

2023 ANNUAL REPORT



**IMPROVING HOW META TREATS PEOPLE
& COMMUNITIES AROUND THE WORLD**

Co-Chairs' Foreword

2023 THE YEAR OF IMPACT AND INNOVATION FOR THE BOARD

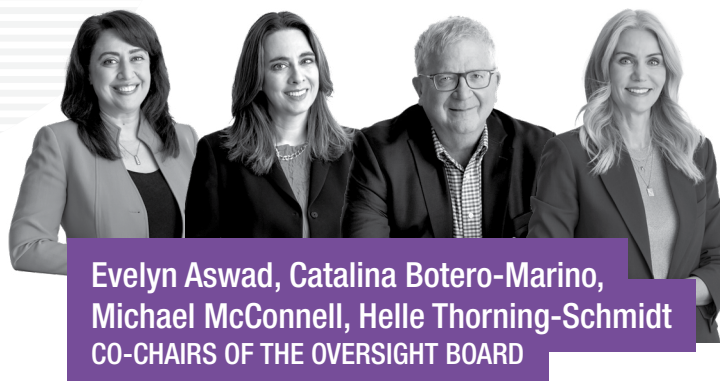


2023 was a year of impact and innovation for the Board. Our recommendations continued to improve how people experience Meta's platforms and, by publishing more decisions in new formats, we tackled more hard questions of content moderation than ever before. From protest slogans in Iran to criticism of gender-based violence, our decisions continued to protect important voices on Facebook and Instagram. For nearly four years, we have pushed Meta to be more transparent with the people who use its platforms. In 2023, we saw further evidence of our impact on users. In response to one of our recommendations, Meta launched Account Status, giving people more information about their current and past penalties. We also saw Meta taking steps to preserve important content by finalizing a new, consistent approach to retaining potential evidence of atrocities.

Following our commitment to publish more cases, faster, we issued our first summary decisions that examine cases in which Meta changed its original decision on a piece of content after it was shortlisted for potential review. We also issued our first expedited decisions which were about the Israel-Hamas conflict. In addition, we published standard decisions related to important topics including prisoners of war, expression by world leaders and discrimination against transgender people. In total, we issued more than 50 decisions in 2023, more than any previous year, overturning Meta's original decision in around 90% of cases.

Looking forward to 2024, we anticipate a busy and impactful year for the Board. As around half of the world's population goes to the polls in 2024, the Board will continue to issue decisions that focus on how Meta's approach to elections and artificial intelligence considers freedom of expression and other human rights. We have expanded our scope to cover Threads and we are interested in exploring major issues that matter to users, such as demoted content. We are also paying attention to the evolving regulatory landscape and the importance of the Board's model — one of an independent, global deliberative body that examines cases through a human rights lens, and provides redress and transparency.

While only four years old, the Board has already tackled some of the toughest questions in content moderation and pushed Meta to treat users fairly. The challenge is enormous, the uncertainty is great, but the work we are doing really matters. Together, we will strive to find answers that improve social media for people all over the world.



**MANAGEMENT & ACCOUNTABILITY
CONTENT GOVERNANCE**

This year, the Board has already expanded its scope to Threads — a major milestone — and published crucial decisions on topics from elections to Holocaust denial. By working with civil society groups, regulators, Meta employees and other platforms, I have no doubt that the Board will continue to go from strength-to-strength in 2024.



4 • Oversight Board 2023 Annual Report

Executive Summary

Meta has fully or partially **implemented 75 of our recommendations** since 2021.

Of these, **34 were implemented** since last year's Annual Report.

