

JULY 2022



Meta Human Rights Report

Insights and Actions 2020 - 2021

About this Report

In March of 2021, Meta adopted its [Human Rights Policy](#) in which we commit to reporting annually on how we are addressing our human rights impacts, including relevant insights arising from human rights due diligence, and the actions we are taking in response. This is our first annual report, covering our learnings and progress from January 1, 2020 through December 31, 2021.

The scope of this report is Meta Platforms, Inc. (formerly known as Facebook, Inc.) and our assessment of what we consider to be

the company's salient human rights risks, defined by the scale, scope, irremediable character and likelihood of impact. Our salience assessments are complemented by an additional materiality assessment included in [Meta's 2021 Sustainability Report](#).

This report is inspired by Principle 15 of the UN Guiding Principles on Business and Human Rights which makes it clear that companies must “know and show” that they respect human rights.

Note on Terminology

On October 28, 2021, Facebook, Inc. changed its name to Meta Platforms, Inc. For consistency, this report uses “Meta” to refer to the company both before and after October 28, 2021. References to “Facebook” apply only to the social media platform, not the company as a whole.

Meta's Corporate Human Rights Policy applies enterprise-wide. Each Meta product and entity has its own policies and procedures with, at times, varying human rights impacts. This report references a particular policy or procedure that may not apply enterprise-wide (e.g., the Community Standards, which apply only to Facebook). Further, this report references actions taken by Meta as a company regarding a specific entity. Such a statement is not intended to imply that Meta took that action regarding all entities. 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 products or entities. If a policy is labeled a “Facebook” policy, it does not apply to other entities, e.g., Instagram or WhatsApp.

For example (and in contrast to other Meta technologies), WhatsApp is an end-to-end encrypted messaging and calling application with unique human rights touchpoints. The report considers the human rights impacts of end-to-end encryption. Likewise, this report's discussion of content moderation and related issues on Facebook and Instagram do not apply to WhatsApp. Unless a policy is specified as applying to WhatsApp, it does not apply to WhatsApp. Further, while many content moderation decisions described in this report apply to Instagram and Facebook, there are intentional distinctions in policies and procedures between the platforms. Unless a policy is specified as applying to Instagram, it does not apply to Instagram.

Executive Summary

Background

Born out of the despair of war, international human rights laws and treaties represent humanity's best hopes for itself.

Meta's mission is to give people the power to build community and bring the world closer together. We seek to advance our mission in a manner consistent with the promotion of humanity's fundamental rights: freedom of expression, privacy, non-discrimination, and more.

Meta recognizes its potential to impact human rights and has taken steps to demonstrate our commitment to human rights principles. We joined the Global Network Initiative, and made commitments in our Corporate Human Rights Policy to respect human rights as set out in the United Nations Guiding Principles on Business and Human Rights (UNGP) and other international human rights standards. This report provides an update on our progress.



Theory of Change

Our human rights policy helps ensure that Meta incorporates consideration of human rights principles into our decisions and actions and supports the evolution of the business and

human rights agenda. Our strategy leverages our mission and values. Simply put, we seek to translate human rights guidance into meaningful action, every day.



Salient Risks

Under the [UNGP](#) framework, we conducted due diligence to identify the most salient risks to human rights that could be associated with the

company's products. In this report, we identify those risks and share the steps we have taken to help manage them.



Right to Freedom of Opinion and Expression

Meta's platforms advance freedom of expression — which is interdependent with all fundamental human rights. For example, freedom of expression requires security of the body, protection from non-discrimination, and privacy. These rights are not in tension; one secures the other. We manage risks related to balancing freedom of expression and other fundamental rights on our social media platforms through our Facebook [Community Standards](#) and Instagram [Community Guidelines](#) (see, the [Community Standards Enforcement Reports](#)).

During the reporting period, we updated our policies as part of our continuing effort to balance these rights in an increasingly dynamic world: including, [specifically referencing human rights principles](#), [clarifying our health misinformation policies](#), [enhancing our bullying and harassment policy to include stronger protections against gender-based harassment](#) for everyone including public figures, enacting prohibitions against certain [mass harassment or brigading](#), and [expanding our policies that prohibit veiled and implicit threats](#), among others. Meta further provides users with a path to disagree with content policy decisions, as described in greater detail in this report under [Providing Access to Remedy](#).



Right to Privacy

The right to privacy is essential to realizing the right to freedom of expression. On [WhatsApp](#), and [expanding to our other messaging apps](#), we offer [end-to-end encryption](#), an [important privacy and security function](#) to protect people's private messages and calls, including journalists and human rights defenders.

We protect users from unlawful or overbroad government data requests via our [Meta Data Policy](#), our [WhatsApp Privacy Policy](#), our dedicated Law Enforcement Response Team, and a UNGP-informed tool assessing our responsiveness to data requests from law enforcement authorities (see, [Government Requests for User Data Transparency Reports](#)). There are many legitimate and lawful reasons why governments request

information from social media companies: international criminal investigations, anti-terrorism work, prevention of child sexual abuse and other efforts intended to protect the lives, security and rights of people in their countries. We respond to such requests for information that are consistent with internationally recognized standards on human rights, including due process, and the rule of law. If we determine that a government request is not consistent with applicable law or our policies, or unlawful (for example, overly broad, or legally deficient in any way), we do not provide governments with direct access or “back doors” to people’s information and we would challenge any order that sought to have us redesign our systems to undermine the encryption we provide to protect people’s data.



Right to Life, Liberty and Security of Person

Our Facebook [Community Standards](#) and Instagram [Community Guidelines](#) seek to protect freedom of expression and the salient risks to freedom of life, liberty, and security of the person, and include rules against [Violence and Incitement](#), [Dangerous Individuals and Organizations](#), [Coordinating Harm and Promoting Crime](#), [Bullying and Harassment](#), [Human Exploitation](#), [Hate Speech](#), [Violent and Graphic Content](#), and [Misinformation](#). We manage risks of abuse of our products for human trafficking and exploitation via in-product features to raise awareness of trafficking, deter violating behavior, and offer support to victims. We are working to increase our ability to identify the illicit actors, networks, organizations, and businesses that perpetrate these activities and disrupt them accordingly.

