our current content review prioritization process across all of our products is publicly outlined in our Transparency Center. On this page, we explain that we primarily consider severity, virality, and likelihood of violation when determining which content our human review teams should prioritize for review. Since Q1 2022, we have undergone a multi-stage process to identify key drivers of trust in appeals to improve their overall effectiveness. Given this, our more granular progress in automatic prioritization and closure of appeals is newly developed and quickly transforming. When considering whether to add additional granularity to our Transparency Center page reflecting these changes, we came to the conclusion that the publication of the new system at this stage would be misleading given the fact that the criteria involved are subject to evolve, often very quickly. In the spirit of transparency, we will be sharing our full assessment of considerations with the Oversight Board. While we will have no further updates on this specific recommendation, we are continuing to refine our automatic prioritization and ranking of

	appeals throughout 2023, and will be providing further updates on the development of these new processes in our responses to [Post Requesting Advice on Pharmaceutical Drugs #2](#) and [Armenian People and the Armenian Genocide #5](#). By continuing to publicly report on each iteration of building, testing, and strengthening automatic prioritization and closure of appeals in our Quarterly Updates, we hope to achieve the spirit of this recommendation by furthering transparency around the process without risking an inaccurate update to the Transparency Center.

Oversight Board Recommendation: In line with the Board's recommendations five and six in the "Iran protest slogan" case (2022-013-FB-UA), the Board specifies that Meta should publicly share aggregated data in its Transparency Centre about the "spirit of the policy" allowances issued, including the number of instances in which they were issued, and the regions and/or languages affected. Meta should keep this information updated as new "spirit of the policy" allowances are issued. The Board will consider this recommendation implemented when Meta makes this information publicly available in the Transparency Centre.

[*(A Call for the Donation of Pharmaceutical Drugs to Sri Lanka #4)*](#)

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	We have not made significant progress on this recommendation since our May 8th, 2023 [response](#), where we shared that we will publish the total number of scaled allowances on our Transparency Center – in line with our response to [Cartoon Depicting Iran's Supreme Leader Ayatollah Ali Khamenei #5](#). We will not, however, publish data about the regions or languages involved in these allowances, due to sensitivities around regional data sharing as well as the significant operational lift required to execute this at a time when any surplus capacity is required for necessary safety functions. In light of this, we will be moving this recommendation to the long-term implementation category going forward, and expect to provide an update on our progress next year.

Appendix B. Policy

Policy Clarity & Accessibility Recommendations	
Oversight Board Recommendation: Revise the Instagram Community Guidelines to specify that female nipples can be shown to raise breast cancer awareness and clarify that where there are inconsistencies between the [Instagram] Community Guidelines and the [Facebook] Community Standards, the latter take precedence. *(Breast Cancer Symptoms & Nudity Recommendation #2 (along with Breast Cancer Symptoms & Nudity Recommendation #7, Support of Abdullah Öcalan Recommendation #10, Post Discussing a Substance with Psychoactive Properties Recommendation #1, and PAO on Sharing Residential Information #9)[5])*	
Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress
May 2023 Update	As described in previous Quarterly Updates, we are continuing work to unify the Instagram Community Guidelines and Facebook Community Standards. As described to the board in a recent briefing, despite regulatory and capacity challenges, our legal, policy, and product teams are continuing foundational work to achieve parity across user experiences and increase learning opportunities about the rules that apply on both Facebook and Instagram. As described in the same briefing to the board's Implementation Working Group, as we continue to build out our unified Community Standards, we are also working to provide people on Instagram with increased transparency and remediation processes for Community Guidelines enforcements. Account Status is a new central hub where people on Instagram can understand the enforcement and restrictions affecting their account in a simple, actionable manner. This product was launched globally in March 2023, in all available languages. While we have had to shift the prioritization of this effort on roadmaps to accommodate complementary foundational product work for urgent regulatory compliance priorities, this work remains a key priority across teams and we will continue to report on our progress in the next Quarterly Update.
Oversight Board Recommendation: 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.	

---

[5] The board issued similar recommendations in recommendation #7 in the Breast Cancer Symptoms & Nudity case, recommendation #10 in the Support of Abdullah Öcalan case, recommendation #1 in the Post Discussing a Substance with Psychoactive Properties case and recommendation #9 in the PAO on Sharing Residential Information case. We are tracking the progress of our work in response to these recommendation as part of our response to recommendation #2 in the Breast Cancer Symptoms & Nudity case.

(*Al Jazeera Post on Tensions Between Israel and Palestine #1* (along with *News Outlet Discussing the Taliban Government in Afghanistan #3*) [6])	
Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress
May 2023 Update	We soon expect to share more details about the policy development and enforcement related to "praise" in our Dangerous Individuals and Organizations Community Standard. As part of this announcement and work, the Community Standards will further clarify the exceptions for neutral discussion, condemnations, and news reporting. We are still assessing the tradeoffs to publishing illustrative examples of non-violating content, and more details about this change and others will be posted to our newsroom in H2 2023.

Oversight Board Recommendation: Meta should define graphic depiction and sexualization in the Child Sexual Exploitation, Nudity and Abuse Community Standard. Meta should make clear that not all explicit language constitutes graphic depiction or sexualization and explain the difference between legal, clinical or medical terms and graphic content. Meta should also provide a clarification for distinguishing child sexual exploitation and reporting on child sexual exploitation. The Board will consider the recommendation implemented when language defining key terms and the distinction has been added to the Community Standard. (*Post Describing Sexual Violence Against Minors #1*)	
Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As part of this definitional work, we are consulting with external experts in adolescent health; child rights and victim advocates; and experts in journalistic standards to ensure that the policy line and any changes properly balance considerations of safety and expression in the form of legal, clinical, or medical terms as well as in reporting context. We are considering their inputs and making final refinements to this language, but expect to have a more robust update in the next Quarterly Update.

Oversight Board Recommendation: Meta should undergo a policy development process, including as a discussion in the Policy Forum, to determine whether and how to incorporate a prohibition on functional identification of child victims of sexual violence in its Community Standards. This process should include stakeholder and expert engagement on functional identification and the rights of the child. The Board will consider this recommendation implemented when Meta publishes the minutes of the Product Policy Forum where this is discussed. (*Post Describing Sexual Violence Against Minors #2*)	

---

[6] The board issued similar recommendations in recommendation #3 in the Post from a News Outlet Discussing the Taliban Government in Afghanistan case. We are tracking the progress of our work in response to that recommendation as part of our response to recommendation #1 in the Al Jazeera Post on Tensions Between Israel and Palestine case.

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As mentioned in previous Quarterly Updates, we have expanded the scope of this work beyond functional identification of child victims of sexual assault and are exploring a definition that spans broader policy areas. We are finalizing research and external engagement work on this broadened policy development and are in the process of preparing options that we will consider for final implementation. We expect to have a more robust update in the next Quarterly Update.

Oversight Board Recommendation: Meta should remove the exception that allows the sharing of private residential information (both images that currently fulfill the Privacy Violations policy's criteria for takedown and 10 addresses) when considered "publicly available". This means Meta would no longer allow otherwise violating content on Facebook and on Instagram if "published by at least five news outlets" or if it contains residential addresses or imagery from financial records or statements of an organization, court records, professional and business licenses, sex offender registries or press releases from government agencies, or law enforcement. The Board will consider this implemented when Meta modifies its Internal Implementation Standards and its content policies.	
(PAO on Sharing Residential Information #1) (along with PAO on Sharing Residential Information #3[7])	

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	We are planning to update our policy guidance, in which we removed residential information from the publicly available carveout. Our plan is to add this carveout to the Community Standards and reviewer guidance, and further define when a property is 'the focus of news story' in guidance. We look forward to providing a more robust update and completing this recommendation in the next Quarterly Report.

Oversight Board Recommendation: Allowing the organization of protests at publicly owned official residences. Meta should allow the publication of addresses and imagery of official residences provided to high-ranking government officials, such as heads of state, heads of federal or local government, ambassadors and consuls. The Board will consider this implemented when Meta modifies its content policies.	
(PAO on Sharing Residential Information #4)	
Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully

---

[7] The board issued a similar recommendation in recommendation #3 in the PAO on Sharing Residential Information case. We are tracking the progress of our work in response to this recommendation as part of our response to recommendation #1 in the PAO on Sharing Residential Information.

Current Status	In Progress
May 2023 Update	We are continuing to refine our policy guidance and expect to provide further details in the upcoming months. These changes are continuing to require time to implement as we expanded scope from the board's initial recommendations as we worked to implement this recommendation. We expect to have an update in the Transparency Center in the upcoming months.

Oversight Board Recommendation: Meta should better explain, in the text of Facebook's Privacy Violations policy, when disclosing the city where a residence is located will suffice for the content to be removed, and when disclosing its neighborhood would be required for the same matter (e.g., by specifically referencing the population threshold at which sharing only the city as part of the content will no longer be considered violating). The Board will consider this implemented when Meta modifies its content policies.

(PAO on Sharing Residential Information #7)

Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress
May 2023 Update	We are continuing to refine our policy guidance, including modifying guidance and training materials to outline criteria for when disclosing the city or the neighborhood is sufficient for identifying where a residence is located, and expect to provide further details in the upcoming months. These changes are continuing to require time to implement as we expanded scope from the board's initial recommendations as we worked to implement this recommendation. We will provide an update on our progress in the next Quarterly Update.

