

Human rights report



Table of contents

About this report	03
Executive summary	06
Human rights risk management	09
1. Freedom of opinion and expression	10
2. Privacy	13
3. Equality and non-discrimination	14
4. Life, liberty and security of person	15
5. Best interests of the child	16
6. Public participation, to vote, and to be elected	16
7. Freedom of association and assembly	17
8. Right to health	17
Issues spotlight	18
Innovating with respect for human rights	18
Advancing accessibility through wearable technology	19
Continuing responsible AI innovation	19
Fraud and scams	22
Disrupting criminal scam syndicates in Southeast Asia	22
How we prepare for and respond to crises	24
Democratic Republic of Congo	25



Stakeholder engagement	26
Trusted Partners	27
Case study: Insights for youth recruitment prevention in Latin America	28
International organizations	29
Transparency and remedy	30
Annex	33
How human rights are governed and managed at Meta	34
Training Meta employees on human rights	34
Links to referenced reports	34





About this report

This annual Human Rights Report covers insights and actions from January 1, 2025, through December 31, 2025. We report on Meta services and products including Facebook, Messenger, Instagram, WhatsApp, Threads and Reality Labs.

It builds on Meta's work to respect human rights and reflects progress made on our commitments to the [United Nations Guiding Principles on Business and Human Rights](#) and our [Corporate Human Rights Policy](#). The report shows how we applied these principles across the company in 2025 and provides guidance on where to find more in-depth information.

The content in this report is grounded in our [Comprehensive Human Rights Salient Risk Assessment](#), which was undertaken in 2022. The purpose of the assessment was to identify and prioritize our most significant potential adverse human rights impacts¹ on people who use our products and others who may be affected by our actions. This report outlines these potential salient risks and examples of our actions and mitigations in 2025.

Human rights continued to be a topic of critical importance to our company and to our stakeholders in 2025. We have shortened and simplified the report while continuing to strive to provide a representative picture of our work across multiple teams and stakeholder engagement around the world.

Policies and progress

In addition to this Human Rights Report, Meta reports annually on policies and progress using the following mechanisms:



-  Annual Report
-  Proxy Statement
-  Responsible Business Practices Report
-  Transparency Center
-  Sustainability Report
-  CDP Climate Change Report
-  UN Global Compact

This report complements [Meta’s Responsible Business Practices Report](#). We [report](#) separately on our efforts to identify and mitigate the risks of modern slavery and human trafficking in our business operations and supply chains. In addition, we comply with mandatory national and European Union reporting, available in our [Transparency Center](#). Links to other Meta disclosures are in the [Annex](#) to this report.

[Go to Annex](#)

¹ The term “adverse human rights impact” is in line with the UN Guiding Principles on Business and Human Rights and means an impact that occurs when an action removes or reduces the ability of an individual to enjoy his or her human rights.



Our [Corporate Human Rights Policy](#) applies enterprise-wide. Each Meta service and entity has its own policies and procedures that may have different human rights impacts. This report references actions taken by Meta as a company regarding one or more Meta entities. Statements are not intended to imply that Meta took that same action regarding all entities and/or in all circumstances.²



² This report's discussion of content moderation and related actions on Facebook and Instagram does not apply to WhatsApp and, unless a policy or action is specified as applying to WhatsApp, it does not apply to WhatsApp. Further, while many actions described in this report apply to Facebook and Instagram, there are intentional distinctions in policies and procedures between the services. If a policy is labeled a "Facebook" policy, it may not apply to Instagram. No statement in this report is intended to create — or should be construed as creating — new obligations (legal or otherwise) regarding the application of a policy or procedure to other services or entities.



About

Executive summary

Managing risks

Issues

Stakeholders

Transparency and remedy

Annex

Executive summary



This is Meta’s fifth annual Human Rights Report. It provides insights into the work Meta did in 2025 to manage human rights risks at scale and live up to our commitments to the [United Nations Guiding Principles on Business and Human Rights](#).



Salient Risk Assessment

Our priorities in 2025 reflected the salient risks that were identified in the 2022 [Comprehensive Human Rights Salient Risk Assessment](#): freedom of opinion and expression; privacy; equality and non-discrimination; life, liberty and security of person; best interests of the child; public participation, to vote, and to be elected; freedom of association and assembly; and right to health.

[→ Read more](#)

Advancing innovation

Meta sought to push the boundaries of artificial intelligence (AI) innovation while embedding human rights considerations into our development process. We introduced new models, safety frameworks and products such as [Meta Ray-Ban Display glasses alongside Meta Neural Band](#). This wristband uses muscle signals for hands-free control of AI glasses. We developed these products through our responsible product process with [accessibility](#) in mind and in line with the right to equality and non-discrimination.

[→ Read more](#)

Addressing fraud and scams

We also expanded our efforts to address fraud and scams. We grounded [these efforts](#) in platform enforcement and cross-regional partnerships with industry, law enforcement and local civil society organizations. Many scam centers rely on human trafficking and forced labor to carry out their work; our mitigation efforts helped respect the right to life, liberty and security, as well as the prohibition on slavery, servitude and forced labor.

[→ Read more](#)

Child and youth safety

Likewise, our commitment to [child and youth safety](#) continued. Among other initiatives, we announced that Instagram was [revamping](#) the content experience on Teen Accounts, inspired by 13+ movie ratings criteria and parent feedback.* We also developed this updated experience in line with expert guidance and the principle of the evolving capacities of the child outlined in the [UN Convention on the Rights of the Child](#).

[→ Read more](#)

Crisis response

Our 2024 Human Rights Report gave detailed insights into how Meta integrates human rights principles into [how we prepare for and respond to crises](#), guided by our [Crisis Policy Protocol](#). This work continued in 2025. We applied the Crisis Policy Protocol to respond to a range of situations, from those in India (Kashmir) and Indonesia to youth-led “Gen Z protests” in Madagascar, Serbia and other countries. In this report, we provide a case study on our crisis response in the Democratic Republic of Congo.

[→ Read more](#)



Stakeholder engagement

Our [Corporate Human Rights Policy](#) supports our proactive engagement with stakeholders. In 2025, we engaged with a diverse range of human rights groups, vulnerable communities, civil society members, academics, think tanks and regulators. Key topics included Meta’s approach to AI and content moderation, the impact of our [January policy and product changes](#), and our efforts to combat fraud and scams.

We continued to engage our [Trusted Partners](#) around the world to help identify trends, and better understand the impact of online content and behavior on local communities. They provided insights and identified potentially violating content in more than 20 countries and regions.

[→ Read more](#)

Oversight Board

We also provide an update on the work of the [Oversight Board](#). The Oversight Board is an independent body that provides binding rulings and non-binding recommendations. During the reporting period, the Oversight Board considered cases concerning our efforts to respect human rights, including freedom of expression, the rights and interests of young people, and the right to equality and non-discrimination, among other topics.

As an illustrative example, in response to a recommendation, Meta [increased](#) the granularity of existing notifications — providing more detailed notice to users when their content is restricted in response to a formal government report stating the content violates local laws.

[→ Read more](#)

Managing government requests

Throughout 2025, we continued to be guided by our commitment to the [Global Network Initiative](#) to respect freedom of expression and privacy, including when responding to government requests to restrict content. Such government requests have been increasing. In addition to our regular integrity reporting, we published [case studies](#) related to government requests to restrict political speech in Austria, Colombia, Côte d’Ivoire, Germany, India, Iraq, Italy, Kazakhstan, Malaysia, New Zealand, Poland, Romania, Singapore, South Africa, Tanzania, Türkiye and Vietnam. In 2025, Meta also underwent its fourth GNI assessment (covering April 2022–December 2024). The GNI Board unanimously determined that Meta had made “good faith efforts to implement the GNI Principles, with improvement over time.”

[View case studies](#)



Human rights risk management

The [United Nations Guiding Principles on Business and Human Rights](#) make it clear that companies should identify their adverse human rights impacts in order to effectively prevent or mitigate them.

Given the scale of Meta's operations and the range of rights it could implicate, anticipating and managing our [salient risks](#) is important but complex. We manage two types of inherent risks: those stemming from our own activities and those stemming from the activities of third parties, including people who use our platforms.



The table below lays out our inherent salient human rights risks³ as defined in our 2022 Comprehensive Human Rights Salient Risk Assessment (CSRA), which we disclosed in our [2022 Human Rights Report](#). This table provides illustrative examples of how we addressed the potential risks in 2025. Later in this report, we take a deeper dive into some of these examples and how we managed potential risks in relation to artificial intelligence (AI), elections and conflicts.

1. Freedom of opinion and expression

The [right to freedom of opinion and expression](#) includes the right to seek, receive and share information and ideas of all kinds. It is a foundational right, essential for the protection of human dignity, individual autonomy and democracy. Freedom of expression is a core part of our mission, consistent with our value of giving everyone a voice.