In response to our recommendations so far, Meta:



Updated and created new classifiers that prevented a total of **3,500 breast cancer** posts being automatically removed over the course of two 30-day periods



Allowed the term “Marg bar Khamenei” (which literally translates as “Death to [Iran’s supreme leader] Khamenei”) to be shared in the context of the **protests in Iran**



Is finalizing a new, consistent approach to preserving potential **evidence of atrocities**



Launched **Account Status** telling people what penalties were applied to their account and why



Made its **‘strikes’ system** fairer and more transparent



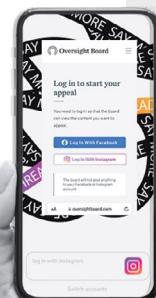
Sent people warnings for **100 million+** pieces of content that were highly likely to be violating – with users deleting their posts 20%+ of the time¹



Launched its **Content Library** – offering researchers in Global Majority countries greater access to data

¹ This data covers a 12-week period. All information is aggregated and de-identified to protect user privacy. All metrics are estimates, based on best information currently available for a specific point in time.

THE BOARD HAS EXPANDED ITS SCOPE TO THREADS



IN 2023, THE OVERSIGHT BOARD...



Published our **first expedited decisions** on the Israel-Hamas conflict **12 days** after we announced the cases.



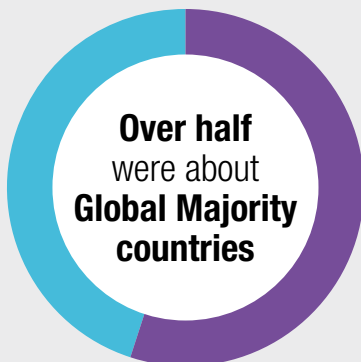
Published our **first summary decisions** examining cases in which Meta changed its original decision after they were appealed to the Board and highlighting areas of improvement for Meta.

ISSUED 53 DECISIONS:

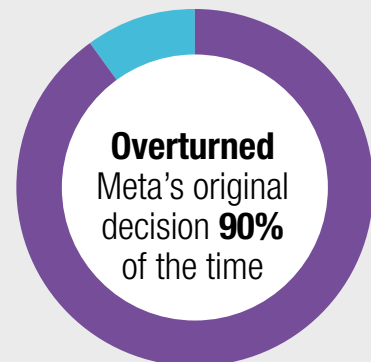
22 STANDARD
DECISIONS

29 SUMMARY
DECISIONS

2 EXPEDITED
DECISIONS



COVERED NEW AREAS
such as
prisoners of war,
extreme diets and
discrimination against
transgender people



Issued a policy advisory opinion
on **COVID-19 misinformation**



Observed Meta **fulfilling its obligations** to implement our decisions within seven days of publication and respond to our recommendations within 60 days



Received **800 public comments**
from organizations and people
around the world

IMPROVING HOW META TREATS PEOPLE & COMMUNITIES AROUND THE WORLD



IN 2024, THE OVERSIGHT BOARD WILL:

1. Explore how our recommendations can identify and mitigate **systemic risks** created by upstream product design choices and the **automated treatment** of content online.
2. Speak out on subjects such as **elections, conflicts** and **manipulated media**.
3. Focus on issues that matter to users – such as **demoted content**.



Recommendations & Impact

TO IMPROVE THE POLICIES APPLIED
TO BILLIONS OF USERS



In our case decisions and policy advisory opinions, we make specific recommendations on how Meta can improve the policies it applies to billions of users.

While our recommendations are non-binding, Meta must respond to them publicly within 60 days. In 2023, Meta made progress on implementing our recommendations, providing us with impact metrics for the first time.



To improve how it deals with elections, protests and health-related content, Meta:

- ➔ Announced it would **allow the term “Marg bar Khamenei” (which literally translates as “Death to [Iran’s supreme leader] Khamenei”) to be shared in the context of the protests in Iran.** After Meta implemented this recommendation in January 2023, Instagram posts using “Marg bar Khamenei” increased by nearly 30% (see pages 23-25 of our [Q1 2023 transparency report](#) for more detail).
- ➔ Updated and created new classifiers, which, over two 30-day periods, **prevented a total of 3,500 pieces of breast cancer-related content from being automatically removed.**
- ➔ Expects to share metrics that will allow the company to **monitor its efforts leading up to, during and after “expected critical events,” such as elections.**