We manage human trafficking and exploitation risks in our supply chain through our [Code of Conduct](#), our [Responsible Supply Chain](#) program, and our membership in the [Responsible Business Alliance](#) and the [Responsible Labor Initiative](#) (see, [Anti-Slavery and Human Trafficking Statement](#)).

Moreover, as we detail in this report, [protecting Human Rights Defenders is a core pillar of our Human Rights Policy](#).



Rights to Equality and Non-Discrimination

Our Community Standards and Community Guidelines address [hate speech](#), and we have [advertising policies on non-discrimination](#). We have dedicated Civil Rights and Human Rights teams; and we have cross-company initiatives for responsible innovation and accessibility.



Best Interests of the Child

In developing youth-focused features, we are listening to [leading experts on child safety](#) (including the [National Center for Missing and Exploited Children](#)). We published a [Parents Guide](#) and other features to [increase child and teen safety](#) on Instagram, and continued our work to fight child exploitation on [WhatsApp](#) and [Facebook and Instagram](#).



Right to Public Participation, to Vote, and to be Elected

Meta is committed to protecting the right to public participation, giving everyone a voice through our products, and empowering people to vote. Over the last few years, Meta [has made significant investments](#) in teams and technologies to better protect free and fair elections, including dedicated teams focused on election integrity and products that bring relevant and reliable [voting information](#) to people. We [continue to learn](#) how to treat different types of speech that have the most meaningful impact on public opinion and how people vote at the polls, such as ads that discuss, debate, or advocate for or against important topics under our [Ads about Social Issues, Elections or Politics Policy](#). We fundamentally believe in freedom of expression, while recognizing the need to set guidelines for ads about elections and to adopt measures to reduce the spread of harmful misinformation.

Meta's Human Rights Work In Practice

I. Applying Relevant Policies;

Meta's respect for human rights — and their underlying principles of equality, safety, dignity, privacy and voice — is applied through existing policies across our company, including the [Facebook Community Standards](#), [Instagram Community Guidelines](#), [WhatsApp Terms of Service](#), the [Responsible Supply Chain program](#), Facebook's [Privacy Principles](#), our [Law Enforcement Guidelines](#), our [Data Policy](#), [Responsible AI pillars](#), and our [Diversity and Inclusion practices](#), among other policies.



II. Conducting Human Rights Due Diligence and Disclosure;

We endeavor to assess human rights impacts that may be linked directly to Meta, take appropriate action and implement findings, monitor implementation, and report annually on “relevant insights” and “actions” under UNGP 21 as defined in our [Corporate Human Rights Policy](#). From 2020 — 2021, our UNGP-informed diligence work included:

- Detailed human rights impact assessments, e.g., [the Philippines](#) and [End-to-End Encryption](#);
- Human rights due diligence on rapidly evolving situations, e.g., Israel-Palestine (to be published in Q3 2022 in response to the Oversight Board), and product initiatives, e.g., [Ray-Ban Stories](#);
- Integrated approaches, embedding UNGP prioritization criteria, and well-defined sets of human rights standards that allow us to assess and to identify human rights concerns for countries at risk and for particularly vulnerable populations.



III. Providing Remedies for Human Rights Impacts

Users can challenge some decisions regarding removal of their account or content via [the help center](#), and appeal up to the [Oversight Board](#). A [detailed human rights report](#) guided the creation of the Oversight Board. The Oversight Board [consists](#) of rights experts from a variety of cultural and professional backgrounds. The Board’s purpose is to: “[p]rotect free expression by making principled, independent decisions about important pieces of content and by issuing policy advisory opinions on [Meta’s] content policies” (See, [Oversight Board Transparency Reports](#)).



IV. Protecting Human Rights Defenders

We are building on our existing work to protect human rights defenders and their accounts. Efforts include: [combatting advanced threat actors](#) targeting them, protecting them from incorrect content removals using [Cross Check](#), offering advanced [security options](#), taking steps to thwart unauthorized access to the accounts of defenders who are arrested or detained, and partnering with human rights organizations on outreach and training. In November 2021, we launched the [Human Rights Defender Fund and Journalist Safety Initiative](#), in partnership with [Civil Rights Defenders](#) and [International Center for Journalists](#), which gives offline assistance to human rights defenders facing critical threats and supports new digital security efforts.



V. Governance, Oversight and Accountability

Meta created a global Human Rights Policy Team and tasked it with guiding the implementation of the Human Rights Policy, training stakeholders on the Policy, and applying the Policy throughout the company by integrating into existing and developing policies, programs, and products. The Team is accountable to the [Audit Risk and Oversight Committee](#) of the Board of Directors. Meta’s [Human Rights Policy](#) and its commitments to protect human rights are enterprise-wide.



The Road Ahead

Our Human Rights Policy laid a strong basis for action. In the coming years, we intend to build the maturity of our program, function, and reporting.

We continue to strive to better protect rights in countries in conflict and crisis, and for particularly disadvantaged groups. We understand that our work must address today's technologies and look towards the future.

Human and civil rights must be at the center of our development of the [metaverse](#).

Protecting rights in the digital space is one of the existential challenges for the human rights movement and for all who care about human rights. In the face of this, we approach our work with humility and determination.



Table of Contents

Introduction	13
Human Rights and Policy Timeline	14
Part 1: Meta’s Human Rights Commitments	16
Background	17
The Global Network Initiative	18
United Nations Guiding Principles on Business and Human Rights	18
Our Human Rights Vision & Strategy: The Meta Human Rights Policy and Our Theory of Change	19
From “Know and Show” to “Show, Not Tell”	20
Our Approach to Due Diligence	24
Meta’s Salient Risks	25
How Meta identifies its Salient Risks	26
Salient Risks	27
Part 2: Meta’s Human Rights Policy in Practice	28
Applying Relevant Policies	29
Content Policies	30
Actions with Regard to Salient Risks	31
Community Standards in Focus: Dangerous Organizations	34
Privacy	35
Privacy Program Overview	35
Salient Risks in Privacy	38
Protecting Users from Unlawful or Overbroad Government Data Requests	39
Properly Collecting, Using, Storing, and Deleting User Data	41
Civil Rights	42
Responsible Innovation and Responsible AI	44
Accessibility	46
Anti-Slavery and Human Trafficking	47
Election Integrity	48
United States Presidential Election 2020	48
Myanmar Election 2020	49
Preparation for Philippines Presidential Election 2022	50
COVID-19	51
Emergency Due Diligence	52
Content Policy Due Diligence	52
Privacy Due Diligence and Data for Good	53