Examples of potential inherent salient human rights risks identified in CSRA	Examples of how Meta addressed the potential risks in 2025
Meta’s content moderation policies and enforcement may limit freedom of expression.	In January 2025, we updated our Hateful Conduct Policy to allow for more discussion on some topics that are part of mainstream discourse, including immigration and gender. We focused our automated systems on tackling illegal and high-severity violations, like terrorism, child sexual exploitation, drugs, fraud and scams, while relying on user reports before taking action for less severe policy violations. We also tuned our systems to generally require a much higher degree of confidence before a piece of content is taken down. Later in 2025, we reported on our enforcement precision progress in our Community Standards Enforcement Report, as part of our Integrity Reports. We ended changes made in 2021 to reduce the amount of civic content people see and began recommending more political content based on personalized signals, e.g., liking a piece of content.

³ For the distinction between “inherent” and “residual” risk, see [Meta’s 2024 Human Rights Report](#), page 11.



Examples of potential inherent salient human rights risks identified in CSRA

Meta's content moderation policies and enforcement may limit freedom of expression. (cont.)

Examples of how Meta addressed the potential risks in 2025

We ended our third-party fact-checking program in the U.S. and instead began moving to a [Community Notes](#) program. In the rest of the world, we continued to rely on [fact-checkers](#) who are independent from Meta and certified through the non-partisan [International Fact-Checking Network](#) or, in Europe, the [European Fact-Checking Standards Network](#) to address misinformation on Facebook, Instagram and Threads.

→ [Read more](#)

Overbroad government limits on content

We strive to follow our [Global Network Initiative](#) commitments. In 2025, Meta underwent its fourth GNI assessment (covering April 2022–December 2024). The GNI Board unanimously determined that Meta had made “good faith efforts to implement the GNI Principles, with improvement over time.” We also reported on our [responses](#) to government requests for data or content restrictions ([here](#) and [here](#)). In 2025, noteworthy transparency [case studies](#) involving political speech included those in Austria, Colombia, Côte d'Ivoire, Germany, India, Iraq, Italy, Kazakhstan, Malaysia, New Zealand, Poland, Romania, Singapore, South Africa, Tanzania, Türkiye and Vietnam.



Launching Community Notes

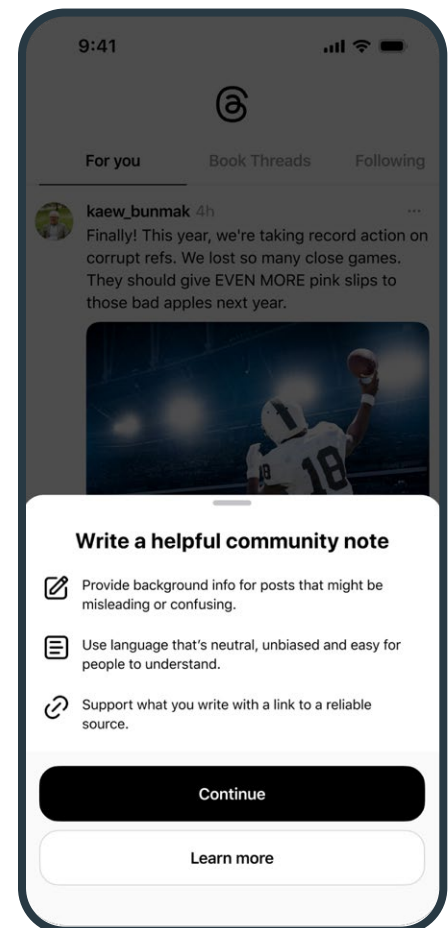
In 2025, Meta launched [Community Notes](#) in the U.S. Community Notes is a new tool for the community to add context to content that may be misleading or confusing. Our approach was informed by human rights principles, including freedom of expression, and aimed at improving access to information broadly trusted by our community of users through proportionate, speech-protective interventions.

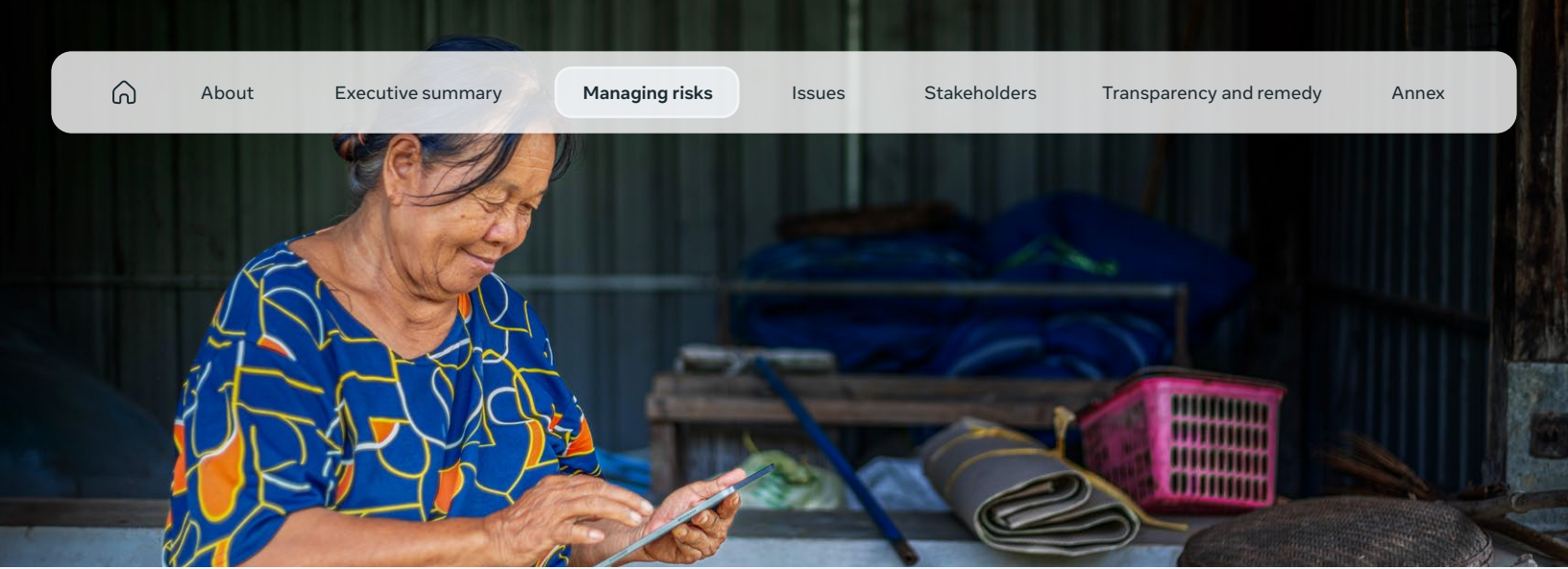
Notes are written and rated by eligible users — those 18 or older with a security-verified account in good standing. A note is only shown when it reaches a “helpful consensus” threshold, meaning enough contributors who historically disagreed now agree the note is helpful. Author names are hidden to reduce peer pressure and retaliation risks and to encourage participation, including around sensitive topics.

Community notes can advance freedom of expression by adding context rather than removing or down-ranking lawful speech, and by widening participation beyond a small set of professional fact-checkers. It is also designed with safeguards to help prevent abuse. These include account integrity requirements, spam and URL protections, and enforcement of Community Standards on notes.

Meta requested the Oversight Board’s guidance on the country-level factors we should consider when deciding whether a country should be omitted from our Community Notes expansion, including low levels of freedom of expression, absence of freedom of the press and levels of digital literacy.

In November 2025, the Oversight Board accepted [Meta’s request for a Policy Advisory Opinion](#) to help inform our approach to expanding the Community Notes program outside of the U.S. The Oversight Board published its [advice](#) in March 2026.





2. Privacy

The [right to privacy](#) is a necessary condition for the realization of other human rights, such as freedom of expression, freedom of assembly and association, and freedom of belief and religion. One of the core principles outlined in our [Corporate Human Rights Policy](#) is to keep people safe and protect privacy.