To improve how it treats the people who use its platforms, Meta:

- ➔ Launched Account Status, an in-product experience that provides **information about the current and past penalties on a person's account, including why Meta applied them.**
- ➔ **Changed its 'strikes' system to make it fairer and more transparent** following our concerns around Meta's opaque penalty system and user concerns about being placed in "Facebook jail."
- ➔ Started warning people if their post is highly likely to be violating, **giving them an opportunity to understand Meta's policies and then delete and post their content again.** Over a 12-week period, Meta issued warnings about more than 100 million pieces of content, with users choosing to delete their posts more than 20% of the time.²
- ➔ Committed to **a new system for tracking government requests** following our calls to review how it collects information on such requests. State actors will have to answer a standard set of questions, which Meta says will lead to improved transparency and capacity for public reporting.
- ➔ Translated its rules into new languages, including Pashto and Somali. Since we originally [recommended in April 2021](#) that Meta translates its rules into all languages widely spoken by its users, it has **translated its Community Standards into more than 20 additional languages spoken by more than a billion people worldwide.**



To improve the experiences of people using its platforms in Global Majority countries, Meta:

- ➔ **Launched its Content Library.** This is open to applications from researchers around the world and will **offer researchers in Global Majority countries greater access to data.**
- ➔ Put in place a **Crisis Policy Protocol** to govern its response to crisis situations. The company [used this to assess](#) whether to restore former President Trump to Facebook and Instagram in January 2023.



To improve its response to crisis and conflict situations, Meta:

- ➔ Set up a **new crisis coordination team** to manage high-risk events and elections. The team is composed of crisis leads who provide 24/7 coverage for both planned and unplanned critical events in all regions.
- ➔ Is finalizing **a new, consistent approach to preserving potential evidence of atrocities** and serious violations of international human rights law and humanitarian law.

² PLEASE NOTE: All information is aggregated and de-identified to protect user privacy. All metrics are estimates, based on best information available for a specific point in time.

Implementation by-the-numbers

The Board has proof that Meta has fully or partially implemented 75 of the recommendations the Board has made since January 2021. This includes 34 recommendations that were fully or partially implemented since the publication of [last year's Annual Report](#). The table shows the full break-down of Meta's implementation of the 266 recommendations the Board had made by the time this report was finalized in May 2024.

Implementation Status	Total	Difference Compared to 2022 Annual Report
Implementation demonstrated through published information	45	+18
Partial implementation demonstrated through published information	30	+16
Progress reported	81	-3
Meta reported implementation or described as work Meta already does but did not publish information to demonstrate implementation	42	+13
Recommendation declined after feasibility assessment	15	+5
Recommendation declined	32	N/A ³
Recommendation omitted or reframed	16	N/A
Awaiting first response from Meta	5	+1

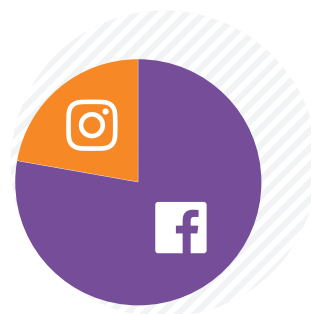
- 3 In April 2024, the Board's Implementation Committee split the category "Recommendation omitted, declined or reframed" into "Recommendation declined" and "Recommendation omitted or reframed." While "Recommendation declined" applies to recommendations which Meta understood and considered, "Recommendation omitted or reframed" highlights instances where the company has misunderstood the Board's recommendation.