Oversight Board Recommendation: Meta should explain, in the text of Facebook's Privacy Violations policy, its criteria for assessing whether the resident is sufficiently identified in the content. The Board will consider this implemented when Meta modifies its content policies.

(PAO on Sharing Residential Information #8)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	We are continuing to refine our policy guidance, including modifying guidance and training materials to outline criteria for when a resident is sufficiently identified, and expect to provide further details in the upcoming months. These changes are continuing to require time to implement as we expanded scope from the board's initial recommendations as we worked to implement this recommendation. We will provide an update on our progress in the next Quarterly Update

Oversight Board Recommendation: Meta should amend the Violent and Graphic Content Community Standard to allow videos of people or dead bodies when shared for the purpose of raising awareness

of or documenting human rights abuses. This content should be allowed with a warning screen so that people are aware that content may be disturbing. The Board will consider this recommendation implemented when Meta updates the Community Standard.

([Video Depicting a Civilian Victim of Violence in Sudan #1](#) (along with [Video Depicting a Civilian Victim of Violence in Sudan #2](#)[8])

Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress
May 2023 Update	We are finalizing work on the policy development that considers the balance between documenting and raising awareness about human rights abuses while protecting the dignity of the person depicted in the imagery as well as considering the potential harm related to exposing users to this type of graphic content. As part of this process, we are completing research and wrapping up conversations with a globally diverse group of external experts. Following this rigorous policy development process, we will share more information and updates about any potential changes in the [Community Standards](#).

Oversight Board Recommendation: Meta should clarify that for content to be removed as a "veiled threat" under the Violence and Incitement Community Standard, one primary and one secondary signal is required. The list of signals should be divided between primary and secondary signals, in line with the internal Implementation Standards. This will make Meta's content policy in this area easier to understand, particularly for those reporting content as potentially violating. The Board will consider this recommendation implemented when the language in the Violence and Incitement Community Standard has been updated.

([Veiled Threat of Violence Based on Lyrics from a Drill Rap Song #2](#))

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As mentioned in our [initial response](#), we are currently considering policy development that seeks to clarify the usage of primary and secondary signals in the veiled threats assessment. We expect to share these changes in our [Violence and Incitement](#) Community Standard and will provide an update on our progress in a future Quarterly Update. In the interim, further details of our policy on [veiled threats](#) can also be found in the overview of our 2020 [Policy Forum](#), in which we explored and established a framework for assessing veiled threats.

Oversight Board Recommendation: Meta should add to the public-facing language of its Violence and Incitement Community Standard that the company interprets the policy to allow content containing

---

[8] The board issued a similar recommendation in recommendation #2 in the [Video Depicting a Civilian Victim of Violence in Sudan](#) case. We are tracking the progress of our work in response to this recommendation as part of our response to recommendation #1 in the Video Depicting a Civilian Victim of Violence in Sudan case.

statements with "neutral reference to a potential outcome of an action or an advisory warning" and content that "condemns or raises awareness of violent threats". The Board expects that this recommendation, if implemented, will require Meta to update the public-facing language of the Violence and Incitement policy to reflect these inclusions.

([*Individual Killed in Ukraine During the Russian Invasion #1*](#))

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	We are currently in the process of refining aspects of our [Violence and Incitement](#) Policy, and as part of this work are working to clarify how a "neutral reference to a potential outcome of an action or an advisory warning" may be interpreted. As shared in our [last Quarterly Update](#), under our Violence and Incitement policy, we remove language that incites or facilitates serious violence. We recognize, however, that people sometimes share content that references or contains violent threats in order to condemn or raise awareness of those threats, and therefore we allow this type of content on our platforms. We also allow people to post content that includes a warning about a potential action, as we believe that users should be empowered to voice concerns of this kind. Although this clarifying work is not yet complete, we will provide an update on our progress in an upcoming Quarterly Update.

__Oversight Board Recommendation:__ Meta should add to the public-facing language of its Violent and Graphic Content Community Standard detail from its internal guidelines about how the company determines whether an image "shows the violent death of a person or people by accident or murder". The Board expects that this recommendation, if implemented, will require Meta to update the public-facing language of the Violent and Graphic Content Community Standard to reflect this inclusion.

([*Individual Killed in Ukraine During the Russian Invasion #2*](#))

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	We are currently in the process of updating our [Violent and Graphic Content Community Standard](#) to clarify how we determine if content depicts "the violent death of a person or people by accident or murder." As stated in our [initial response](#), we often use certain indicators within imagery, such as the presence of blood or injuries on the victim, to help lead to a reasonable conclusion that the person suffered a violent death. We are continuing refinements on this, and expect to provide an update in future Quarterly Reports.

__Oversight Board Recommendation:__ Meta should review the public-facing language in the Violent and Graphic Content policy to ensure that it is better aligned with the company's internal guidance on how the policy is to be enforced. The Board will consider this recommendation implemented when the

policy has been updated with a definition and examples, in the same way as Meta explains concepts such as "praise" in the Dangerous Individuals and Organisations policy.

(*[Aftermath of an Attack on a Church in Nigeria #1](#)*)

Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress
May 2023 Update	We are still in the process of reviewing language for our [Violent and Graphic Content](#) policy and exploring ways to provide further clarity with definitions in the external policy. This work is being conducted in conjunction with broader efforts to review language – incorporating further details and clarification in a way that improves understanding but does not leave our platforms more susceptible to abuse by bad actors. We will provide an update on our progress in a future Quarterly Update.

Oversight Board Recommendation: *Meta should include an exception to the Adult Sexual Exploitation Community Standard for depictions of non-consensual sexual touching, where, based on a contextual analysis, Meta judges that the content is shared to raise awareness, the victim is not identifiable, the content does not involve nudity and is not shared in a sensationalized context, thus entailing minimal risks of harm for the victim. This exception should be applied at escalation only. The Board will consider this recommendation implemented when the text of the Adult Sexual Exploitation Community Standard has been changed.*

(*[Video Depicting Sexual Harassment in India #1](#)*)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	We have finalized and are launching changes to our [Adult Sexual Exploitation Policy](#) to allow depictions of non-consensual sexual touching with a warning screen where the content is shared to raise awareness, the victim is not identifiable, the content does not involve overt nudity and explicit sexual activity, and it is not shared in a sensationalized context. This will be context-specific and applied on escalation only.

Oversight Board Recommendation: *Meta should establish clear and public criteria for list-based mistake-prevention eligibility. These criteria should differentiate between users who merit additional protection from a human rights perspective and those included for business reasons.*

(*[PAO on Meta's Cross-Check Policies #4](#)*)

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress

May 2023 Update	Since we began externally publishing the categories of entities we enroll in the Early Response Secondary Review (ERSR) system on our [Transparency Center](#), we have been exploring possibilities that allow us to further break down our current broad categories into more specific, sub-categories - to better distinguish between those groups of users who have been added because of a business relationship and those who have been added on account of human rights interests. We would, however, like to emphasize as mentioned in our [initial response](#), that we have no current plans to publish all of the more specific criteria we use to determine whether a user qualifies for membership in any of the ERSR categories or potential sub-categories, because doing so could make the system more vulnerable to manipulation and coordinated inauthentic behavior. Maintaining an updated set of publicly available criteria would also be difficult as we will continue to refine and iterate on the criteria for inclusion in ERSR based on the challenges that surface on our platforms. We will provide updates on this work in future Quarterly Updates.

Oversight Board Recommendation: *Meta should provide basic information in its Transparency Center regarding the functioning of any mistake-prevention system it uses that identifies entities or users for additional protections.*

[(PAO on Meta's Cross-Check Policies #31)](#)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	Complete
May 2023 Update	As stated in our [initial response](#), we recently published information on how we detect and enforce violations in our [Transparency Center](#), including an overview of our approach to accurately reviewing high-impact content through cross-check. While we cannot publish the full details of these interventions, due to the high risk of adversarial behavior, we have expanded upon the existing information by including further details around our mistake prevention systems in the "[Detecting Violations](#)" section of our Transparency Center. This recommendation is now complete and will have no further updates.

Oversight Board Recommendation: *Meta's Community Standards should accurately reflect its policies. To better inform users of the types of statements that are prohibited, Meta should amend the Violence and Incitement Community Standard to (i) explain that rhetorical threats such as "death to X" statements are generally permitted, except when the target of the threat is a high-risk person; (ii) include an illustrative list of high-risk persons, explaining that they may include heads of state; (iii) provide criteria for when threatening statements directed at heads of state are permitted to protect clearly rhetorical political speech in protest contexts that does not incite to violence, taking language and context into account, in accordance with the principles outlined in this decision. The Board will consider this recommendation to be implemented when the public-facing language of the Violence and Incitement Community Standard reflects the proposed change, and when Meta shares internal guidelines with the Board that are consistent with the public-facing policy.*

[(A Cartoon Depicting Iran's Supreme Leader Ayatollah Ali Khamenei #1](#))

Previous Commitment	Assessing Feasibility

Updated Commitment	Assessing Feasibility
Current Status	In progress
May 2023 Update	In relation to this recommendation, we are pursuing policy development related to our [Violence and Incitement](#) policy including examining calls for death and high risk persons. As part of this work, we are also scoping out and considering ways to strike a better balance between violent speech and political expression, particularly rhetorical political speech in protest contexts. We expect that this policy development, given its complexity, will take some time to fully complete and implement and therefore expect this recommendation to remain in progress for upcoming Quarterly Updates.