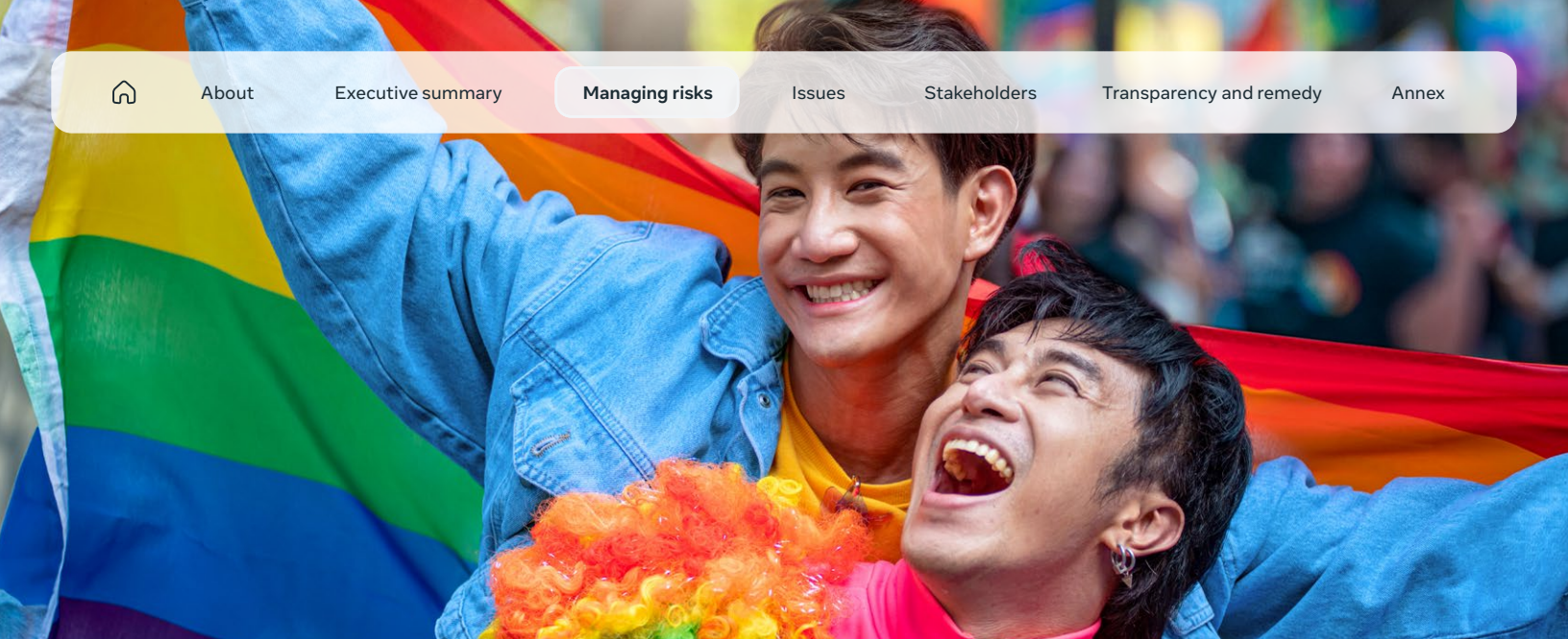
Examples of potential inherent salient human rights risks identified in CSRA

Content or behavior on Meta apps may adversely impact privacy and data protection rights.

Examples of how Meta addressed the potential risks in 2025

In 2025, we developed [Private Processing for Meta AI on WhatsApp](#), which is built on top of a Trusted Execution Environment. Private Processing allows users to direct Meta AI to process their requests in a secure and private cloud environment where no one, not even Meta or WhatsApp, can access their personal messages. This technology now powers certain optional Meta AI features on WhatsApp, such as summarizing messages and providing writing help. Once Private Processing has finished responding to a request, the messages won't be stored. Read more about this work in our [Privacy Progress Update](#).

We filed a [lawsuit](#) against the entity behind CrushAI or “nudify” apps, which allow people to create AI-generated nude or sexually explicit images of individuals without their consent. We further [announced](#) additional steps we took to clamp down on these apps. This work complemented our longstanding efforts to help prevent intimate image abuse, whether real or AI-generated, from spreading online, through tools like [StopNCII.org](#) and the [National Center for Missing and Exploited Children's Take It Down](#).



3. Equality and non-discrimination

The [right to equality and non-discrimination](#) provides for equal protection against discrimination. As part of respecting this right, we don't allow hateful conduct on our platforms, as defined in our [policy](#).

Examples of potential inherent salient human rights risks identified in CSRA

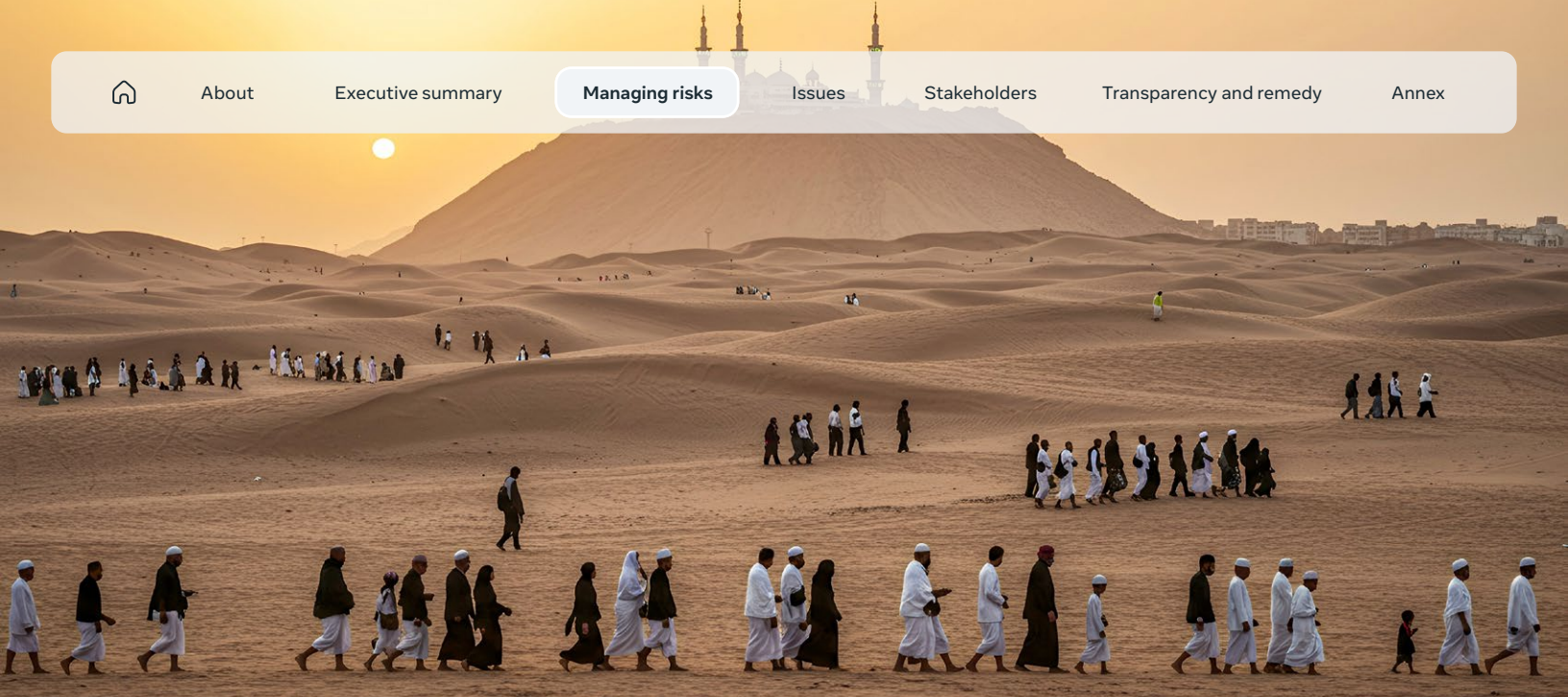
Some languages and dialects may be more challenging to moderate than others.

Examples of how Meta addressed the potential risks in 2025

We [introduced](#) and open sourced [Meta Omnilingual Automatic Speech Recognition \(ASR\)](#), a suite of models providing ASR capabilities for more than 1,600 languages, including 500 low-resource languages. By making spoken content transcribable across previously underserved languages, this technology opens up new education opportunities and further enhances freedom of expression, empowering people to access knowledge and share content in the language they know best.

Content adversely impacting equality and non-discrimination (e.g., hateful conduct)

Ahead of the [policy and product changes](#) announced in January 2025, we assessed the risks of the changes and took into account relevant mitigations, such as the availability of other policies and user reports, to address this type of content. In countries experiencing conflict or crises, we may apply proactive detection by default in some policy areas, including Hateful Conduct violations. For example, we used proactive detection in Libya after deploying the Crisis Policy Protocol.



4. Life, liberty and security of person

The [right to life, liberty and security of person](#) concerns freedom from physical harm and confinement. For Meta, respecting this human right includes mitigating the risk that content may provoke harm, including risks of violence and human trafficking, state-sponsored online threats, and non-state groups engaged in, or advocating for, violence or hate.

Examples of potential inherent salient human rights risks identified in CSRA

Bad actors who:

- Exploit Meta services and apps to coordinate online or offline harm
- Misuse services and apps for cyber-attacks or phishing
- Threaten and harass human rights defenders, activists and other vulnerable groups

Examples of how Meta addressed the potential risks in 2025

Through deep research and partner consultations, Meta identified a potential heightened risk of child soldier recruitment in Sudan and developed youth safety messaging in Arabic, French and English customized for the region. In November 2025, Meta launched two complementary campaigns: a quick promotion delivered to approximately 1 million Sudanese youths (aged 13–18) on Facebook and Instagram, and an interactive custom message pop-up delivered to approximately 40,000 youths connected to [Strategic Network Disruptions](#). Messages warned about recruitment risks and directed the youths to external safety resources.

→ [Read more](#)

We continued supporting the Human Rights Defender Fund and our [Trusted Partner program](#) to improve emergency response for human rights defenders and other vulnerable people.

5. Best interests of the child

The UN Convention on the Rights of the Child (UNCRC) states that in all actions concerning children, “the best interests of the child shall be a primary consideration.” Meta’s [Best Interests of the Child Framework](#) aligns with the fundamental values of the UNCRC. Online child protection is a top priority for Meta. We offer tools for teens, parents and guardians with built-in protections to help keep them safe while providing space for them to exercise their right to freedom of expression and access to information.