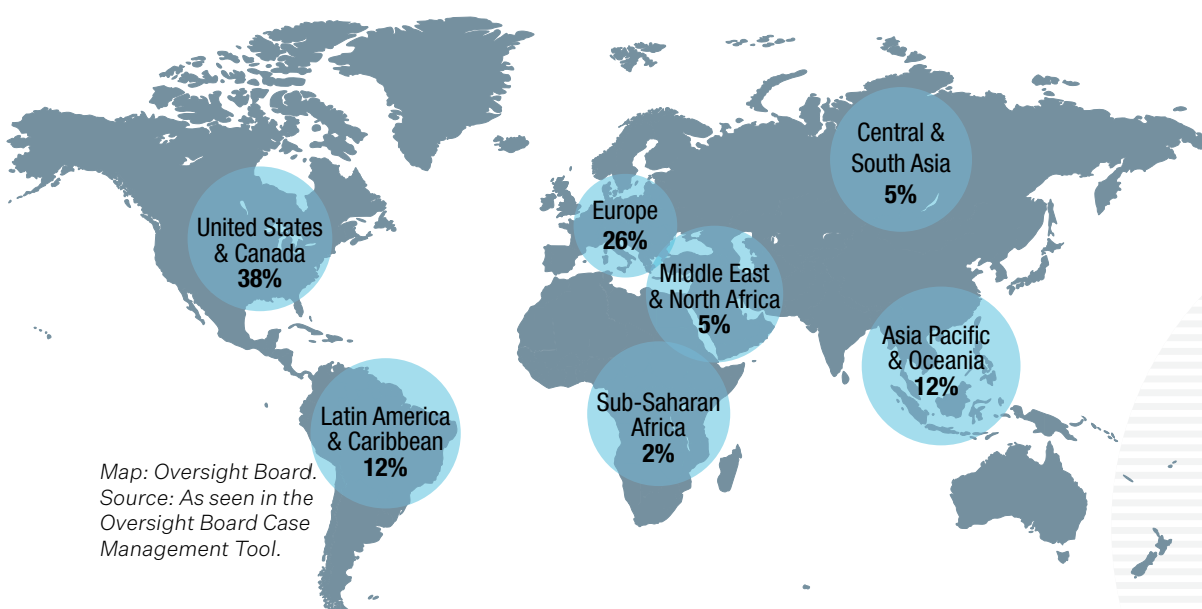
**RECOMMENDATIONS TO META
BY THE BOARD**

Appeals to the Board

In 2023, the Board received 398,597 appeals from around the world — including 39 cases referred by Meta. While we can only review a small number of cases, we continue to select cases that often raise underlying issues facing large numbers of people around the world and make recommendations to address them. More than three-quarters (78%) of appeals submitted in 2023 were about Facebook posts, while less than a quarter (22%) were about posts on Instagram.



Share of Cases Submitted to the Board by User-Selected Region in 2023



In 2023, 38% of cases submitted to the Board came from the US and Canada — a fall of nine percentage points in the share of cases from this region compared to 2022. 26% of submitted cases were from Europe — a rise of 4% compared to 2022. Asia-Pacific and Oceania represented 12% of cases (down from 13% in 2022), while Latin America and Caribbean made up 12% of cases (the same share as in 2022). The Middle East and North Africa represented 5% of cases (up three percentage points compared to 2022) as did Central and South Asia (up two percentage points compared to 2022). 2% of the Board’s appeals came from Sub-Saharan Africa (the same percentage as in 2022). While the share of appeals coming from the Global Majority is still limited, it is encouraging to see small increases for the Middle East and North Africa and Central and South Asia. In 2024, we will aim to increase awareness of our work in the Global Majority — including through advertising to users from specific countries on Facebook and Instagram, and regional events in Asia, Africa and South America.

How the Board's Decisions Have Considered International Human Rights Standards

RECOMMENDATIONS ON HOW META CAN IMPROVE POLICIES



As a Board, we believe that Meta will make decisions in a fairer, more principled way if it bases them on the international human rights standards to which it has committed itself.

We choose cases, issue decisions and provide recommendations to advance Meta's respect for the human rights of all people. In our decisions we provide a detailed analysis of the human rights implications and concerns relevant to the case. Below, we list illustrative examples of how we used international human rights standards in the decisions we published in 2023.