Oversight Board Recommendation: In order to treat all users fairly and provide moderators and the public with a workable standard on nudity, Meta should define clear, objective, rights-respecting criteria to govern the entirety of its Adult Nudity and Sexual Activity policy, ensuring treatment of all people that is consistent with international human rights standards, including without discrimination on the basis of sex or gender identity. Meta should first conduct a comprehensive human rights impact assessment to review the implications of the adoption of such criteria, which includes broadly inclusive stakeholder engagement across diverse ideological, geographic and cultural contexts. To the degree that this assessment should identify any potential harms, implementation of the new policy should include a mitigation plan for addressing them.

([Gender Affirming Surgery Bundle #1](#))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress
May 2023 Update	As mentioned in the initial [60 day response](#) for this recommendation, we are planning to conduct a focused, internal human rights due diligence related to the Adult Nudity and Sexual Activity policy. This work will include usage of the UNGP's [salience criteria](#) and will tie in the various updates that have previously been made or considered for this policy given results may inform future policy development processes. We are in the initial scoping stages of this work.

Oversight Board Recommendation: In order to provide greater clarity to users, Meta should provide users with more explanation of what constitutes an "offer or ask" for sex (including links to third party websites) and what constitute sexually suggestive poses in the public Community Standards. The Board will consider this recommendation implemented when an explanation of these terms with examples is added to the Sexual Solicitation Community Standard.

([Gender Affirming Surgery Bundle #2](#))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress

May 2023 Update	As outlined in our initial [60 day response](#) for this recommendation, we will be conducting policy development that may change or refine our Sexual Solicitation policy to better explain what constitutes an "offer or ask" and "sexually suggestive poses." With that said, we will not not list out examples in our Community Standards as that would require us to surface the type of harmful and negative content the policy seeks to protect against. We are currently in the process of scoping out refinements for this policy line and expect further updates in upcoming Quarterly Updates.

Oversight Board Recommendation: To provide more clarity to users, Meta should explain in the landing page of the Community Standards, in the same way the company does with the newsworthiness allowance, that allowances to the Community Standards may be made when their rationale, and Meta's values, demand a different outcome than a strict reading of the rules. The company should include a link to a Transparency Centre page which provides information about the "spirit of the policy" allowance. The Board will consider this recommendation implemented when an explanation is added to the Community Standards.

[*(A Call for the Donation of Pharmaceutical Drugs to Sri Lanka #1)*](#)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	We have not made significant progress on this recommendation since our May 8th, 2023 [response](#), where we shared that we will add a new page to the Transparency Center detailing our approach to the spirit of the policy allowances. As shared publicly and with the board, we occasionally issue allowances for content that may violate the exact letter of our policies but, upon careful consultation and consideration of important context, does not violate the "spirit" of that policy. We will add a new page to the Transparency Center explaining our approach to these spirit of the policy allowances by the end of the year, and link the page in the introduction to our Community Standards. We will provide an update on our progress in a future Quarterly Update.

Oversight Board Recommendation: In line with the Board's recommendations five and six in the "Iran protest slogan" case (2022-013-FB-UA), the Board specifies that Meta should publish information about the "spirit of the policy" allowance in its Transparency Centre, similar to the information it has published on the newsworthiness allowance. In the Transparency Centre, Meta should: (i) explain that "spirit of the policy" allowances can be either scaled or narrow; (ii) publicize examples of content which benefited from this allowance; (iii) provide criteria Meta uses to determine when to scale "spirit of the policy" allowances; and (iv) include a list of all "spirit of the policy" allowances Meta has issued at scale in the past three years with explanations of why Meta decided to issue and terminate each of them. Meta should keep this list updated as new allowances are issued. The Board will consider this recommendation implemented when Meta makes this information publicly available in the Transparency Centre.

[*(A Call for the Donation of Pharmaceutical Drugs to Sri Lanka #3)*](#)

Previous Commitment	Implementing in Part

Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	We have not made significant progress on this recommendation since our May 8th, 2023 response, where we shared that we will add new information to our Transparency Center about how we make "spirit of the policy" decisions. This will include descriptions of both "narrow" allowances (which apply to a single piece of content) and "scaled" allowances (which apply more broadly). We expect to provide these updates by the end of this year and will provide an update on the status of this work in the next Quarterly Update.

Appendix C. Enforcement

Enforcement Recommendations

	Oversight Board Recommendation: *Facwbook should let users indicate in their appeal that their content falls into one of the exceptions to the Hate Speech policy.* (*Armenian People and the Armenian Genocide #4*, *(along with PAO on Sharing Residential Information #10 and PAO on Sharing Residential Information #14*)[9]*)*
Previous Commitment	Assessing Feasibility
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	As shared in our previous Quarterly Update, we have completed the feasibility assessment for this recommendation, and have begun testing product developments that will allow for users to indicate in their appeal whether their content might fall into a policy exception. Rather than providing an extensive list of every Hate Speech policy exception, (which could potentially discourage people from appealing due to the fact that not everyone is sufficiently familiar with our policy exceptions) our goal is to allow users to choose from broad options designed to indicate some of the intent behind their post. The planned messaging focuses on maximizing accessibility, approachability, and ease of use for people who use our platforms, and we hope to complete these tests this year. We will provide an update on the status of this ongoing work in the next Quarterly Update.
	Oversight Board Recommendation: *To improve the accuracy of Facebook's review in the appeals stage, the company should ensure appeals based on policy exceptions are prioritized for human review.* (*Armenian People and the Armenian Genocide #5*)
Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress
May 2023 Update	As explained in our response to Armenian People and the Armenian Genocide #4, we are currently working on enabling users to indicate that their content falls into a policy exception during an appeal. We are currently designing and testing product messaging allowing users to indicate some of the intent behind their post, which could be inferred as a possible policy exception by proxy. Once we have completed the implementation of that recommendation and aligned on the appropriate level of granularity that would be the most effective and most accessible for people who use our platforms, we will begin to evaluate

---

[9] The board issued similar recommendations in PAO on Sharing Residential Information #10 and PAO on Sharing Residential Information #14. We are tracking the progress of our work in response to these recommendations as part of our response to recommendation #4 in the Armenian People and the Armenian Genocide case.

	the tradeoffs of prioritizing appeals based on policy exceptions. We will provide an update on our progress in the next Quarterly Update.

Oversight Board Recommendation: *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).* (*Support of Abdullah Öcalan, Founder of the PKK #9*)	
Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	In 2022, we launched updated user notifications for situations in which we restrict access to content for users in a particular jurisdiction or jurisdictions on the basis of local law following a government request. In the first quarter of 2023, we updated the language of those notifications to include which government authority (in specific jurisdictions) sent the takedown request resulting in the restriction. These changes are an important part of our commitments under our Corporate Human Rights Policy and as a member of the Global Network Initiative. More details about our approach to content restrictions can be found in the Transparency Center. Our goal is to share more detail in user notifications for content restricted based on local laws. We note that we already notify people when their content is removed based on a violation of our Community Standards, regardless of the way that violation was identified. Work to specifically indicate in this notification when content removed for violating the Community Standards was reported by a government entity identifiable as such is ongoing. Our teams have identified some barriers for disclosure in particular countries that prevent user notice with the same level of transparency due to confidentiality obligations in those jurisdictions. We will provide updates on this work in future Quarterly Updates.

Oversight Board Recommendation: *Notify all users who reported content assessed as violating but left on the platform for public interest reasons that the newsworthiness allowance was applied to the post. The notice should link to the Transparency Center explanation of the newsworthiness allowance.* (*Post Depicting Protests in Colombia While Using a Slur #4* (along with *Video Depicting a Civilian Victim of Violence in Sudan #4*[10]))	
Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress

---

[10] The board issued a similar recommendation in recommendation #4 in the Video Depicting a Civilian Victim of Violence in Sudan case. We are tracking the progress of our work in response to this recommendation as part of our response to recommendation #4 in the Post Depicting Protests in Colombia While Using a Slur case.

May 2023 Update	We are continuing to prioritize the foundational work on post-report communications on Facebook and Instagram, and are increasing the scope and availability of more granular user messaging on our platforms in tandem with other compliance requirements. As explained in our last Quarterly Update, this foundational work is intended to increase the quality of notifications and make sure users feel supported by Meta. This work includes auditing the current user notifications content and increasing the percentage of reports that prompt follow up communications alongside granular messaging design. Once we have completed this foundational work, we will be able to assess the best way to expand these notifications to inform reporters when content assessed as violating is left on our platform because it is considered newsworthy. Since beginning this foundational work last year, we have successfully expanded the availability of reporter notifications and are aiming to have full coverage by the end of the year on Facebook and Instagram. We will continue to report on the status of this multi-stage effort in future Quarterly Updates.

Oversight Board Recommendation: Provide users with timely and accurate notice of action being taken on the content their appeal relates to. Where applicable, including in enforcement error cases like this one, the notice to the user should acknowledge that the action was a result of the Oversight Board's review process. Meta should share the user messaging sent when board actions impact content decisions appealed by users, to demonstrate it has complied with this recommendation.

(Depicting Indigenous Artwork and Discussing Residential Schools #1)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	Complete
May 2023 Update	In our last Quarterly Update, we shared that we currently notify all reporters about enforcement decisions following their initial reporting of a piece of content – including a secondary notification about further decisions should the original reporter appeal the initial review outcome. In February 2023, we expanded upon these secondary notifications by launching more specific messaging for incorrect outcomes in all available languages. These notifications are available globally on both Facebook and Instagram, and we now consider this recommendation complete. We will have no further updates on this recommendation.