Conducting Human Rights Due Diligence and Disclosure	54
Human Rights Impact Assessments	55
Philippines HRIA	56
Summary Disclosure - India Human Rights Impact Assessment	57
Product-Focused Human Rights Impact Assessments	61
End-to-End Encryption (“E2EE”)	61
Due Diligence Exercises	63
Ray-Ban Stories	64
Integrated Analytical Approaches	65
Countries At Risk	65
Our Program in Action: Case Studies	67
Myanmar	67
Ethiopia	68
Security Policies	69
Combating Emerging Threat Actors	70
Protecting Users Against Espionage	71
<hr/>	
Providing Access to Remedy	72
The Oversight Board	73
Code of Conduct	74
<hr/>	
Protecting Human Rights Defenders	75
<hr/>	
Governance, Oversight, Accountability	77
Governance Model	77
Stakeholder Engagement	79
Program Spotlight: Trusted Partners	81
Engagement with International Organizations	82
<hr/>	
A Final Note	83

Introduction

While technology has changed much about our world, the principles of human rights remain timeless. As the [Universal Declaration of Human Rights](#) makes clear: “All human beings are born free and equal in dignity and rights” and “everyone has the right to freedom of opinion and expression.”

When first drafted, these principles were meant to guide the behavior of governments. Today, they reach far beyond, influencing the policies and behaviors of global companies like Meta, whose decisions can make a real difference in how people’s rights are exercised, protected and respected.

Indeed, the digital sphere represents a unique challenge and opportunity for the human rights movement. Companies, the United Nations, governments, and civil society must work together to protect basic human rights online, so that as we continue to advance and innovate, we ensure human rights are respected at the frontier of new technologies.

To protect human rights online, there are rarely simple answers, only a careful balance to be struck between competing values. With billions of people using Meta’s apps and services across the world every day, it is incumbent on us to have the right policies in place, and the right processes for acting on them.

Last year, we launched our first [Corporate Human Rights Policy](#), which formalizes the company’s commitment to human rights and explains how we apply its principles to our platforms, products, policies, transparency, and programming efforts.

Under this policy, we committed to publishing an annual human rights report that would keep the public informed of our progress.

This report is inspired by Principle 15 of the UN Guiding Principles on Business and Human Rights which makes it clear that companies must “know and show” that they respect human rights.

It sums up our progress over the past two years, and explains how we think about our human rights responsibilities and what we do to uphold them. We seek to embed our commitments in a governance model which supports integration of our human rights work with ongoing activities and policies on civil rights and Environmental, Social and Corporate Governance (ESG) efforts, as part of the company’s culture, governance, decision-making processes and communication strategies.

As a next step, we’ll build on preliminary work to carry out a comprehensive human rights risk (also known as a salient risk) assessment to help prioritize our approaches and align our resources and commitments.

Most of all, we strive for our efforts to make a difference, and for our human rights commitments to be reflected in today’s technologies and in the metaverse to come.

Human Rights and Policy Timeline

2011

Adoption of the [UN Guiding Principles on Business and Human Rights](#) by the UN Human Rights Council



2016

WhatsApp completes its [rollout of end-to-end encryption](#)

2019

Creation of Meta Human Rights Policy Team

2020 May.

Announcement of [first 20 members of the Oversight Board](#)



2020 July.

Release of [Civil Rights Audit Report](#)

2020 Oct.

Oversight Board [begins operations](#)

2013

Meta [joins the Global Network Initiative](#)



2018

Public release of [Myanmar independent human rights impact assessment](#)

2020

Voluntary disclosure of data to the [UN mechanism](#) investigating events in Myanmar begins

2020 May.

Public release of executive summaries and Meta responses to human rights impact assessments for [Cambodia, Indonesia, and Sri Lanka](#)



2021 Jan.

Meta appoints first VP of Civil Rights



2021 June.

Launch of revised Meta Code of Conduct

2021 Oct.

Launch of revised content policies for public figures and human rights defenders



2021 Nov.

Launch of pilot human rights defender fund and journalist safety initiative in APAC

2021 Dec.

Launch of Philippines Human Rights Impact Assessment



2021 March.

Launch of the Meta Human Rights Policy



2021 Sep.

Meta joins the UN Global Compact



2021 Nov.

Publication of the Civil Rights Audit progress report



2022 April.

Publication of End-to-End Encryption HRIA

PART 1:

Meta's Human Rights Commitments

Background

In April of 2004 — two months after Meta’s initial founding — Shi Tao, a Chinese journalist, logged into his Yahoo! account and emailed a document about the 15th anniversary of the 1989 Tiananmen Square protests to Democracy Forum, a US-hosted website. Intense [controversy](#) erupted when it was revealed that Yahoo!, a US-based internet company, acceded to requests by the Chinese government for Shi Tao’s personal data. Tao was convicted of sharing state secrets, and sentenced to 10 years in prison.

Following this incident, [information and communications technology companies and their human rights stakeholders aligned on a framework](#) to protect people’s rights to freedom of expression and privacy against overbroad or unlawful government requests; essentially, building

a way to encourage rights-respecting practices by telecommunication and technology companies.

In 2008, this effort became the [Global Network Initiative](#) (“GNI”). Directly interpreting the relevant provisions of the [International Covenant on Civil and Political Rights](#), its members committed to implementing certain human rights protections in their decision making and processes. It was — and is — an important innovation in the field of business and human rights. Meta became a GNI member in 2013.

Building on GNI frameworks, Meta [adopted its Corporate Human Rights Policy](#), and [created a policy team](#) that works to integrate human rights principles across the company.



A. The Global Network Initiative

Global Network Initiative members must commit to its Principles on Freedom of Expression and Privacy (“GNI Principles”). These Principles are based on internationally recognized laws and standards and are intended to “provide direction and guidance to the ICT [Information and Communications Technology] industry and its stakeholders in protecting and advancing the enjoyment of these human rights globally.”

Meta joined the GNI in 2013, recognizing how “advancing human rights, including freedom of expression and the right to communicate freely, is core to our mission” and that by joining, we hoped to “shed a spotlight on government practices that threaten the economic, social and political benefits the internet provides.”