Examples of potential inherent salient human rights risks identified in CSRA

Children can be exposed to unwanted, inappropriate content or predatory behavior.

Examples of how Meta addressed the potential risks in 2025

Building on [Teen Accounts](#), we added [new protections](#) and unveiled [updated content settings](#), inspired by 13+ movie ratings criteria and parent feedback. This means Instagram’s goal is for teens to see content that’s similar to what they’d see in age-appropriate movies by default, and we introduced a new, stricter setting for parents who prefer a more restrictive experience for their teen. We also announced new controls to help parents manage how their [teens interact with AI](#) on our apps, [took legal action](#) against companies behind “nudify” apps and [partnered with schools to address cyberbullying](#).

6. Public participation, to vote, and to be elected

The [right to public participation, to vote, and to be elected](#) in free and fair elections is a cornerstone of democracy. Protecting the integrity of elections on our services and apps is one of our highest priorities. We work hard to protect elections online before, during and after election periods.

Examples of potential inherent salient human rights risks identified in CSRA

Violating content can adversely impact public participation, voting or running for office. This can stem from activities including, but not limited to:

- Coordinated bad actors interfering with elections
- Threat of offline harm and violence to candidates
- Individual efforts to prevent people from voting, increases in spam, foreign interference or reports of content that violates our policies

Examples of how Meta addressed the potential risks in 2025

We continued our election efforts, building upon our scaled and risk-based election approach detailed in our [2024 Human Rights Report](#).

In 2025, our election preparations included those in Australia, Cameroon, Chile, Côte d’Ivoire, Ecuador, Egypt, Germany, Iraq, Japan, Myanmar, the Philippines, Poland, Romania and Tanzania, among other countries.

7. Freedom of association and assembly

The [right to freedom of association and assembly](#) is essential to democracy and interdependent with many other rights guaranteed under international human rights law, including the right to freedom of expression and to take part in the conduct of public affairs. For Meta, this right relates to our core values of giving people a voice and building connection and community.

Examples of potential inherent salient human rights risks identified in CSRA	Examples of how Meta addressed the potential risks in 2025
Certain types of content on Meta platforms may lead some people to feel unable to freely connect on Meta apps or offline.	We deployed our Crisis Policy Protocol to help resource our efforts to address violating content regarding mass demonstrations, including “Gen Z protests,” for example in Iran, Madagascar, Nepal and Serbia.

8. Right to health

The [right to health](#) is the right of everyone to the highest attainable standard of physical and mental health. Meta respects this right by increasing access to health information, enabling people with similar health issues to connect with one another, and empowering people to make informed decisions about their health and well-being.

Examples of potential inherent salient human rights risks identified in CSRA	Examples of how Meta addressed the potential risks in 2025
Policy-violating content that incites or is intended to cause offline harm	We announced protections for teens in our AI offerings, including designing them to respond safely to prompts about self-harm, suicide and disordered eating. We also announced new controls including launching insights to parents about topics their teens are chatting with AI about. In the U.S., we further sought to increase our efforts to address online bullying in schools by giving teachers, educators and administrators a prioritized system for reporting instances of teen safety issues.

Issues spotlight

Innovating with respect for human rights

Innovations in wearable technology and artificial intelligence (AI) provided important opportunities to advance human rights and support responsible risk mitigation.

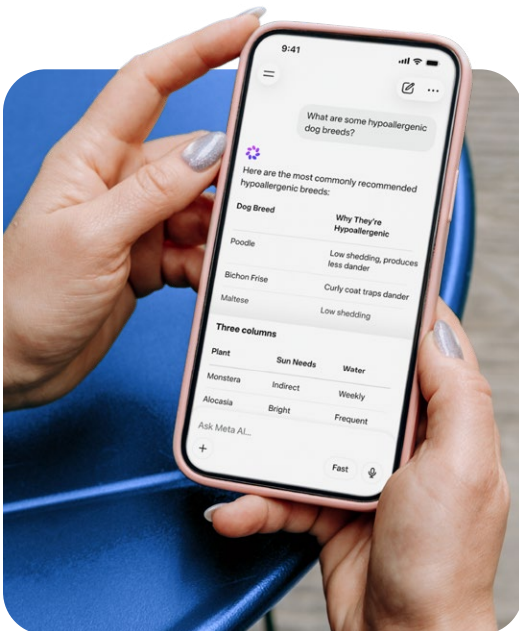
Advancing accessibility through wearable technology

In 2025, we [launched Meta Ray-Ban Display and Meta Neural Band](#) — a combination of next-generation AI glasses and a wristband that uses [surface electromyography \(sEMG\)](#) technology to convert electrical signals created by subtle movements in your muscles into commands allowing hands-free control.

Prototypes of the wristband were developed with [accessibility](#) in mind and in line with the right to equality and non-discrimination. Based on research involving nearly 200,000 participants, the wristband enables [new possibilities](#) for individuals with mobility limitations due to spinal cord injuries, strokes or other disabling events, as well as those who experience tremors or have fewer than five fingers. To protect privacy, the machine learning that translates EMG signals into digital commands is [processed locally on-device](#).



Together with AI glasses, our technology [empowers people with disabilities](#) to navigate daily life with greater independence. Features such as hands-free calling, real-time translation and live captioning break down communication barriers, while our “[call a volunteer](#)” feature — created in partnership with [Be My Eyes](#) — allows blind and low-vision individuals to connect with volunteers for real-time descriptions of their surroundings. In building these tools, we engaged directly with disability stakeholders, including the [Blinded Veterans Association](#) and U.S. Department of Veterans Affairs’ [Blind Rehabilitation Centers](#), to develop training guides and support accessibility use cases.



Continuing responsible AI innovation

We released the [standalone Meta AI assistant](#) alongside an updated open-source large language model, [Llama 4](#). We also launched [Private Processing](#), a new [technology leveraging a trusted execution environment](#), to enable Meta AI assistant on WhatsApp while preserving privacy protections, including encrypted communications. For details, see our prior due diligence that discusses the [human rights benefits of end-to-end encryption](#).

To mitigate risks of misuse and potentially harmful outputs from our AI products and models, the development of Llama 4 included pre-training data filtering, post-training mitigations and system-level mitigations. These are outlined in our [developer use guide on AI protections](#).

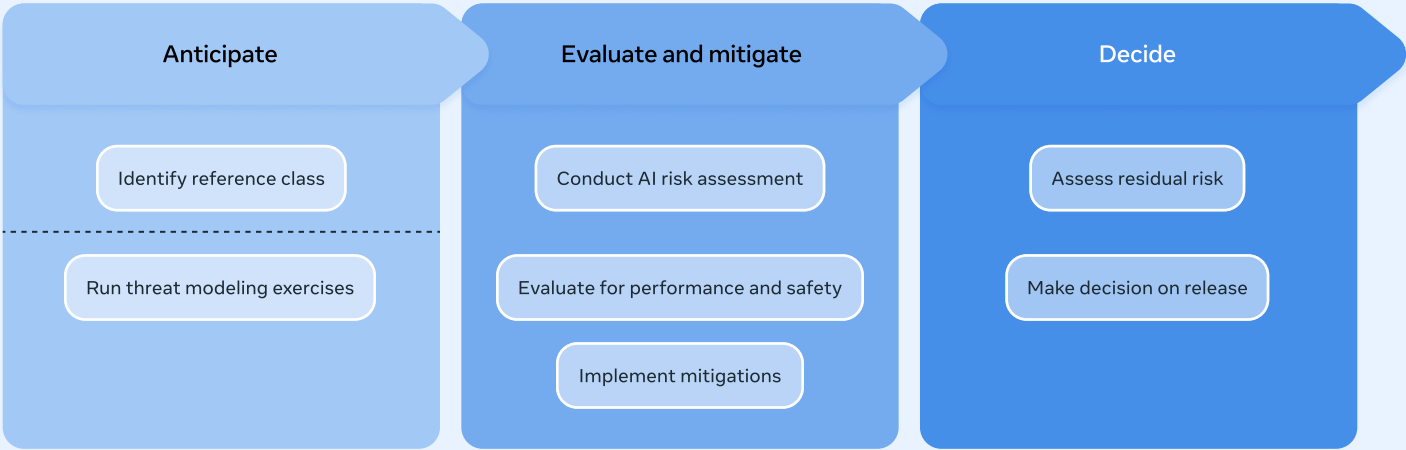
We also published our [Frontier AI Framework](#), focused on the most critical risks in the areas of cybersecurity threats and risks from chemical and biological weapons. Our framework outlined a number of processes we follow to anticipate and mitigate risk when developing frontier AI systems, including:

Identifying catastrophic outcomes to prevent 	Our framework identified potential catastrophic outcomes related to cyber, chemical and biological risks that we strive to prevent. It focused on evaluating whether these catastrophic outcomes were enabled by technological advances and, if so, identifying ways to mitigate those risks.
Threat modeling exercises 	We conducted threat modeling exercises to anticipate how different actors might seek to misuse frontier AI to produce those catastrophic outcomes, working with external experts as necessary. These exercises were fundamental to our outcomes-led approach.
Establishing risk thresholds 	We defined risk thresholds based on the extent to which our models could facilitate the threat scenarios. We have processes in place to keep risks within acceptable levels, including applying mitigations.