Gender Identity and Nudity

In these cases, we overturned Meta's decisions to remove two posts showing a transgender and non-binary couple bare chested but covering their nipples. While the International Covenant on Civil and Political Rights guarantees the right to freedom of expression to all people without discrimination as to "sex" or "other status," we found that Meta's policies on adult nudity result in greater barriers to expression for women, trans and non-binary people.

In response, we urged Meta to define clear, objective, rights-respecting criteria to govern its Adult Nudity and Sexual Activity Community Standard, so that all people are treated in a manner consistent with international human rights standards without discrimination based on sex or gender.

Cambodian Prime Minister

In this case, we overturned Meta’s decision to leave up a Facebook video in which the then-Prime Minister of Cambodia, Hun Sen, threatens his political opponents with violence. We concluded that leaving this content up was at odds with Meta’s human rights responsibilities. It represented a severe risk to the rights to vote and participate in public affairs, to peaceful assembly, to physical security and to life in Cambodia.

After assessing the relevant factors using the six-part Rabat Plan of Action, we found that removing the content was a necessary and proportionate limit on expression.

Armenian Prisoners of War

In this case, we upheld Meta’s decision to leave up, with a warning screen, a Facebook post that included a video showing identifiable prisoners of war. Our decision stressed the importance of social media companies preserving content depicting grave human rights violations or atrocity crimes, such as those specified under the Rome Statute of the International Criminal Court. In our view, this content informed the public and contributed to pressure on the detaining power to protect the prisoners’ rights. The decision to apply a warning screen to the post was necessary and proportionate, showing respect for the rights of prisoners and their families.

In our decision, we urged Meta to establish a new, consistent approach to retaining potential evidence of atrocities and serious violations of international human rights law. The company is currently implementing this recommendation.

The Value of a Human Rights-Based Approach

As a Board, we now have more than three years of practical experience of applying international human rights standards to a private company moderating the content of billions of people around the world. These standards have many benefits. They place freedom of expression and human dignity at the center of our analysis, offer a cross-cultural reference point and foster transparency, making our work part of an ecosystem of human rights stakeholders.

However, there are also challenges. For example, the kinds of harms found online are also different to in the real world. In our decisions, we have expressed concern at how the scale and speed of online content can create cumulative harms that would not exist offline. Our experience as a Board shows that an approach based on international human rights standards can be useful. Moving forward, we will continue to adapt our approach given the challenges above. This will be difficult, but immensely valuable, and we will rely on feedback from academics and civil society as we continue down this path.

What's Next – 2024 & Beyond

THE REGULATORY LANDSCAPE AND THE IMPORTANCE OF THE BOARD



From new tech regulation to elections and conflicts that will determine the world's future, 2024 is already seeing some big changes. As a Board, we will strive to make our work relevant to ongoing debates, including in the following areas:

- ➔ While the EU's Digital Services Act represents the most wide-ranging and potentially impactful piece of **regulation**, other markets such as the US, the UK, Brazil and India are pursuing different legislative approaches. As a Board, we have helped Meta tackle systemic risks to human rights associated with its content moderation. This has prevented important expression from being unnecessarily removed, including political speech in Iran and breast cancer-awareness content. We believe our work on systemic risk could also be of value to the wider industry. Our work identifies and helps to address systemic risks to free expression and other human rights, ensuring compliance with new regulation and, ultimately, building trust with the people who use online platforms.
- ➔ 2024 is already proving to be a year of **political and geopolitical uncertainty**. On top of the ongoing conflicts around the world, this year will see more than 50 national-level elections, with around half the world's population heading to the polls. In response, we will take on more cases in these areas, on an expedited basis where necessary.

-
- ➔ We will continue to **focus on issues that matter to people**. We have already expanded our scope to Threads and will soon start addressing the issues facing its millions of users. Moving forward, we will continue to expand our work to new areas to ensure that Meta treats users fairly. One area we are interested in exploring is demoted content, where a platform limits a post's visibility without telling the user.



IMPROVING HOW META TREATS PEOPLE & COMMUNITIES AROUND THE WORLD



www.oversightboard.com

© 2024 Oversight Board LLC