To assess whether GNI member companies are “making good-faith efforts to implement the GNI Principles with improvement over time,”

member companies are required to undergo periodic independent assessments conducted by accredited third-party assessors. GNI members are accountable for identifying, preventing, and mitigating risks of overbroad or legally deficient government requests to restrict freedom of expression or violate users’ right to privacy.

Since joining the GNI, Meta has cooperated fully with two assessments, an assessment in 2015—2016, and an assessment in 2018—2019, the latter of which found Meta is making “good-faith efforts to implement the GNI Principles with improvement over time” and commended Meta for “strengthen[ing] its systematic review of both privacy and freedom of expression.” Meta’s next GNI assessment is ongoing as of this report’s publication.

Further information regarding the GNI’s mission and principles can be found on the GNI website.

United Nations Guiding Principles on Business and Human Rights

In 2011, the United Nations Human Rights Council unanimously adopted the United Nations Guiding Principles on Business and Human Rights (“UNGP”).

Historically, governments were obliged to protect human rights. The UNGPs: changed this by articulating a role for businesses in the rights framework. For example, under UNGP II(A)(13), businesses have a responsibility to:



- a. Avoid causing or contributing to adverse human rights impacts through their own activities, and address such impacts when they occur; and



- b. Seek to prevent or mitigate adverse human rights impacts that are directly linked to their operations, products or services by their business relationships, even if they have not contributed to those impacts.

Unlike treaty law, the UNGPs are voluntary and non-binding. Companies seeking to follow the UNGPs signal that commitment in a clear and public policy statement.

B. Our Human Rights Vision & Strategy: The Meta Human Rights Policy and Our Theory of Change

Meta's mission — to give people the power to build community and bring the world closer together — is strongly aligned with human rights principles. Our mission guides how we work, what we prioritize, and which principles we use to make decisions. In this sense, our mission is inherently, closely associated with the values and rights of the Universal

Declaration of Human Rights, as well as the many treaties, practices, and norms that succeeded it. We strive for a world where businesses respect rights, governments protect rights, and people prosper.

Simply put — we seek to translate human rights guidance into meaningful action, every day.



From “Know and Show” to “Show, Not Tell”

[UNGP 21](#) states that companies should have in place policies and processes by which they can “know and show” their human rights responsibilities, including transparent and accountable communication to people or groups whose rights may be impacted by a company’s operations.

With this in mind, and in pursuit of meaningful transparency, we worked to build an ethos of human rights action that is grounded in “show, not tell.” e.g.:

- Newsroom Posts from the Global Human Rights Policy Team on human rights issues, including [Our Commitment to Human Rights](#), [Our Approach to Countries at Risk](#).
- Our [Transparency Center](#) provides a hub for Facebook’s and Instagram’s integrity and transparency work, acting as a central destination for all updates on how we [enforce Facebook’s Community Standards](#) and how we respond to decisions, recommendations, and case updates from the [Oversight Board](#).
 - We published [an end-of-year threat report](#) on six adversarial networks we found and removed for Coordinated Inauthentic Behavior, Brigading and Mass Reporting.
 - We publish a [quarterly Community Standards Enforcement Report](#) covering Facebook and Instagram.
 - We semi-annually publish data on how we enforce our [Intellectual Property](#) policies.
 - We regularly publish data on [Government Requests for User Data](#).
 - We semi-annually publish data on [Content Restrictions Based on Local Law](#).
 - We semi-annually publish data on [Internet Disruptions](#).
 - We report quarterly on [Widely Viewed Content Report: What People See on Facebook](#).
- We publish reports that apply human rights principles:
 - [Conflict Minerals Policy and Annual Report](#)
 - [Annual Diversity Report](#)
 - [Annual Sustainability Report](#)
 - [Privacy Progress Update](#)
 - [Progress on Civil Rights Audit Commitments](#)
 - [Anti-Slavery and Human Trafficking Statement](#)



Defining What Matters: Our Human Rights Policy

Building on our mission, we adopted our [Corporate Human Rights Policy](#) on March 16, 2021, committing ourselves to respecting human rights as set out in the UNGPs and other human rights

standards.¹ The Policy guides teams to build rights-respecting products, respond to emerging crises, and work with speed and agility to embed human rights at scale. Also noteworthy:

We call the policy a *policy* — not a “statement.” It is an enterprise-wide commitment; part of our governance, not just a reflection of a position.



- Our [Code of Conduct](#) requires personnel to respect the Policy.
- Meta recognizes an expansive set of relevant human rights standards and instruments — not just freedom of expression — and other rights contained in the International Covenant on Civil and Political Rights (ICCPR).² In particular, the policy is framed to recognize the importance of non-discrimination, equity, and equality — and links to the work of the Meta Civil Rights team.

Section Two (“How We Implement”) describes the Policy’s five main pillars and includes a commitment to regularly reviewing and updating the Policy.

01. Meta joins other technology companies in making human rights responsibilities more explicit. For example, Apple published its [human rights commitment](#) in August 2020, the same year that Google amended the charter of their [Audit and Compliance Committee](#) to include oversight of civil and human rights issues.

02. E.g. the [United Nations Guiding Principles on Business and Human Rights](#), which encompasses internationally recognized human rights as defined by the International Bill of Human Rights, including of the Universal Declaration of Human Rights; the International Covenant on Civil and Political Rights; and the International Covenant on Economic, Social and Cultural Rights — as well as the [International Labour Organization Declaration on Fundamental Principles and Rights at Work](#).



Our commitments: In the Human Rights Policy, we recognize:

- [The United Nations Guiding Principles on Business and Human Rights \(UNGPs\)](#)
- [The International Bill of Human Rights \(the Universal Declaration of Human Rights; the International Covenant on Civil and Political Rights; the International Covenant on Economic, Social and Cultural Rights\);](#)
- [The International Labour Organization Declaration on Fundamental Principles and Rights at Work](#);
- [The International Convention on the Elimination of All Forms of Racial Discrimination](#);
- [The Convention on the Elimination of All Forms of Discrimination Against Women](#);
- [The Convention on the Rights of the Child](#);
- [The Convention on the Rights of Persons with Disabilities](#);
- [The Charter of Fundamental Rights of the European Union](#);
- [The American Convention on Human Rights](#);
- [The Global Network Initiative \(GNI\) Principles on Freedom of Expression and Privacy, and their associated Implementation Guidelines](#);
- [The OECD Principles on Artificial Intelligence](#); and
- [The UN Declaration on Human Rights Defenders](#).