We also continued to support a broad ecosystem for responsible AI development and deployment. To empower developers to build safe applications, we expanded our investment in [system-level safeguards](#), including [Llama Guard 4](#) for unified safeguards across text and image modalities and [Llama Prompt Guard 2](#) for improved jailbreak and prompt injection detection.

Our work was informed by wide-ranging engagements and consultations with civil society, international organizations, academia and the public on the human rights dimensions of AI. This included [Community Forums](#) — conducted in partnership with Stanford University's Deliberative Democracy Lab — that gathered diverse perspectives on the future of AI agents and their governance.

2025 Frontier AI framework





Open-source well-being tools for content review

In order for our reviewers to apply our policies accurately, they must also focus on their [resilience and well-being](#).

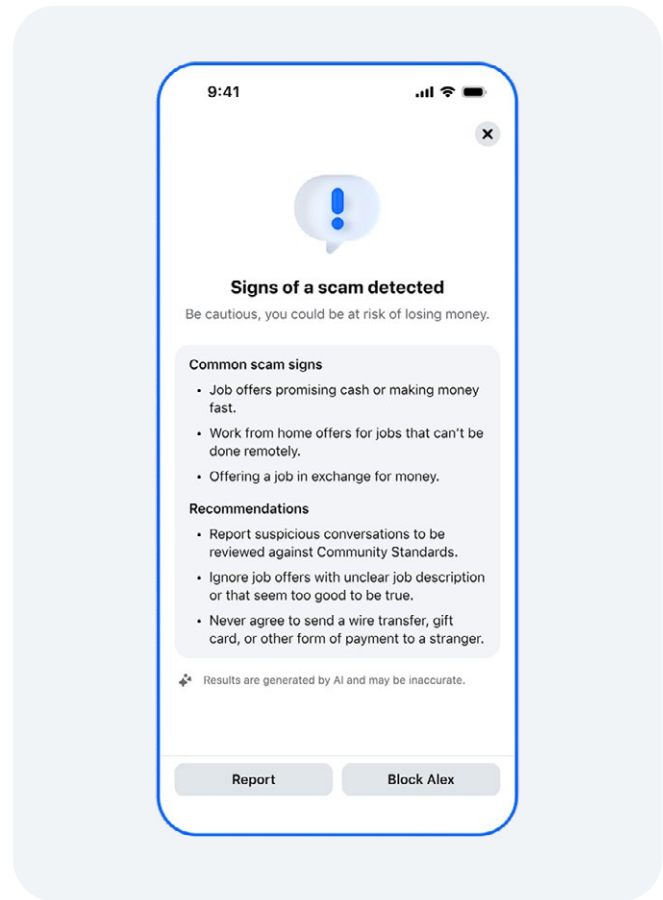
In 2025, we [open sourced a suite of content review filters](#) designed to support moderator well-being so other developers can benefit from our experience in this space. These customizable image and video filters — with features such as blurring, grayscale/sepia conversion, detail reduction and AI-enabled predictive warning screens — can help reduce the potential adverse mental health impacts of exposure to graphic or disturbing content.

Fraud and scams

Fraud and scam networks have grown considerably in the last decade and have had a negative impact on people's human rights. Many scam centers rely on human trafficking and forced labor to carry out their work, and scam victims can experience considerable financial loss and other harms. In 2025, we [removed](#) over 159 million scam ads for violating our policies, and we also [took down](#) 10.9 million accounts on Facebook and Instagram associated with criminal scam centers.

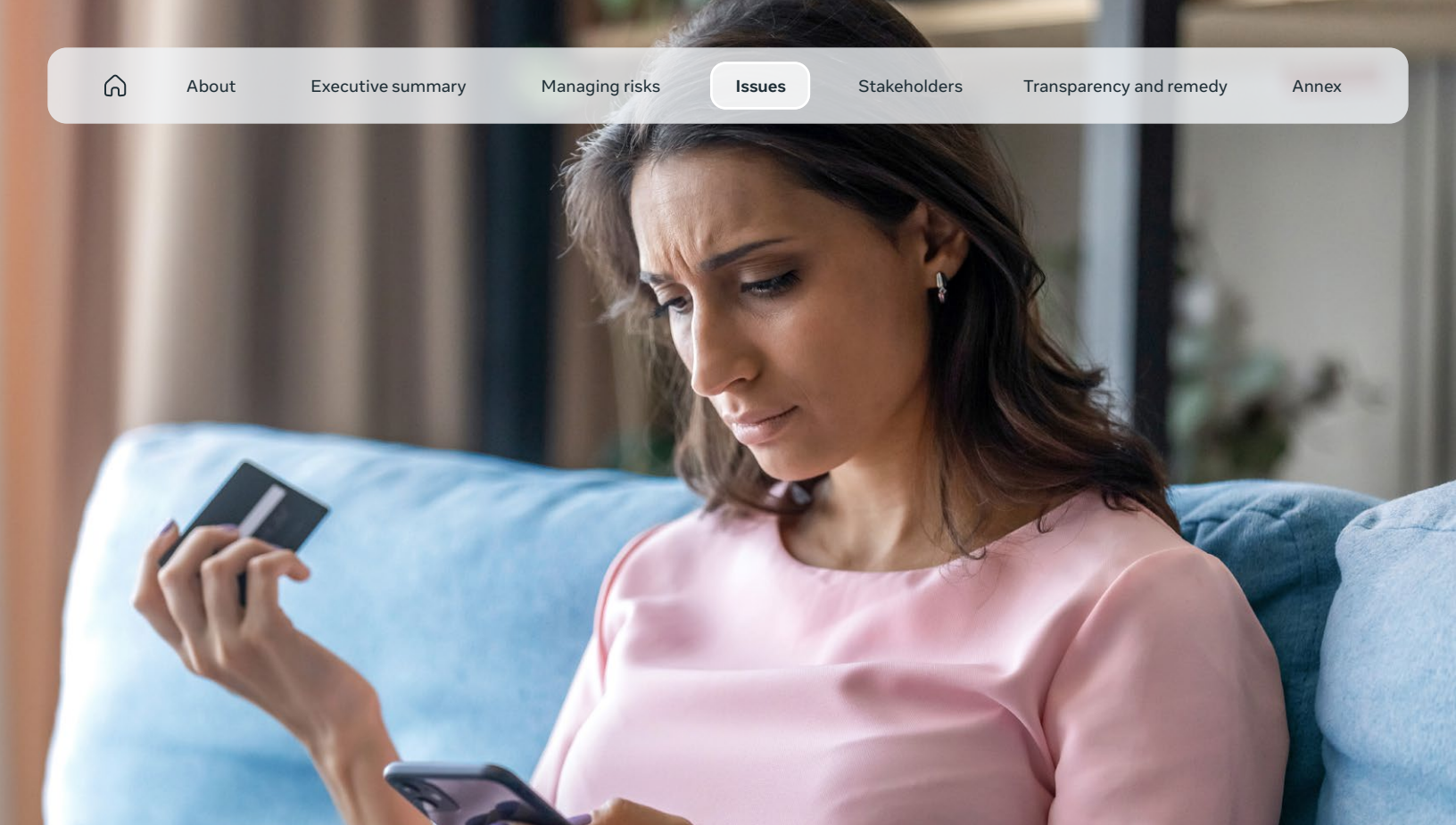
We prioritized safety through [platform defenses](#) such as [facial recognition](#) to combat impersonation, [scam detection warnings](#) on Messenger and [screenshare warnings](#) on WhatsApp. Our [work](#) disrupting scammers also involved litigation and supporting law enforcement to hold them accountable. Balancing the safety and well-being of our communities with proportionate, rights-respecting mitigations can be challenging. We took these actions with privacy in mind and in line with international human rights standards, including the [Global Network Initiative Principles](#).

Fraud and scams are global challenges. Signal sharing, collaboration across defender communities and a whole-of-society approach are necessary for effective work. Throughout 2025, Meta engaged with signal sharing partnerships such as the [Global Signal Exchange](#), [Global Anti-Scam Alliance](#) and [Fraud Intelligence Reciprocal Exchange](#). We also [advocated](#) for publishing lists of countries that fail to take adequate action and prompting international penalties, while being mindful of human rights and unintended consequences for vulnerable populations, including trafficked individuals forced into scam operations.



Disrupting criminal scam syndicates in Southeast Asia

Industrial-scale, sophisticated scam networks and compounds have [emerged](#) over the past decade. Across [the Asia Pacific region](#), these compounds employ hundreds of thousands of people and operate with near impunity. Criminal scam hubs lure job seekers with deceptive job postings and force them to work as online scammers, often under the threat of physical abuse — implicating the prohibition on slavery, servitude and forced labor under international human rights standards.



We [removed](#) over 6,400 Facebook accounts and Pages between January and October 2025 for [violating](#) our policies against [Fraud, Scams and Deceptive Practices](#) and [Dangerous Organizations and Individuals](#). One disrupted network originating in Cambodia [impersonated](#) global law enforcement to falsely offer loss-recovery services. This work required cross-border collaboration and closely monitoring the impact of geopolitical changes such as conflict in Myanmar and along the border in northern Cambodia. We discuss these efforts in more detail in our H1 2026 [Adversarial Threat Report](#).