Oversight Board Recommendation: Meta should study the consequences and trade-offs of implementing a dynamic prioritization system that orders appeals for human review, and consider whether the fact that an enforcement decision resulted in an account restriction should be a criterion within this system. The Board will consider this recommendation implemented when Meta shares the results of these investigations with the Board and in its quarterly Board transparency report.

(Post Requesting Advice on Pharmaceutical Drugs #2)

Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress

May 2023 Update	Since launching the first iteration of our new prioritization system in Q4 2022, we have continued our commitment to test, develop, and expand on the capabilities of this nascent system. The system currently orders appeals using a variety of indicators. As the system is evolving we will continue to update this work. As mentioned in our response to [Armenian People and the Armenian Genocide #5](), we are completing the foundational work that will enable us to understand and evaluate how we should prioritize appeals – including assessing the feasibility of ordering appeals based on the severity of enforcement decisions. We will provide an update on the progress of this multi-stage work in a future Quarterly Update.

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

([Video of an Edited Cartoon Depicting a Croatian City #2]())

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	Complete
May 2023 Update	As explained in our response to [Depicting Indigenous Artwork and Discussing Residential Schools #1](), in February 2023 we launched more specific messaging for when we notify people of a change to the status of their content because of an appeal to the Oversight Board. The messaging for reporters reads: "Because your case was being considered by the Oversight Board, [Facebook/Instagram] reviewed the [content] you reported again and decided to [remove/restore] it. The Oversight Board may still use the information you provided to issue recommendations that can improve [Facebook/Instagram]'s decisions on other cases like yours. Thanks again for coming to the Oversight Board. Reports like yours help hold [Facebook/Instagram] to account." These notifications are available globally on both Facebook and Instagram. We will share the efficacy of introducing this change to the public through our future regulatory disclosures and now consider this recommendation complete. We will have no further updates on this recommendation.

Oversight Board Recommendation: Meta should revise its internal Implementation Standards to
make clear that the "reporting" allowance in the Dangerous Individuals Organizations policy allows for
positive statements about designated entities as part of the reporting, and how to distinguish this
from prohibited "praise." The Known Questions document should be expanded to make clear the
importance of news reporting in situations of conflict or crisis and provide relevant examples, and that
this may include positive statements about designated entities like the reporting on the Taliban in this

case. The Board asks Meta to share the updated Implementation Standards with the Board within 60 days.

(*[Post from a News Outlet Discussing the Taliban Government in Afghanistan #4](#)*)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In progress
May 2023 Update	As mentioned in our response to [Al Jazeera Post on Tensions Between Israel and Palestine #1](#), this half, we expect to share more details about the policy development and enforcement related to "praise" in our [Dangerous Individuals and Organizations](#) Community Standard. As part of this announcement and work, the Community Standards will further clarify the exceptions for neutral discussion, condemnations, and news reporting. More details about this change and others will be posted to our newsroom in Q2 2023.

Oversight Board Recommendation: Meta should assess the accuracy of reviewers enforcing the reporting allowance under the Dangerous Individuals and Organizations policy in order to identify systemic issues causing enforcement errors. The Board asks Meta to inform the Board within 60 days of the detailed results of its review of this assessment, or accuracy assessments Meta already conducts for its Dangerous Individuals and Organizations policy, including how the results will inform improvements to enforcement operations, including for HIPO.

(*[Post from a News Outlet Discussing the Taliban Government in Afghanistan #5](#)*)

Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress
May 2023 Update	The successful review of our definition of "Praise" in the [Dangerous Organizations and Individuals](#) policy, along with its carveout on news reporting through a policy forum process has led us to begin the implementation phase of the update. This phase involves training human reviewers on the new policy and updating our external facing community standards to reflect that update. As proposed in our [previous update](#), we will continue to explore developing new classifiers that will help track this policy allowance enforcement onwards and update the board on our progress in future Quarterly Updates.

Oversight Board Recommendation: Meta should conduct a review of the HIPO ranker to examine if it can more effectively prioritize potential errors in the enforcement of allowances to the Dangerous Individuals and Organizations Policy. This should include examining whether the HIPO ranker needs to be more sensitive to news reporting content, where the likelihood of false-positive removals that impacts freedom of expression appears to be high. The Board asks Meta to inform the Board within 60 days of the results of its review and the improvements it will make to avoid errors of this kind in the future.

(*[Post from a News Outlet Discussing the Taliban Government in Afghanistan #6](#)*)

Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In progress
May 2023 Update	As we shared in our response to [Post from a News Outlet Discussing the Taliban Government in Afghanistan #7](#), we have been conducting several experiments aimed at improving the efficiency of our mistake prevention tools – including the High Impact False Positive Override (HIPO) ranker. This quarter, we focused on expanding the capacity allocation of the HIPO ranker as a first step in expanding upon the system. This year, we are continuing to review the HIPO ranker for opportunities to improve its prioritization and effectiveness to address false positive removals. We will update the board on the status of this work in future Quarterly Updates.

Oversight Board Recommendation: To improve Meta's ability to remove non-violating content from banks programmed to identify or automatically remove violating content, Meta should ensure that content with high rates of appeal and high rates of successful appeal is re-assessed for possible removal from its Media Matching Service banks. The Board will consider this recommendation implemented when Meta: (i) discloses to the Board the rates of appeal and successful appeal that trigger a review of Media Matching Service-banked content, and (ii) confirms publicly that these reassessment mechanisms are active for all its banks that target violating content.

([Cartoon Depicting Violence by the Police in Colombia #1](#))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress
May 2023 Update	We remain committed to the process of improving the governance model for MMS banking towards standardization across the banks. In efforts towards this, we're driving towards the implementation of an internal system designed to address appeals spikes in order to investigate these and diagnose inefficiencies in our ecosystem. We seek to understand the root cause of any emerging errors across our media matching, classifier or country specific tooling. Under the guidance of this effort, we will be able to sample specific segments such as violation types or markets which may experience sudden spikes in appeal rates and sample these for urgent human review. Some violation areas have operational teams that manually review appeal jobs that were overturned and identify clusters that may have been incorrectly banked by the original bank. We aim to automate this review process in 2023, to reduce the time-lag between when a review accepts an appeal and when the associated cluster is actually removed from the bank. Additionally, some violation groups now have human review queues where content is added to banks after a designated set of skilled human reviewers who are trained on banking eligible elements have reviewed the incoming content and completed the Dynamic Multi-Review (DMR) process where a minimum of three reviewers must agree on the specific violation for the content to be added to a bank.

Across our banks, we conduct periodic bank clean-up sprints to remove poor quality or misclassified content, have set up virality alerts for banks and continue to refine our metrics assessment of banking accuracy. As explained in our previous quarterly update, our banking system is dynamic due to differing use cases, policy types, and banking strategies. We will continue to drive uniformity across these and share our progress towards improving appeals processes with the board.

Oversight Board Recommendation: To ensure that inaccurately banked content is quickly removed from Meta's Media Matching Service banks, Meta should set and adhere to standards that limit the time between when banked content is identified for re-review and when, if deemed non-violating, it is removed from the bank. The Board will consider this recommendation implemented when Meta: (i) sets and discloses to the Board its goal time between when a re-review is triggered and when the non-violating content is restored, and (ii) provides the Board with data demonstrating its progress in meeting this goal over the next year.

(*Cartoon Depicting Violence by the Police in Colombia #2*)

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress
May 2023 Update	Many of our MMS banking teams already have strict standards for the re-review and potential removal of flagged content within their banks and they continue to improve across the board. As mentioned in our previous Quarterly Update, the various instances of cases, policy types, and banking strategies lead to differentiation in the standards across violation groups. Currently, we have enabled spike detection alerts systems to catch newly banked content that was incorrectly removed in "real-time" for certain violation groups. These systems measure the rate of enforcement for specific content types and alerts human reviewers to review above specific enforcement thresholds. This mechanism enables our teams to improve banking precision by using virality to review and remove incorrectly removed content in real-time, thereby minimizing the lag time between content being incorrectly banked, reviewed and then removed. Additionally, we have built designated queues that have assigned metrics to assess incorrectly banked content and monitor these with assigned metrics. We continue to work towards consolidating these efforts across our banks and will provide further updates on our progress in future Quarterly Updates.

Oversight Board Recommendation: To improve enforcement of its content policies during periods of armed conflict, Meta should assess the feasibility of establishing a sustained internal mechanism that provides the expertise, capacity and coordination required to review and respond to content effectively for the duration of a conflict. The Board will consider this recommendation implemented when Meta provides an overview of the feasibility of a sustained internal mechanism to the Board.

(*Post Calling for Violence in Ethiopia #2*)

Previous Commitment	Assessing Feasibility
Updated Commitment	Implementing Fully

Current Status	In progress
May 2023 Update	In our [last Quarterly Update](), we committed to assessing the feasibility of a new crisis coordination team to provide dedicated Operations oversight of all tactical execution of support efforts before, during, and after crises. Following this assessment, we are in the process of setting up such a team to address crisis coordination and provide dedicated Operations oversight throughout the lifecycle of imminent and emerging crises. This team will be responsible for operational elements of critical events including high risk events and elections. The team will be composed of crisis leads in our three main regions (APAC, EMEA and NA) who will provide 24/7 coverage to crisis events. This will be initially executed through a pilot team that will deal with planned and unplanned critical events. All Meta regions will be covered within the scope of the team, with dedicated regional representation. We have finalized approvals and have structurally prepared to operationalize the cross functional team by the second half of 2023, pending the finalization of restructuring efforts within the organization. We will provide further updates on our progress in future Quarterly Updates.