We recognize that the universal obligation of non-discrimination is a necessary — but not sufficient — condition for real, lived, equality.

We strive to respect domestic laws. When faced with conflicts between such laws and our human rights commitments, we seek to honor the principles of internationally recognized human rights to the greatest extent possible. In these circumstances we seek to promote international human rights standards by engaging with governments, and by collaborating with other stakeholders and companies.

Meta is anchored in the value “Live in the Future”

Bringing the Policy to Life: A Story of Culture and Progress

Meta is anchored in the value “[Live in the Future](#),” where we build the future of distributed work and where opportunity isn’t limited by geography; to do this, we develop products using agile software development practices.³

Our human rights strategy is built from a well-known theory of change, the Knowledge, Attitudes, and Practices Model (KAP). Researchers assume that knowledge, attitudes, and practice are related, and that knowledge and attitudes directly influence practice.⁴

We are a mission-driven company where employees are typically aligned with human rights norms. In turn, this consensus leads to a company-wide community that wants to protect and advance human rights. Our Policy and integrated analytical approaches provide guidance, expectations and guardrails for employees to apply in their work.

Human rights defenders, subject matter experts, external stakeholders, and our Human Rights Policy Team contribute to our **knowledge**. The Human Rights Policy Team shares that knowledge by advising teams and decision-makers; participating in external fora; developing role-specific and enterprise-wide training; and by performing useful and durable human rights due diligence.

Together, we create rights-respecting **practices**, decisions, approaches, and products.

To work at scale, we are guided by UNGP prioritization criteria, and look for opportunities to affect a broader system that touches multiple products and countries. By way of example, the Meta-created Oversight Board’s role in content policy development and feedback advances human rights as well as other content moderation imperatives. Likewise, so does our workstream focused on mitigating risks of harm in countries experiencing or at risk of violence.

Using this theory of change, we are building an impactful and durable human rights risk management system; one that has greater strength because it is closely aligned with our mission.

03. According to the [Agile Alliance](#), agile software development is “an umbrella term for a set of frameworks and practices based on the values and principles expressed in the [Manifesto for Agile Software Development](#) and the [12 Principles](#) behind it.” It refers to an environment where “Solutions evolve through collaboration between self-organizing cross-functional teams utilizing the appropriate practices for their context.” For an overview of agile and relevance across a broad range of industries, see Darrell Rigby, Jeff Sutherland, and Hiroataka Takeuchi “Embracing Agile: How to Master the Process that’s Transforming Management” [Harvard Business Review](#), May 2016. See also, [Nikola Jurisic](#), [Michael Lurie](#), [Philippine Risch](#), and [Olli Salo](#): [Doing vs Being: Practical Lessons on Building an Agile Culture](#). McKinsey and Company, August 4 2020
04. Originally pioneered in public health research, “K” stands for knowledge of a problem or disease, “A” for attitude towards a problem or disease, and “P” for practice or preventive behavior to protect against the problem or disease. A relevant bibliography is cited in T.H. Wan et al., [A Systematic Review of KAP-O Framework for Diabetes](#), [ESMED Medical Research Archives Vol 3 No 9 \(2016\): April, Vol.3 Issue 9](#)

C. Our Approach to Due Diligence

[UNGP 17](#) states that companies should carry out human rights due diligence in order to “identify, prevent, mitigate and account for how they address their human rights impacts.” [UNGP 24](#) guides companies to prioritize their actions to address actual and potential human rights impacts according to severity.

We based our early approaches to human rights due diligence on our obligations as members of the GNI; specifically, we put into place strong operational, policy, and legal measures designed to scrutinize governmental takedown requests and requests for personal data of users.

We expanded our approach with the November 2018 release of an independent human rights impact assessment (“HRIA”) of Facebook’s impact in [Myanmar](#).⁵

Methodologies for human rights due diligence are evolving. HRIAs are only one form of diverse human rights due diligence offerings. Historically, many HRIAs stemmed from the extractive or natural resource industries and focused on a specific defined physical location, or on supply chain issues. Today, initiatives such as the [UN B-Tech](#) project seek to build on these methodologies to better fit risks faced by technology companies.

Global frameworks agree that the defining element of a good faith HRIA is that it involves direct consultation with affected rights-holders, as made clear in [UNGP 21](#) which requires companies to have in place policies and processes by which they can “know and show” their human rights responsibilities, including transparent and accountable communication to people or groups whose rights may be impacted by a company’s operations.⁶ Since 2018, Meta has rapidly experimented with ways to “show, not tell” our commitment to human rights risk management through our due diligence program. Building on country-based HRIAs from 2018 and 2019, Meta tested a variety of due diligence approaches

during this reporting period. Our goal has been to turn insights into actionable approaches that work for a company of Meta’s size, scale, and the culture of agile software development.

In addition, we engaged in crisis support; continuous due diligence processes, where we focus on a key challenge for an ongoing period; as well as ongoing content policy feedback, and real time decision-making support.

We also sought to be an industry leader in assessment and transparency. On an ongoing basis, we shared insights and actions from our human rights due diligence. Just prior to the reporting period, Meta disclosed a [detailed human rights report guiding the creation of the Oversight Board](#). During the reporting period, Meta disclosed due diligence on [Cambodia](#), [Indonesia](#), the [Philippines](#), [Sri Lanka](#), and — just after the reporting period — on its transition of Messenger and Instagram Direct Messages to [End-to-End Encryption](#).

05. For more on HRIAs, see Part 02.II of this report. Meta’s reference in this Report to third party diligence assessments cannot be construed as admission, agreement with, or acceptance of any of the findings, conclusions, opinions or viewpoints identified in those assessments, or the methodology that was employed to reach such findings, conclusions, opinions or viewpoints. Likewise, while Meta references steps it has taken, or plans to take, which may correlate to points assessors raised or recommendations they made, these also cannot be deemed an admission, agreement with, or acceptance of any findings, conclusions, opinions or viewpoints.