Our 2025 disruption efforts culminated in our participation in a [Joint Disruption Week](#) pilot program in Bangkok, where Meta collaborated with the FBI, Royal Thai Police and other law enforcement agencies. This pilot [resulted in](#) the removal of 59,000 accounts, Pages and groups from Meta's platforms and six arrest warrants.



In addition to disruptions, we updated our [Community Standards on Human Exploitation](#) to capture more potential criminal scam syndicates' recruitment content. The update restricts job recruitment posts in locations that are high risk for labor exploitation as confirmed by law enforcement and local non-governmental organizations.

Meta also partnered with the United Nations Office on Drugs and Crime on the [#TrappedInScamCrime](#) campaign to help counter online recruitment and trafficking of persons for forced criminal activity linked to Southeast Asian scam centers. The campaign amplified survivor-informed prevention messaging to help people recognize deceptive job offers and access support resources.

We further engaged with the Association of Southeast Asian Nations to develop [principles](#) for tackling online scams. This work was in addition to our broader [anti-trafficking efforts](#).

How we prepare for and respond to crises

We prepare for and respond to many crises, including conflicts, intra-communal violence, civil unrest, mass protests and environmental disasters, as well as terrorist attacks and shootings, around the world. In 2025, we initiated and coordinated crisis responses in Cameroon, India (Kashmir), Indonesia, Libya, Madagascar, Nepal, Serbia, Tanzania and Thailand-Cambodia, among other countries and territories. We continued our efforts for designated crises under the [Crisis Policy Protocol](#) for the conflicts in Ukraine, Sudan and the Middle East. We address these designated crises in our [previous Human Rights Reports](#).

During a period of crisis, we rely on the Crisis Policy Protocol as a key tool. The protocol guides our expedited use of levers to mitigate potential harm as described in our [2024 Human Rights Report](#). The specific types of responses deployed are consistent with observed risks. They are informed by past crisis interventions, human rights principles and the law of armed conflict.

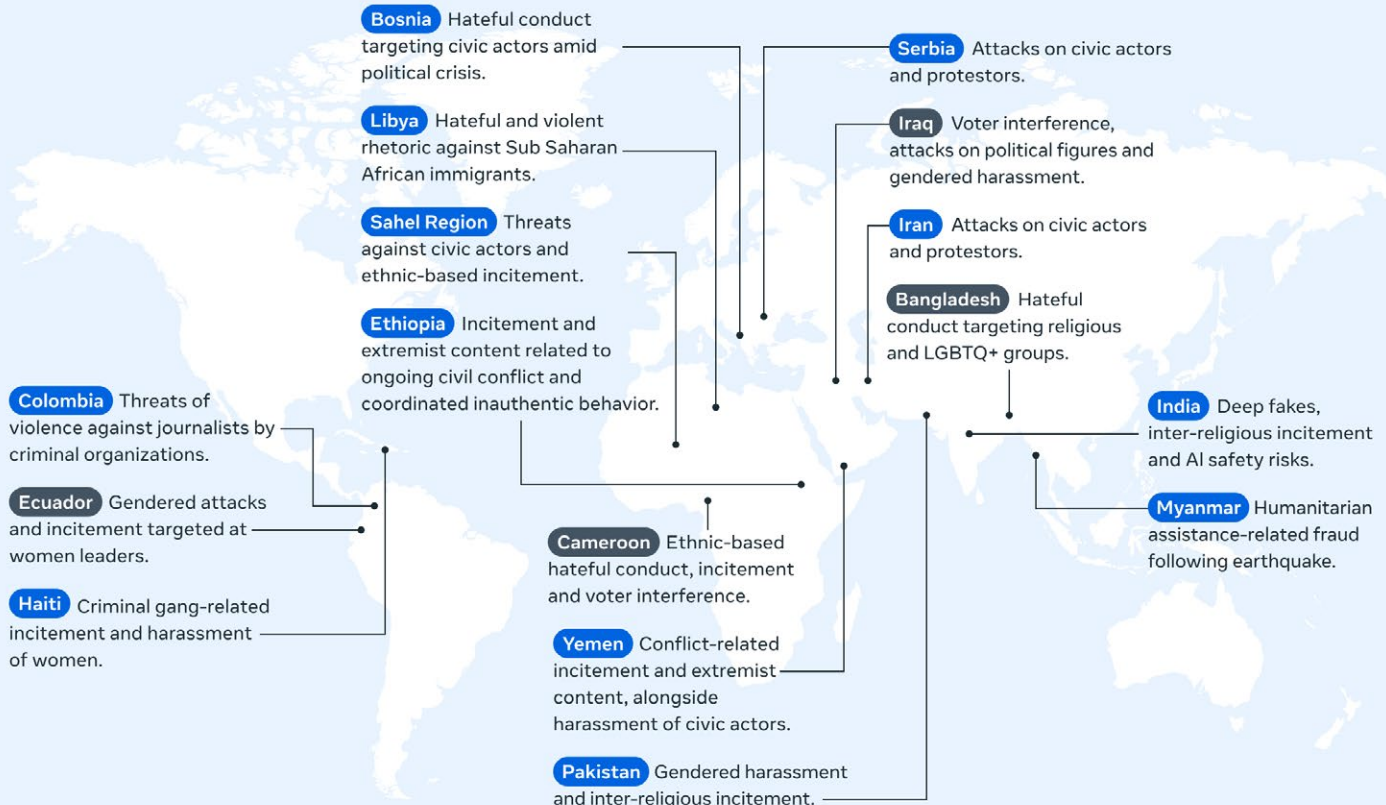
On the next page, we illustrate how we prepare for and respond to crises and conflict-affected and high-risk areas, using an example from a region not featured in our previous reports.

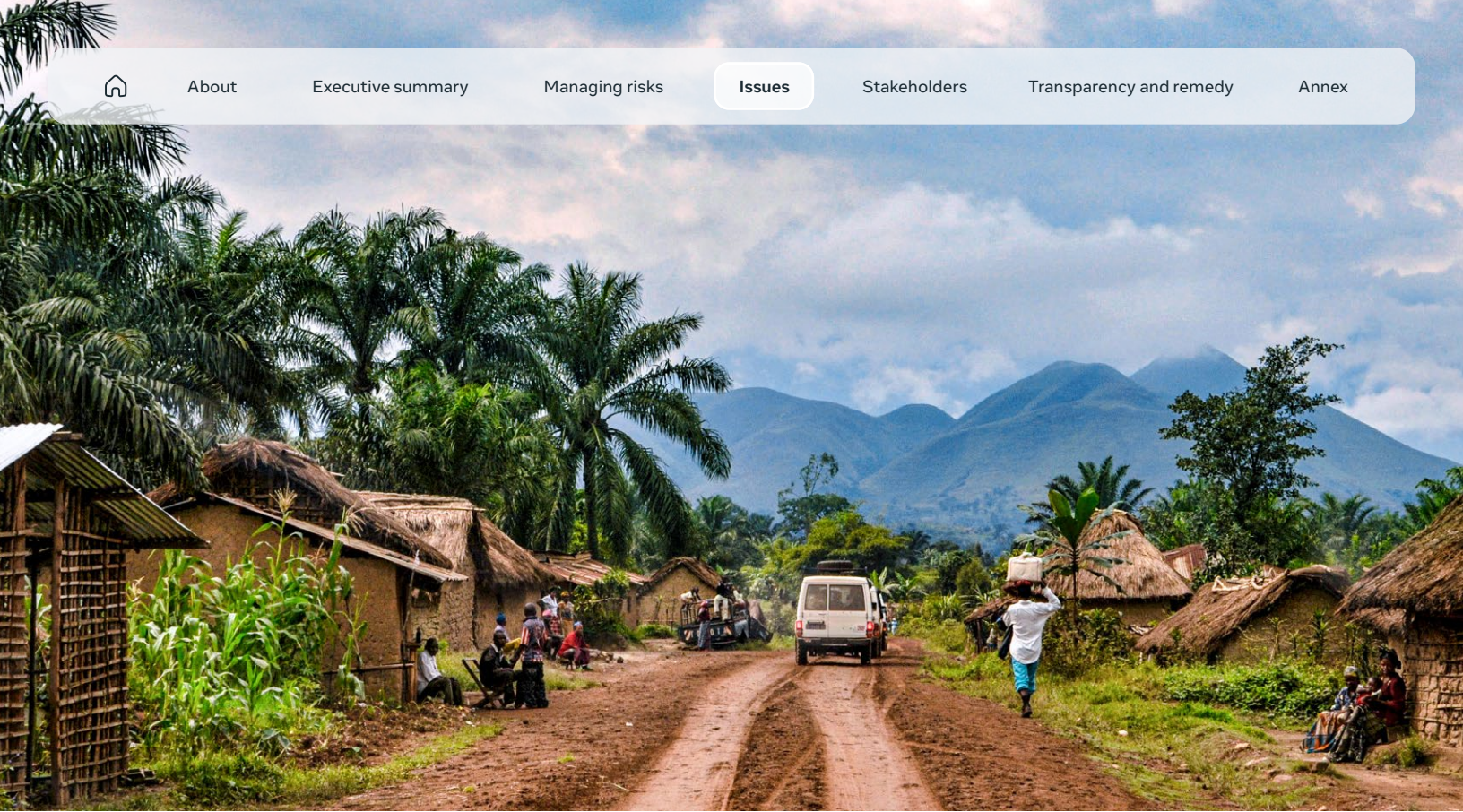
Spotlight on crisis response and elections integrity in 2025

Legend

Crisis Response

Election





Democratic Republic of Congo

In 2025, fighting in eastern Democratic Republic of Congo (DRC) intensified as the M23 armed group, reportedly [backed](#) by Rwanda, seized territory in North Kivu, driving mass displacement of civilian populations and [serious human rights violations](#). Trusted Partners alerted us to escalating violence in the East of the country and rising risks of inter-communal violence, leading Meta to designate the crisis under our Crisis Policy Protocol.