Oversight Board Recommendation: Meta should provide users with the opportunity to appeal to the Oversight Board for any decisions made through Meta's internal escalation process, including decisions to remove content and to leave content up. This is necessary to provide the possibility of access to remedy to the Board and to enable the Board to receive appeals for "escalation-only" enforcement decisions. This should also include appeals against removals made for Community Standard violations as a result of "trusted flagger" or government actor reports made outside in-product tools. The Board will consider this implemented when it sees user appeals coming from decisions made on escalation and when Meta shares data with the Board showing that for 100% of eligible escalation decisions, users are receiving reference IDs to initiate appeals.

([Veiled Threat of Violence Based on Lyrics from a Drill Rap Song #3]() (along with [PAO on Meta's Cross-Check Policies #25]()[11]))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	Complete
May 2023 Update	Starting on April 19, we completed the development of an alternate pathway which expands users' ability to appeal eligible content decisions, including those made on escalation, to the Oversight Board. Previously, content decisions that we made during an internal escalation process were often not appealable because of separate pathways used for specialized contextual review. This is a product solution and therefore, inclusive of eligible appeals made in-product only. We expect this major product update to significantly expand the volume and scope of content that users are able to appeal to the Oversight Board, further increasing the board's impact. We now consider this recommendation complete and will have no further updates.

[11] The board issued a similar recommendation in recommendation #25 in the [PAO on Meta's Cross-Check Policies](). We are tracking the progress of our work in response to this recommendation as part of our response to recommendation #3 in the Veiled Threat of Violence Based on Lyrics from a Drill Rap Song case

Oversight Board Recommendation: *Meta should implement and ensure a globally consistent approach to receive requests for content removals (outside in-product reporting tools) from state actors by creating a standardized intake form asking for minimum criteria, for example, the violated policy line, why it has been violated, and a detailed evidential basis for that conclusion, before any such requests are actioned by Meta internally. This contributes to ensuring more organized information collection for transparency reporting purposes. The Board will consider this implemented when Meta discloses the internal guidelines that outline the standardized intake system to the Board and in the Transparency Centre.*

(Veiled Threat of Violence Based on Lyrics from a Drill Rap Song #4)

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	In our Q4 2022 Quarterly Update, we shared that we are working to consolidate and standardize intake of content reports by state actors. We highlighted that we have begun the rollout of a platform that allows for a more consistent intake approach for incoming requests from some state actors. We have since onboarded the majority of global regulators and plan to conclude onboarding our regulatory partners in H2 2023. The work of unifying intake processes continues to be informed and affected by regionally-specific compliance, practical, and legal obligations, which makes it difficult to adopt a globally uniform intake system. We remain aligned with the spirit of the board's recommendation and are committed to adopting a consistent approach to the extent possible. We will provide further updates on this multi-stage work in future Quarterly Updates.

Oversight Board Recommendation: *Meta should mark and preserve any accounts and content that were penalised or disabled for posting content that is subject to an open investigation by the Board. This prevents those accounts from being permanently deleted when the Board may wish to request content that is referred for decision or to ensure that its decisions can apply to all identical content with parallel context that may have been wrongfully removed. The Board will consider this implemented when Board decisions are applicable to the aforementioned entities and Meta discloses the number of said entities affected for each Board decision.*

(Veiled Threat of Violence Based on Lyrics from a Drill Rap Song #5)

Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility
Current Status	In Progress
May 2023 Update	In our last Quarterly Update we shared that we are assessing the feasibility of a mechanism that will allow us to extend the preservation period for accounts and content subject to open board investigations in selected cases, while upholding our obligations to user data privacy. In the status quo, as outlined in our privacy policy, account data may be preserved in limited scenarios including where we are legally obligated to do so, if retention is necessary for investigating certain violations of our policies in relation to a legal claim or

	litigation. We will be working on this assessment in tandem with other regulatory obligations. We will provide further updates on our progress in future Quarterly Updates.

Oversight Board Recommendation: Meta should notify Instagram users when a warning screen is applied to their content and provide the specific policy rationale for doing so. The Board will consider this recommendation implemented when Meta confirms that notifications are provided to Instagram users in all languages supported by the platform. *(Aftermath of an Attack on a Church in Nigeria #2)*

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As shared in our initial response, we are currently working on improving the availability and granularity of information shared in user messaging across all violation areas and enforcement types. We are continuing to develop a variety of efforts dedicated toward improving the experience for people on our platforms when we make enforcement decisions or safety actions on their content. In light of these regulatory requirements taking priority, we hope to begin work on warning screens in 2024, and will share updates on our progress in a future Quarterly Update.

Oversight Board Recommendation: Meta should update its internal guidance to at-scale reviewers on when to escalate content reviewed under the Adult Sexual Exploitation Community Standard, including guidance to escalate content depicting non-consensual sexual touching, with the above policy exception. The Board will consider this recommendation implemented when Meta shares with the Board the updated guidance to at-scale reviewers. *(Video Depicting Sexual Harassment in India #2)*

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As shared in our response to Video Depicting Sexual Harassment in India #1, we have finalized changes to our Adult Sexual Exploitation Policy to allow depictions of non-consensual sexual touching with a warning screen where the content is shared to raise awareness, the victim is not identifiable, the content does not involve overt nudity and explicit sexual activity, and it is not shared in a sensationalized context. Once this policy is launched, we will reflect these changes in policy guidance we provide to our at scale reviewers. We will share an update on our progress in the next Quarterly Update.

Oversight Board Recommendation: Meta should split, either by distinct pathways or prioritization, any list-based over-enforcement prevention program into separate systems: one to protect expression in line with Meta's human rights responsibilities, and one to protect expression that Meta views as a business priority that falls outside that category.

	([PAO on Meta's Cross-Check Policies #1](#))
Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	In our March 6th, 2023 [response](#), we shared that we will partner with our Product and Policy teams, as well as our Global Operations (GO) team to establish an objective process that would allow us to distinguish entities which are included to respect freedom of expression in line with our human rights responsibilities and those that are purely business interests, wherever possible. This is while recognizing that an absolute distinction between our business interests and human rights responsibilities is not always possible. It is worth noting that this effort is highly complex and therefore requires significant understand work and process investments. So far, we have started efforts to label advertisers and roadmapped how we will be collaborating with relevant stakeholders going forward. We expect more progress here in upcoming quarters, and will continue to update the Board in future Quarterly Updates.

Oversight Board Recommendation: Meta should ensure that the review pathway and decision making structure for content with human rights or public interest implications including its escalation paths, is devoid of business considerations. Meta should take steps to ensure that the team in charge of this system does not report to public policy or government relations teams or those in charge of relationship management with any affected users.

([PAO on Meta's Cross-Check Policies #2](#))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	We are constantly evaluating and exploring ways that would better refine our guardrails at each level of the review pathway & decision-making structure for content with human rights or public interest implications, to prevent bias and error. We are piloting this effort through our Early Response Secondary Review (ERSR) program. As we shared in our [initial response](#), our content review processes & policies are intrinsically designed to prioritize the protection of communities and people on our platform, and we will continue to hone them to eliminate bias and error as much as possible. We expect to share more details on our progress in future Quarterly Updates.

Oversight Board Recommendation: Meta should improve how its workflow dedicated to meet Meta's human rights responsibilities incorporates context and language expertise on enhanced review, specifically at decision making levels.

([PAO on Meta's Cross-Check Policies #3](#))

Previous Commitment	Implementing Fully

Updated Commitment	Implementing Fully
Current Status	In Progress

May 2023 Update	In Q3 2022, we scaled the highest level of Early Response Secondary Review (ERSR) to 14 of our regional teams with the highest volume of ERSR jobs - a first of several phases to further incorporate context and language expertise at this level. As we measure the success of phase one, we are exploring ways to continue expansion of this initiative. Simultaneously, we have taken steps to train a portion of regional experts who remain at the General Secondary Review (GSR) level on escalation-only policies so they may apply context-specific decisions. As explained in our initial update, historically, we have not scaled context-specific decisions to larger groups of reviewers because quality and consistency of review suffers. However, in areas where we have confidence in high-quality decisions, we will continue identifying opportunities to:
	
	    A. Increase the number of regional experts participating in ERSR level review; and
	    B. Increase the number of GSR reviewers enabled to apply context-specific decisions.
	
	We will provide further updates on our progress in future Quarterly Updates.

Oversight Board Recommendation: *Meta should establish clear criteria and processes for audit. Should entities no longer meet the eligibility criteria, they should be promptly removed from the system. Meta should review all included entities in any mistake prevention system at least yearly. There should also be clear protocols to shorten that period where warranted.* *(PAO on Meta's Cross-Check Policies #11)*

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	Since our March 6th response, we have established initial protocols for applying certain ERSR tags for a short-term period in cases where the increase in risk is temporal in nature, making the standard one year eligibility period too long. This is particularly helpful for human rights entities who may become eligible for additional review during significant world events, such as the World Cup or an election. Additionally, we are in the process of creating protocols for auditing certain entities within 30 days of the tag application to confirm eligibility. Since this tag is often applied during a short-term sensitive escalation, we want to ensure that the entities are truly eligible for long-term protection earlier than the standard timeline. We will continue to provide updates on this work in a future Quarterly Update.