06. See also, UNGP 18 (“In order to gauge human rights risks, business enterprises should identify and assess any actual or potential adverse human rights impacts with which they may be involved either through their own activities or as a result of their business relationships. This process should: (a) Draw on internal and/or independent external human rights expertise; (b) Involve meaningful consultation with potentially affected groups and other relevant stakeholders, as appropriate to the size of the business enterprise and the nature and context of the operation.”); OHCHR, [The Corporate Responsibility to Respect Human Rights: An Interpretive Guide](#) (2012), at 33 (“Human rights due diligence is about people. It reflects the entitlement of every human being to be treated with dignity. It therefore involves relationships — between an enterprise and those on whom it may have an impact. Hence, the key to human rights due diligence is the need to understand the perspective of potentially affected individuals and groups. Where possible and appropriate to the enterprise’s size or human rights risk profile, this should involve direct consultation with those who may be affected or their legitimate representatives, as discussed further under Guiding Principle 18.”).

Meta's Salient Risks

Every day, billions of people all over the world use Meta's social media platforms and messaging apps.

In this report, we aim to show the most salient human rights issues identified by our due diligence efforts to date, the actions taken across product and policy teams to prevent or mitigate these risks, and how we developed processes to assess the effectiveness of those measures.



“Human rights risks are understood to be the business enterprise’s potential adverse human rights impacts.

Potential impacts should be addressed through prevention or mitigation, while actual impacts — those that have already occurred — should be a subject for remediation.”

COMMENTARY TO UNGP 17

A. How Meta Identifies its Salient Risks

As the Key Concepts section of the [Interpretive Guide to the UNGPs](#) clarifies:

The most salient human rights for a business enterprise are those that stand out as being most at risk. This will typically vary according to its sector and operating context. The Guiding Principles make clear that an enterprise should not focus exclusively on the most salient human rights issues and ignore others that might arise. But the most salient rights will logically be the ones on which it concentrates its primary efforts.

Meta uses a thorough human rights assessment process to identify salient risks in a given or particular context. Drawing on these assessments — as well as our ongoing day-to-day human rights due diligence — we established an initial list of salient human rights issues.

We commit to conducting a company-wide salient risk analysis in the near future and reporting on it in a forthcoming annual report.



B. Salient Risks

Based on due diligence efforts to date, the categories of salient risks for Meta can include, for example:

Rights	Human Rights Standards
Right to Freedom of Opinion and Expression	Universal Declaration Of Human Rights, Art. 19 International Covenant On Civil And Political Rights, Arts. 19, 20 Convention On The Rights Of The Child, Arts. 13, 17 International Convention On The Elimination Of All Forms Of Racial Discrimination, Art. 4
Right to Privacy	Universal Declaration of Human Rights, Art. 12 International Covenant on Civil and Political Rights, Art. 17
Right to Life, Liberty and Security of Person	Universal Declaration of Human Rights, Art. 3 International Covenant on Civil and Political Rights, Arts. 6, 9 Convention on the Rights of the Child, Art. 6
Rights to Equality and Non-Discrimination	Universal Declaration of Human Rights, Arts. 1, 2, 7 International Covenant on Civil and Political Rights, Arts. 2, 3, 26 International Covenant on Economic, Social and Cultural Rights, Arts. 2, 3 Convention on the Elimination of All Forms of Discrimination Against Women, Art. 2 International Convention on the Elimination of All Forms of Racial Discrimination, Art. 2
Best Interests of the Child	International Covenant on Civil and Political Rights, Art. 24 Convention on the Rights of the Child, Art. 3
Rights to Public Participation, to Vote, Be Elected	Universal Declaration of Human Rights, Art. 21 International Covenant on Civil and Political Rights, Arts. 25 Convention on the Elimination of All Forms of Discrimination Against Women, Art. 7, 8 International Convention on the Elimination of All Forms of Racial Discrimination, Art. 5

PART 2:

Meta's Human Rights Policy in Practice

Our approach to human rights is grounded in the five pillars of our [Human Rights Policy](#):

- I. Applying Relevant Policies;
- II. Conducting Human Rights Due Diligence and Disclosure;
- III. Providing Remedies for Human Rights Impacts;
- IV. Protecting Human Rights Defenders; and
- V. Governance, Oversight and Accountability

As we reflect on the Policy's first year, we address our progress and our room for growth across each area of commitment.

Applying Relevant Policies

Under our [Human Rights Policy](#), our respect for human rights — and their underlying principles of equality, safety, dignity, privacy and voice — is applied through our existing policies, including the Facebook [Community Standards](#), Instagram [Community Guidelines](#), [WhatsApp Terms of Service](#), the [Responsible Supply Chain](#) program, our [Privacy Principles](#), our [Law Enforcement Guidelines](#), our [Data Policy](#), our [WhatsApp Privacy Policy](#), our [Code of Conduct](#), and our [Diversity and Inclusion](#) practices, among other policies so included.⁷

In the following section, we highlight key human-rights related developments in the application of our policies.



7. Moreover, as discussed *infra*, we extend our human rights commitments to additional, internal product-focused decision-making approaches. These efforts require coordination, dependencies and alignment with Products, Operations, and Corporate partners.

A. Content Policies

Freedom of expression is a fundamental right, one that is interdependent with other fundamental rights. We look to international human rights experts when developing our standards for what content is and is not allowed on our social media platforms, and when deciding how to implement these standards in practice.

These rules are known as the [Community Standards](#) for Facebook and [Community Guidelines](#) for Instagram. They were — and

are — developed based on feedback from our community and the advice of experts in fields such as technology, public safety [and human rights](#). In seeking to ensure that everyone's voice is valued equally, we take care to create standards that include different views and beliefs, especially from people and communities that might otherwise be overlooked or marginalized. For these reasons, when we limit expression, we do it in service of one or more of a set of values: Authenticity, Safety, Privacy, and Dignity.

Our Commitment to Voice:

The goal of our Community Standards is to create a place for expression and give people a voice. Meta wants people to be able to talk openly about the issues that matter to them, even if some may disagree or find them objectionable.

THE FACEBOOK COMMUNITY STANDARDS

The full set of Community Standards are [available](#) in 66 language options to-date, as are our [detection and enforcement policies](#). Details on the effectiveness of our enforcement of these policies can be found in our quarterly [Community Standards Enforcement Reports](#).

Our work to infuse our Community Standards with human rights principles is ongoing and responsive to changing global salient risks, consistent with Meta's iterative approach to policy development.