A cross-functional team of experts from Meta, spanning policy, operations, security and public policy, had been monitoring the offline situation in DRC since the resurgence of the conflict in late 2021, enforcing against violating content under our [Misinformation and Harm](#), [Violence and Incitement](#), [Coordinating Harm and Promoting Crime](#), and [Dangerous Organizations and Individuals](#) policies. We had also deployed mitigating measures to address potential online content risks, including the designation of pre-reviewed potentially harmful claims under the Misinformation and Harm policy and the use of our [Temporary High-Risk Locations policy](#). Additionally, we had previously designated outing risk groups related to “Rwandaphones, Rwandans, Banyamulenge or Tutsis” to help prevent the risk of offline harm.

As the conflict escalated in 2025, we assessed and mitigated related-content risks. We provided policy clarifications and guidance, reviewed keywords to ensure effective detection and precise enforcement of our policies, and applied our Temporary High-Risk Locations policy.

Meta also consulted with Trusted Partners to help inform the potential designation of new pre-reviewed harmful claims — false or misleading statements that might directly contribute to a risk of imminent physical harm under our Misinformation and Harm policy. Trusted Partner insights enabled us to more effectively detect and remove policy-violating content, for example material likely to inflame local divisions.

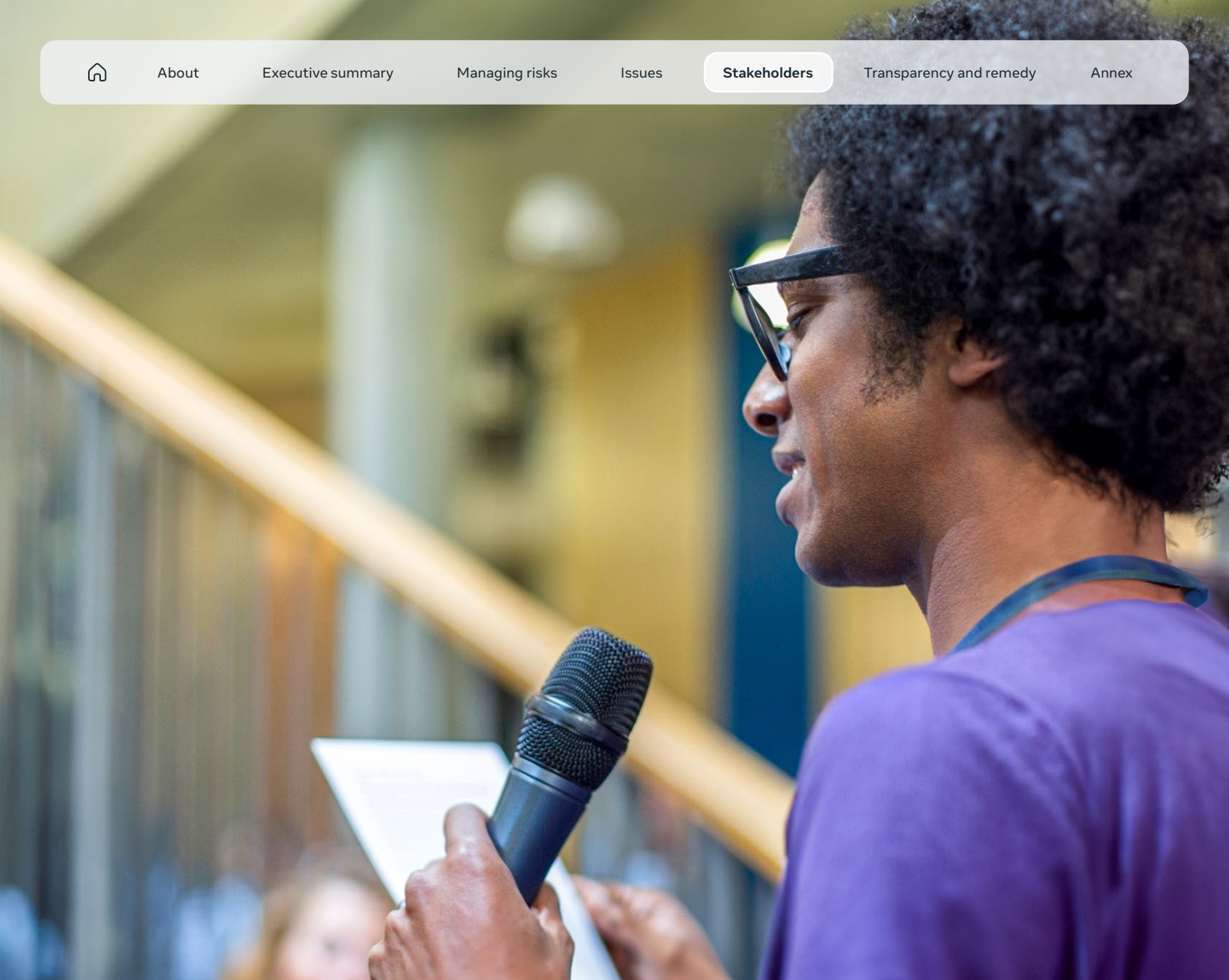
Stakeholder engagement

Proactive, structured [engagement](#) with our global community of users helps shape Meta policies and is central to our human rights risk management.

We engaged with a broad range of stakeholders in 2025, including civil society members, academics, think tanks, human rights experts and regulators. Key policy questions included our approach to artificial intelligence (AI) and content moderation and the impact of the product and policy changes introduced in January 2025, including [Community Notes](#), as well as fraud and scams.

We engaged with over 615 external stakeholders from 65 countries to inform policy development work in several areas including Dangerous Organizations and Individuals as well as health and wellness advertising policies. We consulted with hundreds more for risk forecasting, trend analysis, enforcement and prevention workstreams.

In addition, we engaged with civil society on a range of freedom of expression and regulatory issues, including in Argentina and Brazil, as well as in Malaysia as part of the [Digital Rights Asia Pacific Assembly Conference](#). We also conducted training on our policies for civil society. In North Africa, for example, we trained civil society groups on our [Misinformation and Harm](#) and [Hateful Conduct](#) policies to help them identify potentially violating content under those policies. The training also helped equip human rights defenders and vulnerable groups with tools to protect their digital presence.



Trusted Partners

Trusted Partners are important allies in identifying high-severity violations of our Community Standards. We continued to engage our [Trusted Partners](#) to identify trends, better understand the impact of online content and behavior on local communities, and explore how we can strengthen our civil society escalation channels.

In 2025, Trusted Partners provided insights and identified harmful content in countries and territories that experienced heightened unrest or elections. These included Afghanistan, Cambodia, Côte d'Ivoire, Democratic Republic of Congo, India, Indonesia, Israel, Kenya, Lebanon, Madagascar, Mexico, Nigeria, Palestine, Somalia, Sudan, Syria, Thailand, Tanzania, Ukraine and Venezuela, among other countries and regions.

We removed more than 60,000 pieces of policy-violating content as a result of our Trusted Partner program. An example of Trusted Partner impact in Latin America can be found on the next page.