Oversight Board Recommendation: *Meta should notify all entities that it includes on lists to receive enhanced review and provide them with an opportunity to decline inclusion.* *(PAO on Meta's Cross-Check Policies #14)*

Previous Commitment	Assessing Feasibility
Updated Commitment	Assessing Feasibility

Current Status	In Progress
May 2023 Update	In our March 6, 2023 response, we shared that we will collaborate with our Human Rights and Civil Rights teams to assess options to address the issue of some users who may not wish to be included on cross-check lists, even if for their benefit. While we believe that marking or informing the entities included on cross-check lists would lead to serious security concerns and potentially impact the integrity of the program as outlined in our initial response to PAO on Meta's Cross-Check Policies #12; we also recognize the importance of transparency and user autonomy. Since publishing the response, we have assembled a team to explore the feasibility of implementing this recommendation via creative solutions that would mitigate those same unintended consequences. We will provide an update on our assessment in a future Quarterly Update.

Oversight Board Recommendation: Meta should consider reserving a minimum amount of review capacity by teams that can apply all content policies (e.g., the Early Response Team) to review content flagged through content based mistake-prevention systems. (*PAO on Meta's Cross-Check Policies #15*)	
Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	As explained in our March 6th response, we have refrained from scaling context-specific decisions to larger groups of reviewers because the quality and consistency of content review can suffer as the size of the reviewer pool increases. It is important that we ensure consistency of high quality policy implementation before doing so. Regardless, as mentioned in PAO on Meta's Cross-Check Policies #3, we have begun training a portion of regional experts who remain at the General Secondary Review (GSR) level on escalation-only policies, so they may apply context-specific decisions. Additionally, we are working to develop the necessary measurement infrastructure to assess various aspects of decision making at the GSR and Early Response Secondary Review (ERSR) level, and identify key areas for improvement. We will continue to report on our progress in the next Quarterly Update.

Oversight Board Recommendation: Meta should take measures to ensure that additional review decisions for mistake-prevention systems that delay enforcement are taken as quickly as possible. Investments and structural changes should be made to expand the review teams so that reviewers are available and working in relevant time zones whenever content is flagged for any enhanced human review. (*PAO on Meta's Cross-Check Policies #16*)	
Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress

May 2023 Update	In alignment with the commitments we made in our initial response, we are in the process of conducting important foundational work to understand more advanced metrics before establishing robust Service-Level Agreements (SLAs) for review decisions across our mistake-prevention systems. We are also in the exploratory stages of developing contingency efforts to tackle potential surges in volumes that prevent us from abiding by these SLAs. We will provide an update on the status of this recommendation in the next Quarterly Update.

Oversight Board Recommendation: Meta should not delay all action on content identified as potentially severely violating and should explore applying interstitials or removals pending any enhanced review. The difference between removal or hiding and downranking should be based on an assessment of harm, and may be based, for example, on the content policy that has possibly been violated. If content is hidden on these grounds, a notice indicating that it is pending review should be provided to users in its place.

(*PAO on Meta's Cross-Check Policies #17*)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	Since our March 6th, 2023 response, teams within Global Operations (GO) collaborated to eliminate our backlogs in cross-check reviews. Additionally, we are continuing to explore further options to best protect the users of our platform from harm while content flagged for certain violations is pending cross-check review. However, we will not be implementing the board's recommendation to provide users with a notice that indicates their content is pending review, should it be flagged for possibly violating our Community Standards. This is because doing so could absolve potential violators from our strike penalty systems and thresholds, and users could proactively self-delete potentially violating content once notified that their content is pending review - before they are reviewed internally. As iterated earlier, all entities on Facebook & Instagram, including those whose content is removed through cross-check, are subject to a strike threshold and are disabled once that threshold is reached. We will continue to update the board on our progress in future Quarterly Updates.

Oversight Board Recommendation: Meta should not operate these programs at a backlog. Meta should not, however, achieve gains in relative review capacity by artificially raising the ranker threshold or having its algorithm select less content.

(*PAO on Meta's Cross-Check Policies #18*)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As mentioned in our update to PAO on Meta's Cross-Check Policies #17, we worked with

our operations teams to significantly reduce our backlogs in cross-check reviews within the Early Response Secondary Review (ERSR) program. We are also continuing to explore further options to best protect the users of our platform from harm while content flagged for certain violations is pending cross-check review. In the coming quarter, we not only plan to prioritize re-evaluating the current strategy for how our ERSR & GSR systems select & prioritize content for review, but also redefine how we can best optimize our existing Operations teams' capacity to review more content efficiently with modifications to our content review infrastructure. This is to prevent the occurrence of future cross-check backlogs, without raising the ranker threshold artificially or having its algorithm select less content. We will continue to report on our progress in the next Quarterly Update.

Oversight Board Recommendation: *Meta should not automatically prioritize entity-based secondary review and make a large portion of the algorithmically selected content-based review dependent on extra review capacity.*

(*PAO on Meta's Cross-Check Policies #19*)

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	In our March 6th, 2023 response, we shared that in the cross-check program's current state, only entities on the ERSR list receive the benefit of automatic specialized review. We've maintained this approach; however, we have invested significantly in expanding scope for those reviewing content escalated by the General Secondary Review (GSR) ranker system. To address the Board's recommendation that we should not prioritize ERSR review over GSR in terms of review capacity, we launched a series of training efforts to upskill a portion of the regional experts who staff the GSR level on escalation-only policies, as mentioned in PAO on Meta's Cross-Check Policies #3. We will continue exploring opportunities to scale certain context-specific review decisions to larger groups of reviewers as we progress toward greater consistency of high quality policy implementation. We will continue to provide updates on this recommendation in future Quarterly Updates.

Oversight Board Recommendation: *Meta should ensure that content that receives any kind of enhanced review because it is important from a human rights perspective, including content of public importance, is reviewed by teams that can apply exceptions and context.*

(*PAO on Meta's Cross-Check Policies #20*)

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As stated in our response to PAO on Meta's Cross-Check Policies #3, we began leveraging language experts from 14 regional teams at the highest level of review - Early Response Secondary Review (ERSR), as the first of several phases to further incorporate context and

	language expertise at this level. We have efforts underway to measure the success of phase one, while simultaneously exploring ways to continue expansion of this initiative. Alongside this, a portion of regional experts who remain at the General Secondary Review (GSR) level, have also been trained on escalation-only policies so that they may apply context-specific decisions. As we shared in our initial response, scaling context-specific decisions to larger groups of reviewers is highly contingent on the quality and consistency of review. We will continue to invest in increasing the number of regional experts participating in ERSR level review as well as the number of GSR reviewers we empower to apply context-specific decisions at a pace that allows high-quality and consistent policy implementation. Refining our ERSR categories, notably Higher Context, Significant World Events, and Historically Overenforced to increase coverage for content and entities with human rights or public interest implications continues to be a priority and work in progress. We will update the board on our efforts in this space in future Quarterly Updates.

Oversight Board Recommendation: *Meta should establish clear criteria for the application of any automatic bars to enforcement ('technical corrections'), and not permit such bars for high severity content policy violations. At least two teams with separate reporting structures should participate in granting technical corrections to provide for cross-team vetting.*

([PAO on Meta's Cross-Check Policies #21](#))

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	Since our [March 6th response](#), we've established more stringent requirements for entities already included in automatic bars to enforcement: 'technical corrections' lists now require a risk-based approach at every 6-month renewal cycle. The risk based approach requires teams responsible for governing the technical corrections list renewal process to identify high risk actions and violations that may be barred from enforcement. These high risk actions are identified using the pre-established risk spectrum for technical corrections governance, and, if the team identifies a high risk action, they're responsible for removing those entities from the list. The risk spectrum for entities currently included and those to be included in the future ranges from low to medium severity violations. We currently do not permit bars to enforcement for high severity violations. This approach is expected to significantly decrease the overall list size of technical corrections where high risk inclusion has been established. We currently have teams with separate reporting structures participating in both the initial approval as well as the renewals processes, where the initial review is done by Operations teams and then Legal/Policy teams. We will provide further updates on this progress in our next Quarterly Update.

Oversight Board Recommendation: *Meta should ensure that all content that does not reach the highest level of internal review is able to be appealed to Meta.*

([PAO on Meta's Cross-Check Policies #24](#))

Previous Commitment	Implementing in Part

Updated Commitment	Implementing in Part
Current Status	In Progress
May 2023 Update	As mentioned in our initial response to this recommendation, we are rolling out multiple systemic changes across our infrastructure, business process, strategy, and partnership functions to expand the availability of appeals for Global Operations (GO) enforcement. Since our initial update, Meta has been designated as a VLOP (Very Large Online Platform) under the EU DSA, with compliance effective from the second half of 2023. As such, we are working to ensure the necessary infrastructure is in place for eligible GO enforcement to be appealable by that date. Additionally, in our response to Veiled Threat of Violence Based on Lyrics from a Drill Rap Song #3, we shared that we have completed the development of an alternate pathway which expands users' ability to appeal eligible content decisions, including those made on escalation, to the Oversight Board. We will continue to mature this initiative alongside upcoming regulatory requirements, and will provide an update on our progress in a future Quarterly Update.