01. Actions with Regard to Salient Risks



Right to Freedom of Opinion and Expression

[Universal Declaration of Human Rights](#), Art. 19

[International Covenant on Civil and Political Rights](#), Arts. 19, 20

[Convention on the Rights of the Child](#), Arts. 13, 17

[International Convention on the Elimination of All Forms of Racial Discrimination](#), Art. 4

Our Community Standards and Community Guidelines are highly informed by international human rights standards (including the [Rabat principles](#)). Writing these policies involves extensive stakeholder engagement, including with human rights activists and academics. The implementation of these policies is informed by trusted partners who provide important context, especially in high-risk environments.

For example, Meta drew on international human rights law in determining how to address the complicated issue of [bullying and harassment of public figures](#). For this issue, the right to freedom of expression and legitimate public discourse around people in the public eye must be carefully balanced against public figures' rights to security and non-discrimination, honor, and reputation.

To develop a consistent approach, Meta consulted with a [diverse set of global stakeholders including](#) free speech advocates, human rights experts, women's safety groups, our Women's Safety Expert Advisors, cartoonists and satirists, female politicians and journalists, representatives of the LGBTQ+ community, content creators and public figures.

Based on stakeholder feedback, and guided by international human rights principles, the resultant [Community Standards Bullying and Harassment policy](#) maintains a distinction between public figures and non-public figures, and requires removal of [attacks on public figures](#) that encompass a wide range of harms to their fundamental rights. During the reporting period, this policy was further updated to remove certain severe sexualizing, degrading or derogatory content. The policy also provides additional protections to public figures like journalists and human rights defenders who have become famous involuntarily or because of their work.

In another, very important example: when a government sends Meta a take-down request, we typically conduct three levels of [review](#): against the Community Standards, against applicable local law, and against international human rights standards and our company [principles](#). No matter which government makes a given request, we will scrutinize it to make sure it is legally valid, proportionate and legitimate. In the interests of full transparency, we publish reports that set out statistics on [content we restricted based on local law](#), including in response to government requests. We also publish case studies of content we have restricted based on government takedown requests and [recently announced](#) we may also share these requests with [Lumen](#).



Right to Privacy

[Universal Declaration of Human Rights, Art. 12](#)

[International Covenant on Civil and Political Rights, Art. 17](#)

Privacy is one of the core values of the [Community Standards](#). As we note: it “gives people the freedom to be themselves, choose how and when to share on Facebook and connect more easily.” In addition to prohibiting personally identifiable information such as social security numbers, or other personal information directly identifying an individual, the Community Standards also prohibit the posting of personal contact information about oneself or others, including phone numbers or financial information about businesses or organizations, except when publicly available. In 2021, Meta requested a policy advisory opinion from the Oversight Board on the sharing of private residential addresses and images. In 2022, the Board recommended that we [disallow](#) this kind of sharing, and Meta [acceded](#).



Right to Life, Liberty and Security of Person

[Universal Declaration of Human Rights, Art. 3](#)

[International Covenant on Civil and Political Rights, Arts. 6, 9](#)

[Convention on the Rights of the Child, Art. 6](#)

We [expanded our policies that prohibit veiled and implicit threats](#). We added a policy to address hate speech attacks on concepts, ideas, practices, beliefs and institutions related to protected characteristics when those attacks pose an imminent risk of harm, intimidation or discrimination. We enhanced our bullying and harassment policy to include [stronger protections against gender-based harassment](#) for everyone, including public figures. We created new policies against certain [mass harassment or “brigading.”](#) We also revisited our [violence and incitement policies](#) — specifically, the allowance around the discussion of State use of force.



Rights to Equality and Non-Discrimination

[Universal Declaration of Human Rights, Arts. 1, 2, 7](#)

[International Covenant on Civil and Political Rights, Arts. 2,3, 26](#)

[International Covenant on Economic, Social and Cultural Rights, Arts. 2, 3](#)

[Convention on the Elimination of All Forms of Discrimination Against Women, Art. 2](#)

[International Convention on the Elimination of All Forms of Racial Discrimination, Art. 2](#)

We make country-specific decisions informed by [extensive stakeholder engagement](#) through our work with [Trusted Partners](#) and human rights activists and academics. Using existing policies and relationships, we improved our abilities to enforce at scale.

In 2021, we provided [an update on our work in Ethiopia](#) identifying and removing a number of persistent harmful false claims and out of context imagery that make false allegations about the perpetrators, severity or targets of violence in Ethiopia. This decision was based on guidance from over 50 local partners and independent experts.



Best Interests of the Child

[International Covenant on Civil and Political Rights, Art. 24](#)
[Convention on the Rights of the Child, Art. 3](#)

The [Philippines HRIA](#) spurred improvements to our existing policy language prohibiting multiple forms of human exploitation, and enforcement of our policy that [prohibits](#) all sales of people, including children. As we described in [Meta's Response](#), we updated policies around our proactive detection tools with the goal to reduce discoverability of exploitative content and incorporate emerging trends in consultation with the Philippines Department of Health and the National Kidney and Transplant Institute, as well as the Inter-Agency Council Against Trafficking, Overseas Workers Welfare Administration, Blas Ople Policy Center, Department of Foreign Affairs and the Philippine Overseas Employment Administration. We have ongoing coordination with the Philippines Inter-Agency Council Against Trafficking, Department of Justice, Department of ICT and the Department of Education as well as organizations advocating child safety, to quickly respond to reports and gather local insights and trends, and help us improve our policies, tools, and resources to better protect children online.



Rights to Public Participation, to Vote, Be Elected

[Universal Declaration of Human Rights, Art. 21](#)
[International Covenant on Civil and Political Rights, Arts. 25](#)
[Convention on the Elimination of All Forms of Discrimination Against Women, Art. 7, 8](#)
[International Convention on the Elimination of All Forms of Racial Discrimination, Art. 5](#)

Our [human rights-informed approach to election misinformation](#) balances freedom of expression against threats to the right to public participation. For example, [in preparation for the 2020 election in the United States](#), we took steps aimed at preventing the spread of misinformation, (including [clearer fact-checking labels](#)), fighting voter suppression and interference (including [banning paid ads that suggest voting is useless or advise people not to vote](#)), and helping people better understand the information they see online (including an [initial investment of \\$2 million](#) to support a media literacy project). We also [partnered with third-party fact-checkers on WhatsApp](#).

02. Community Standards in Focus: Dangerous Organizations

One area that has drawn significant attention from human rights stakeholders has been our [Dangerous Individuals and Organizations Policy](#) (“DOI” Policy), which addresses networks of people who “proclaim a violent mission or are engaged in violence.” We do not want Meta to be a platform for hate.