CASE STUDY

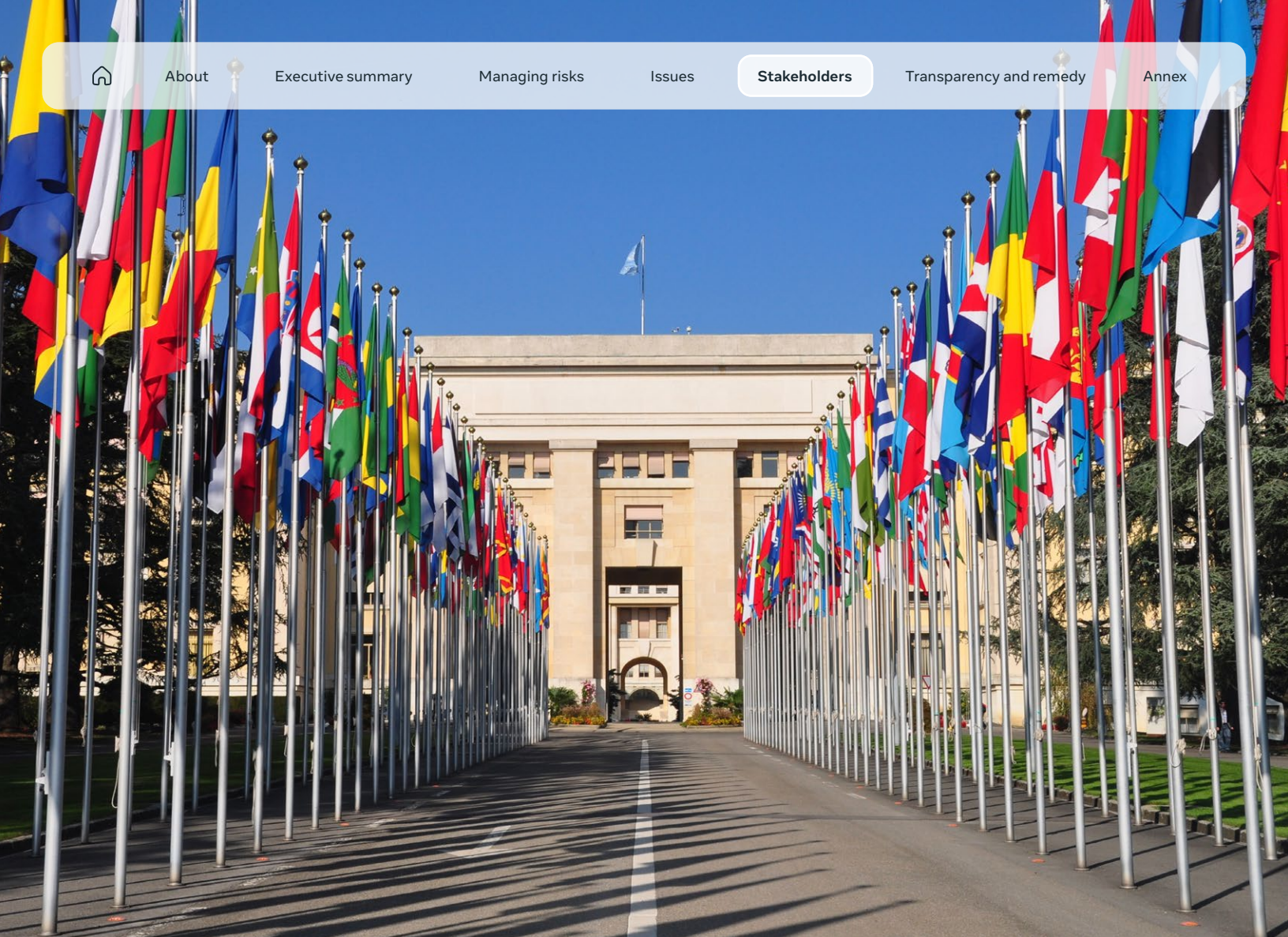
Insights for youth recruitment prevention in Latin America



In Latin America, Trusted Partners played a critical role in reporting content including glorification, support and representation of [dangerous organizations and individuals](#). They also provided feedback on youth recruitment prevention initiatives.

In Brazil, Trusted Partners reported content on Instagram portraying members of a designated criminal organization, including depictions of prohibited firearms, acts of violence and drugs. In Colombia, Trusted Partners reported profiles representing criminal groups, in the context of threats and violent attacks by these groups against journalists in the Guaviare department. In Haiti, the reported content glorified criminal acts by designated individuals associated with gangs. These insights supported our efforts to remove violating content and mitigate the risk of dangerous organizations and individuals on our platforms, as well as of physical attacks.

Meta partnered with Trusted Partners and regional experts to launch youth-focused messaging campaigns across Latin America. Delivered to 80,000 users aged 13–25 in Mexico and an additional 55,000 across Colombia, Haiti and Brazil, these campaigns were designed to inform users about platform violations, educate them on the risks of engaging with dangerous organizations and individuals, and empower them to make their own more informed choices online. The campaign led to a measurable increase in account deactivation and unfriending behavior.



International organizations

Meta continued to engage across the UN system throughout the year. We joined policy discussions at the 80th United Nations General Assembly (UNGA). Topics included global digital and AI governance, the use of AI for [national security](#) and bridging the AI capacity gap. We also participated in the launch of the UN Global Dialogue on AI Governance on the sidelines of UNGA High Level Week.

Meta continued to work closely with the [Office of the High Commissioner for Human Rights](#), including on the [B-Tech Project](#). We responded to requests for submissions and consulted with the [African Commission on Human and People's Rights](#) and UN Human Rights Special Procedures — independent human rights experts — including the Special Rapporteurs for Freedom of Opinion and Expression. We also engaged with other UN agencies and humanitarian organizations to help inform our crisis response work.

In addition, we participated in the 2025 [Internet Governance Forum](#), the [UN Forum on Business and Human Rights](#) and the [RightsX Summit](#), where we spoke about positive use cases of AI for human rights defenders, among other topics. We continued to participate and engage with stakeholders in several other multistakeholder forums, including the [Global Internet Forum to Counter Terrorism](#), [RightsCon](#) and [Tech Against Trafficking](#).

Transparency and remedy



The [Oversight Board](#), as an independent body, helps us resolve some of the most difficult questions around freedom of expression online: what to take down, what to leave up, and why. It reviews cases referred by Meta or appealed by individuals on Facebook, Instagram or Threads who disagree with our content moderation decisions and provides binding rulings on whether to remove or leave up content. The Oversight Board also provides recommendations to enhance our content moderation practices and offers Policy Advisory Opinions upon request.



In 2025, the Oversight Board considered cases concerning our content enforcement in light of the international human rights framework, including freedom of expression, the rights of children, and the right to equality and non-discrimination, among others.

Examples of Oversight Board decisions:



The Oversight Board [overturned](#) Meta's decisions not to label a likely manipulated audio clip of two Iraqi Kurdish politicians discussing rigging parliamentary elections. The Board considered that, to be consistent with Meta's human rights responsibilities, Meta should have placed a "High Risk" label on the content in this case as well as on all instances of the audio clip.



A majority of the Oversight Board [upheld](#) Meta's decision to leave up two posts that included videos in which a transgender woman was confronted for using a women's bathroom and a transgender athlete who won a track race.

Examples of actions following Oversight Board recommendations:



Following [recommendations](#) in the case of an [Iranian woman confronted on the street](#) and in line with our efforts to reduce over-enforcement, we updated our approach to more reactive enforcement. This helped mitigate risks of enforcing on figurative speech or banter that is potentially non-threatening.



Following [recommendations](#) to provide more details about its content decisions, Meta [increased](#) the granularity of existing notifications to users — providing more detailed notice when user content is restricted in response to a formal government report stating the content violates local laws.

Details are in [Meta's Bi-Annual Reports on the Oversight Board](#) and our [Oversight Board Impact page](#).

Impact: Increasing transparency through user notifications

+9%

Q1 2024 – Q1 2025

Increase in notification rates for eligible Community Standards enforcement on Facebook

+16%

Q1 2024 – Q1 2025

Increase in notification rates for eligible Community Standards enforcement on Instagram

~100%

in Q1 2025

Notification rate for content restricted for violating local laws on Instagram and Facebook



We also helped support the Oversight Board's [thought leadership](#) related to freedom of expression and systemic risk assessment so regulators could learn from the Board's experience. Throughout the year, the Board participated in policy discussion at the United Nations General Assembly and engaged with the African Commission on Human and Peoples' Rights, among other stakeholders.

***Disclaimer:** There are lots of differences between social media and movies. We didn't work with the MPA when updating our content settings, they're not rating any content on Instagram, and they're not endorsing or approving our content settings in any way. Rather, we drew inspiration from the MPA's public guidelines, which are already familiar to parents, as well as feedback from parents. Our content moderation systems are not the same as a movie ratings board, so the experience may not be exactly the same.



Annex



How human rights are governed and managed at Meta

Our human rights experts guide the implementation of our [Corporate Human Rights Policy](#), which is overseen by our Chief Global Affairs Officer and Chief Legal Officer.

The human rights experts' tasks include promoting the policy's integration into existing and developing policies, programs and services; undertaking due diligence; and supporting the training of employees on the policy. The policy gives guidance to build rights-respecting products, respond to emerging crises, and work with speed and agility to embed human rights at scale.

Training Meta employees on human rights

How we build at Meta is as important as what we build. Our human rights training highlights the potential and actual real-life impacts of our services, policies and business decisions on human rights. It seeks to promote a human rights mindset in our day-to-day work, encouraging respect for human rights to the benefit of all who use our services.

We launched our *Bigger than Meta: Human Rights* training in 2022, which continued throughout 2025. Our privacy training also supports our human rights training objectives by focusing on developing our collective ability to protect individuals — including, in particular, marginalized categories of individuals — against harms emanating out of the processing of people's data.



Links to referenced reports

[Responsible Business Practices Reports](#)

[Sustainability Reports](#)

[2024 Human Rights Report, 2023 Human Rights Report, 2022 Human Rights Report, 2021 Human Rights Report](#)

[2026 Anti-Slavery and Human Trafficking Statement](#)

[2025 Conflict Minerals Report](#)

[Meta Transparency Reports](#)

[Regulatory and Other Transparency Reports](#)

[Adversarial Threat Reports](#)

[Meta's Bi-Annual Reports on the Oversight Board](#)