Oversight Board Recommendation: *Meta should use the data it compiles to identify "historically over-enforced entities" to inform how to improve its enforcement practices at scale. Meta should measure over-enforcement of these entities and it should use that data to help identify other over-enforced entities. Reducing over-enforcement should be an explicit and high-priority goal for the company.* (*PAO on Meta's Cross-Check Policies #26*)
Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	In Progress
May 2023 Update	As explained in our March 6th, 2023 response, we view this recommendation to be a long-term priority for the success of the cross-check program. We are still in the exploratory stages of implementing our long-term commitment to leverage both metric-based models and non data-based analysis to strengthen enforcement practices at scale, which will enable us to objectively identify historically over-enforced entities on our platform. We are continuing research in this space and refining our metric-based models, and will provide further insights in future Quarterly Updates.

Oversight Board Recommendation: *Meta should use trends in overturn rates to inform whether to default to the original enforcement within a shorter time frame or what other enforcement action to apply pending review. If overturn rates are consistently low for particular subsets of policy violations or content in particular languages, for example, Meta should continually calibrate how quickly and how intrusive an enforcement measure it should apply.* (*PAO on Meta's Cross-Check Policies #27*)
Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part

Current Status	In Progress
May 2023 Update	As stated in our [March 6th, 2023 response](#), we will continue evaluating patterns of enforcement decisions for larger improvement within our systems. At this time, teams within our Global Operations (GO) organization are beginning to break down overturn rate, false positive rate, and decision discrepancy rate by region and policy violation type in order to identify key areas for quality improvements. We look forward to sharing more progress on these efforts and how these metrics are being used to make improvements in future Quarterly updates.

Oversight Board Recommendation:Meta should err on the side of issuing scaled allowances where (i) this is not likely to lead to violence; (ii) when potentially violating content is used in protest contexts; and (iii) where public interest is high. Meta should ensure that their internal process to identify and review content trends around protests that may require context-specific guidance to mitigate harm to freedom of expression, such as allowances or exemptions, are effective. The Board will consider this recommendation implemented when Meta shares the internal process with the Board and demonstrates through sharing data with the Board that it has minimized incorrect removals of protest slogans.

([A Cartoon Depicting Iran's Supreme Leader Ayatollah Ali Khamenei #2](#))

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress
May 2023 Update	In relation to this recommendation, we are pursuing policy development related to our [Violence and Incitement](#) policy – including examining calls for death and high risk persons. As part of this work, we are also scoping out and considering ways to strike a better balance between violent speech and political expression, particularly rhetorical political speech in protest contexts. We expect that this policy development, given its complexity, will take some time to fully complete and implement and therefore expect this recommendation to remain in progress for upcoming Quarterly Updates.

Oversight Board Recommendation: Pending changes to the Violence and Incitement policy, Meta should issue guidance to its reviewers that "marg bar Khamenei" statements in the context of protests in Iran do not violate the Violence and Incitement Community Standard. Meta should reverse any strikes and feature limits for wrongfully removed content that used the "marg bar Khamenei" slogan. The Board will consider this recommendation implemented when Meta discloses data on the volume of content restored and number of accounts impacted.

([A Cartoon Depicting Iran's Supreme Leader Ayatollah Ali Khamenei #3](#))

Previous Commitment	Implementing Fully
Updated Commitment	Implementing Fully
Current Status	Complete
May 2023 Update	In our [January 23rd, 2023](#) response, we fully implemented the Board's recommendation to

	allow the phrase "marg bar Khamenei" in the context of the ongoing protests in Iran. We have also run searches to identify previous enforcement actions and reverse strikes made on the basis of this type of content to the most feasible extent possible, given our considerations pertaining to capacity and data retention. These searches surfaced a sample of qualifying organic content for review and led to the reversal of strikes on accounts, within the sample, that had been penalized for using the phrase in the designated context. We will continue to pursue further reversals where they are applicable and feasible. We now consider this recommendation complete and will have no further updates.

Oversight Board Recommendation: Meta should revise the indicators that it uses to rank appeals in its review queues and to automatically close appeals without review. The appeals prioritization formula should include, as it does for the cross-check ranker, the factors of topic sensitivity and false-positive probability. The Board will consider this implemented when Meta shares with the Board their appeals prioritization formula and data that shows that it is ensuring review of appeals against the incorrect removal of political expression in protest contexts.

(A Cartoon Depicting Iran's Supreme Leader Ayatollah Ali Khamenei #4)

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress
May 2023 Update	As described in our previous update, our teams have launched the first iteration of a solution for the ranking and automation of content takedown appeals jobs. We continue to assess the impact of this rollout and refine our ranking approach and process as we uncover efficiencies. These assessments will aid us in future roadmapping processes as our internal optimization and appeals experience teams continue to understand the incentives and tradeoffs of appeals prioritization. This is an ongoing process to be developed throughout 2023. We will continue to share updates on the development of our efforts in future Quarterly Updates, including expansion on the feasibility of sharing confidential data with the board to verify implementation.

Oversight Board Recommendation: In order to ensure that Meta's internal criteria for its Sexual Solicitation policy do not result in the removal of more content than the public-facing policy indicates and so that non-sexual content is not mistakenly removed, Meta should revise its internal reviewer guidance to ensure that the criteria reflect the public-facing rules and require a clearer connection between the "offer or ask" and the "sexually suggestive element." The Board will consider this implemented when Meta provides the Board with its updated internal guidelines that reflect these revised criteria.

(Gender Affirming Surgery Bundle #3)

Previous Commitment	Implementing in Part
Updated Commitment	Implementing in Part
Current Status	In progress

May 2023 Update	As shared in our update for [Gender Affirming Surgery Bundle #2](#), we are in the process of scoping out refinements for our Sexual Solicitation policy to better explain what constitutes an "offer or ask" and "sexually suggestive element." Though we are still in the early stages of conducting this policy development, we will align our internal reviewer guidance with our publicly-facing [Sexual Solicitation](#) policy should we make any policy changes once the development process is complete. We will share updates on our progress in future Quarterly Updates.

Appendix D. Long Term Recommendations

Long-Term Transparency Recommendations	
Oversight Board Recommendation: *Expand transparency reporting to disclose data on the number of automated removal decisions per Community Standard, and the proportion of those decisions subsequently reversed following human review.* (*Breast Cancer Symptoms & Nudity #6*)	
Commitment	Assessing Feasibility
Workstream Summary	Our current focus for this work is on improving what we internally call "data readiness", by aligning on a consistent accounting methodology across metrics. We are working to define binaries for each metric as a first step towards aggregating public-facing enforcement metrics. To do this, we are discussing complexities such as how to quantify instances of enforcement conducted by human review and automated tools (e.g. quantifying cases where a human reviewer determined that an image was violating and then a machine scaled that decision more broadly). Concurrently, we are resolving gaps in our logging infrastructure to allow us to pull those metrics once we've decided on how to report it.
Next Expected Update	Q4 2024
Oversight Board Recommendation: *In its transparency reporting, Facebook should include numbers of profile, page, and account restrictions, including the reason and manner in which enforcement action was taken, with information broken down by region and country.* (*Former President Trump's Suspension #18*)	
Commitment	Implementing in Part
Workstream Summary	We are currently working on two long-term initiatives prompted by this recommendation: measuring our enforcement actions on profile, page, and account restrictions; and measuring enforcement data by location. Both of these initiatives fit into our overall vision for the Community Standards Enforcement Report (CSER).
Next Expected Update	Q4 2023
Oversight Board Recommendation: *Facebook should improve its transparency reporting to increase public information on error rates by making this information viewable by country and language for each Community Standard.* (*Punjabi Concern Over the RSS in India #3*)	
Commitment	Implementing in Part
Workstream Summary	We are conducting long-term work to define our accuracy metrics, alongside our work on Breast Cancer Symptoms & Nudity #6. As we continue to develop the necessary measurement infrastructure and data validation protocols to report high-quality, consistent information, we are continuing to engage with the board on our more incremental roadmaps, challenges, and expansion opportunities.

Next Expected Update	Q4 2024

Oversight Board Recommendation: Meta should regularly review the data on its content moderation decisions prompted by state actor content review requests to assess for any systemic biases. Meta should create a formal feedback loop to fix any biases and/or outsized impacts stemming from its decisions on government content takedowns. The Board will consider this recommendation implemented when Meta regularly publishes the general insights derived from these audits and the actions taken to mitigate systemic biases.

(*Veiled Threat of Violence Based on Lyrics from a Drill Rap Song #7*)

Commitment	Assessing Feasibility
Workstream Summary	We are working to develop a process to re-review a randomized sample of government requests on an ongoing basis to ensure that our review of those requests was accurate, fair, and consistent with Meta's policies and commitments. This approach will likely include, among other elements, a review of accuracy of any enforcement actions taken under our Community Standards, a review of any actions taken on the basis of local law, and assurance of consistency with our human rights commitments as a member of the Global Network Initiative.
Next Expected Update	Q4 2023

Long-Term Enforcement Recommendations

Oversight Board Recommendation: Meta should conduct regular assessments on reviewer accuracy rates focused on the Restricted Goods and Services policy. The Board will consider this recommendation implemented when Meta shares the results of these assessments with the Board, including how these results will inform improvements to enforcement operations and policy development, and summarize the results in its quarterly Board transparency reports. Meta may consider if these assessments should be extended to reviewer accuracy rates under other Community Standards.

(*Post Requesting Advice on Pharmaceutical Drugs #3*)

Commitment	Implementing in Part
Workstream Summary	In the interest of transparency around how we identify and address potential mistakes in the enforcement of our Restricted Goods and Services Policy, we report on the amount of appealed content and content that is restored on Facebook and Instagram under that policy in our quarterly Community Standards Enforcement Report. Our work defining reviewer accuracy metrics more broadly is underway.
Next Expected Update	Q4 2024

IV. Index

Case	Recommendation	Updated Category	Status	Section	Page
Breast cancer symptoms and nudity	2020-004-IG-UA-2	Assessing feasibility	In progress	Policy	29
	2020-004-IG-UA-6	Assessing feasibility	In progress	Long Term Transparency	60
	2020-004-IG-UA-7	Assessing feasibility	In progress	Policy Footnote 5	29
Former President Trump	2021-001-FB-FBR-18	Implementing in part	In progress	Long Term Transparency	60
Punjabi Concerns Over the RSS in India	2021-003-FB-UA-3	Implementing in part	In progress	Long Term Transparency	60
Armenian People and the Armenian Genocide	2021-005-FB-UA-4	Implementing in part	In progress	Enforcement	40
	2021-005-FB-UA-5	Assessing feasibility	In progress	Enforcement	40
Support of Abdullah Ocalan	2021-006-IG-UA-9	Implementing fully	In progress	Enforcement	41
	2021-006-IG-UA-10	Assessing feasibility	In progress	Policy Footnote 5	29
	2021-006-IG-UA-11	Implementing in part	In progress	Transparency	21
Al Jazeera Post on Tensions Between Israel and Palestine	2021-009-FB-UA-1	Assessing feasibility	In progress	Policy	30
	2021-009-FB-UA-4	Implementing in part	In progress	Transparency Footnote 4	21
Post Depicting Protests in Colombia While Using a Slur	2021-010-FB-UA-4	Assessing feasibility	In progress	Enforcement	41
Depicting Indigenous Artwork and Discussing Residential Schools	2021-012-FB-UA-1	Implementing fully	Complete	Enforcement	42
Post Discussing a Substance with Psychoactive Properties	2021-013-IG-UA-1	Assessing feasibility	In progress	Policy Footnote 5	29
Post Requesting Advice on Pharmaceutical Drugs	2021-015-FB-UA-2	Assessing feasibility	In progress	Enforcement	42
	2021-015-FB-UA-3	Implementing in part	In progress	Long Term Enforcement	61
Post Describing Sexual Violence Against Minors	2021-016-FB-FBR-1	Implementing fully	In progress	Policy	30
	2021-016-FB-FBR-2	Implementing fully	In progress	Policy	30

PAO on Sharing Residential Information	2021-001-FB-PAO-1	Implementing fully	In progress	Policy	31
	2021-001-FB-PAO-3	Implementing fully	In progress	Policy Footnote 7	31
	2021-001-FB-PAO-4	Implementing fully	In progress	Policy	31
	2021-001-FB-PAO-7	Assessing feasibility	In progress	Policy	32
	2021-001-FB-PAO-8	Implementing fully	In progress	Policy	32
	2021-001-FB-PAO-9	Assessing feasibility	In progress	Policy Footnote 5	29
	2021-001-FB-PAO-10	Implementing in part	In progress	Enforcement Footnote 9	40
	2021-001-FB-PAO-14	Implementing in part	In progress	Enforcement Footnote 9	40
Video of an Edited Cartoon Depicting a Croatian City	2022-001-FB-UA-2	Implementing fully	Complete	Enforcement	43
Video Depicting a Civilian Victim of Violence in Sudan	2022-002-FB-MR-1	Assessing feasibility	In progress	Policy	33
	2022-002-FB-MR-2	Assessing feasibility	In progress	Policy Footnote 8	33
	2022-002-FB-MR-4	Assessing feasibility	In progress	Enforcement Footnote 10	41
News Outlet Discussing the Taliban Government in Afghanistan	2022-005-FB-UA-3	Assessing feasibility	In progress	Policy Footnote 6	30
	2022-005-FB-UA-4	Implementing fully	In progress	Enforcement	44
	2022-005-FB-UA-5	Assessing feasibility	In progress	Enforcement	44
	2022-005-FB-UA-6	Assessing feasibility	In progress	Enforcement	44
Cartoon Depicting Violence by the Police in Colombia	2022-004-FB-UA-1	Implementing in part	In progress	Enforcement	45
	2022-004-FB-UA-2	Implementing in part	In progress	Enforcement	46
Post Calling for Violence in Ethiopia	2022-006-FB-MR-2	Implementing fully	In progress	Enforcement	46
Veiled Threat of Violence Based on Lyrics from a Drill Rap Song	2022-007-IG-MR-2	Implementing fully	In progress	Policy	33
	2022-007-IG-MR-3	Implementing in part	Complete	Enforcement	47
	2022-007-IG-MR-4	Implementing in part	In progress	Enforcement	48
	2022-007-IG-MR-5	Assessing feasibility	In progress	Enforcement	48
	2022-007-IG-MR-6	Implementing in part	In progress	Transparency Footnote 4	21

	2022-007-IG-MR-7	Assessing feasibility	In progress	Long Term Transparency	61
Individual Killed in Ukraine During Russian Invasion	2022-008-FB-UA-1	Implementing fully	In progress	Policy	34
	2022-008-FB-UA-2	Implementing fully	In progress	Policy	34
Aftermath of an Attack on a Church in Nigeria	2022-011-IG-UA-1	Assessing feasibility	In progress	Policy	35
	2022-011-IG-UA-2	Implementing fully	In progress	Enforcement	49
Video Depicting Sexual Harassment in India	2022-012-IG-MR-1	Implementing fully	In progress	Policy	35
	2022-012-IG-MR-2	Implementing fully	In progress	Enforcement	49
PAO on Meta's Cross-Check Policies	2021-002-FB-PAO-0	Implementing Fully	In progress	Transparency	21
	2021-002-FB-PAO-1	Implementing in part	In progress	Enforcement	50
	2021-002-FB-PAO-2	Implementing in part	In progress	Enforcement	50
	2021-002-FB-PAO-3	Implementing fully	In progress	Enforcement	50
	2021-002-FB-PAO-4	Implementing in part	In progress	Policy	35
	2021-002-FB-PAO-5	No further action	No further updates	Footnote 3	14
	2021-002-FB-PAO-6	No further action	No further updates	Footnote 3	14
	2021-002-FB-PAO-7	Implementing in part	In progress	Transparency	22
	2021-002-FB-PAO-8	Implementing in part	In progress	Transparency	22
	2021-002-FB-PAO-9	Implementing in part	In progress	Transparency	22
	2021-002-FB-PAO-10	Implementing in part	In progress	Transparency	23
	2021-002-FB-PAO-11	Implementing in part	In progress	Enforcement	51
	2021-002-FB-PAO-12	No further action	No further updates	Footnote 3	14
	2021-002-FB-PAO-13	No further action	No further updates	Footnote 3	14
	2021-002-FB-PAO-14	Assessing feasibility	In progress	Enforcement	51
	2021-002-FB-PAO-15	Implementing in part	In progress	Enforcement	52
	2021-002-FB-PAO-16	Implementing in part	In progress	Enforcement	52
	2021-002-FB-PAO-17	Implementing fully	In progress	Enforcement	53
	2021-002-FB-PAO-18	Implementing fully	In progress	Enforcement	53
	2021-002-FB-PAO-19	Implementing in part	In progress	Enforcement	54
	2021-002-FB-PAO-20	Implementing fully	In progress	Enforcement	54
	2021-002-FB-PAO-21	Implementing fully	In progress	Enforcement	55

	2021-002-FB-PAO-22	Implementing in part	In progress	Transparency	24
	2021-002-FB-PAO-23	Implementing fully	In progress	Transparency	24
	2021-002-FB-PAO-24	Implementing in part	In progress	Enforcement	55
	2021-002-FB-PAO-25	Implementing in part	Complete	Enforcement Footnote 11	47
	2021-002-FB-PAO-26	Implementing fully	In progress	Enforcement	56
	2021-002-FB-PAO-27	Implementing in part	In progress	Enforcement	56
	2021-002-FB-PAO-28	Implementing fully	In progress	Transparency	25
	2021-002-FB-PAO-29	No further action	No further updates	Footnote 3	14
	2021-002-FB-PAO-30	Implementing in part	In progress	Transparency	25
	2021-002-FB-PAO-31	Implementing fully	Complete	Policy	36
	2021-002-FB-PAO-32	Implementing fully	In progress	Transparency	26
Depicting Iran's Supreme Leader Ayatollah Ali Khamenei	2022-013-FB-UA-1	Assessing feasibility	In progress	Policy	37
	2022-013-FB-UA-2	Implementing in part	In progres	Enforcement	57
	2022-013-FB-UA-3	Implementing fully	Complete	Enforcement	57
	2022-013-FB-UA-4	Implementing in part	In progress	Enforcement	58
	2022-013-FB-UA-5	Implementing in part	In progress	Transparency	26
	2022-013-FB-UA-6	Implementing fully	In progress	Transparency	27
	2022-013-FB-UA-7	No further action	No further updates	Transparency	27
Gender Affirming Surgery Bundle	2022-009-IG-UA-1	Implementing in part	In progress	Policy	37
	2022-009-IG-UA-2	Implementing in part	In progress	Policy	37
	2022-009-IG-UA-3	Implementing in part	In progress	Enforcement	58
Donation of Pharmaceutical Drugs to Sri Lanka	2022-014-FB-MR-1	Implementing fully	In progress	Policy	38
	2022-014-FB-MR-2	No further action	No further updates	Footnote 3	14
	2022-014-FB-MR-3	Implementing in part	In progress	Policy	39
	2022-014-FB-MR-4	Implementing in part	In progress	Transparency	28