While terrorism and hate are global issues, there is currently no globally recognized and accepted definition of terrorist or hate organizations. So we developed [the definitions](#) in our [DOI Policy](#) to assist our decision-making on enforcing against such organizations:

- Our designations are divided into three tiers that indicate the level of content enforcement, with Tier 1 resulting in the most extensive enforcement, because we believe these entities have the most direct ties to offline harm.
- We designate Dangerous Individuals and Organizations under our policy after a rigorous process that takes into account both online and offline behavior. During this process, we work to identify an organization’s goals and if it has a track record of offline violence.
- As for **terrorist groups**, we designate individuals and organizations based on their behavior and whether they engage in, advocate, or lend substantial support to purposive and planned acts of violence with the intent to coerce, intimidate and/or influence a civilian population, government, or international organization in order to achieve a political, religious, or ideological aim.
- In addition to the definition of terrorist groups, and in recognition of the political complexity associated with the issue, we define as **violent non-state actors** those groups who primarily direct violence toward state or military actors (e.g. military targets).
- As for **hate groups**: We ban groups that proclaim a hateful and violent mission from having a presence on Facebook and Instagram and we remove content that represents, praises or supports them. To date, we identified a range of groups across the globe

as hate organizations because they engage in coordinated violence against others based on characteristics such as religion, race, ethnicity or national origin and we routinely evaluate groups and individuals to determine if they violate our policy.

- In addition, our policy covers Militarized Social Movements, Violence-Inducing Conspiracy Networks, and individuals and groups banned for promoting hatred.

Our process for assessing organizations for possible designation includes structured review of each case by a wide range of relevant teams, including policy, legal, and security. As the [DOI Policy](#) demonstrates, our definitions and thresholds are agnostic to region or ideology. Every designation goes through the process.

14

Strategic network
disruptions to remove

23

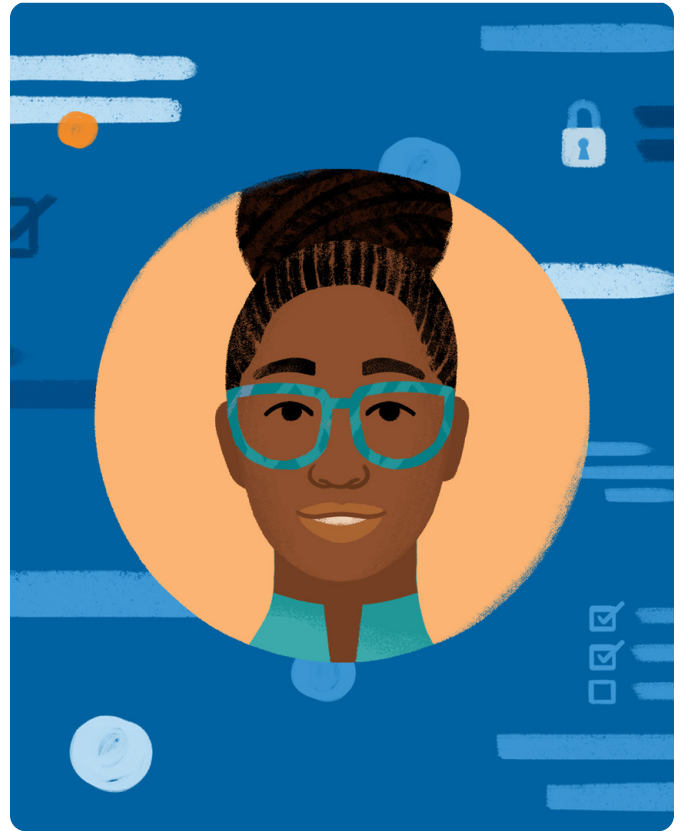
Different organizations
designated under our DOI policy

Finally, as part of our enforcement efforts against DOIs, we undertake certain strategic network disruptions, to help make sure that these groups cannot find ways back on our platforms and proliferate on our platforms. As outlined in our [Community Standards Reporting](#), between October 2019 and August 2020, we conducted 14 strategic network disruptions to remove 23 different organizations designated under our DOI Policy, including [three U.S.-based white supremacist groups](#) in October 2019.

B. Privacy

The right to privacy is a foundational right; an essential requirement for the realization of other rights, like the right to freedom of opinion and expression. Particularly in this digital age, interference with this right can both directly and indirectly limit the free development and exchange of ideas.

Privacy is one of the defining social issues of our time and is central to Meta's vision. We are [building a new privacy foundation to guide us now and in the future](#). We are committed to seeking feedback from and working with stakeholders across industry, civil society, think tanks and academia to improve our program.



Privacy Program Overview

Our [Privacy program](#) rests on a foundation of:



Governance: A cross-functional group of organizations across the company provides engineering, legal, policy, compliance and product expertise that enable the design and implementation of our privacy program. This includes an independent committee of Meta's Board of Directors that meets quarterly to oversee our efforts to live up to our privacy commitments.



Privacy Education. To make privacy a core responsibility of every Meta employee, we drive continuous privacy learning and education that spans training, internal campaigns, regularly provided privacy content and other dynamic resources. We do this through internal Workplace channels, lightning talks with privacy leadership, internal Q&A sessions, a dedicated Privacy Week and an internal hub of on-demand privacy content to help guide decisions and processes.



Accountability in Practice. In order to put our accountability foundation into practice, we designed processes, escalation paths and technical mechanisms that embed privacy across all facets of our company operations. This includes:

- **A privacy risk assessment program** that performs an annual assessment to identify, assess, and address privacy risk across the company, as well as a process to assess privacy risk after an incident occurs.
- **Safeguards and controls**, including operational activities, policies, and technical systems.
- A **privacy review process** by which we assess privacy risks that involve the collection, use, or sharing of people’s information and external representations about our privacy and security practice; along with technical requirements and tools to enhance accountability and operate the privacy review process at scale, such as: a centralized tool that is used throughout the project lifecycle for Privacy Review and a technical implementation review that conducts review, verification, and documentation of the technical implemental of privacy mitigations and commitments prior to product launch.
- **Incident Management**, which includes: **proactive identification** of potential privacy issues, a **bug bounty** program that has, over the past 10 years, awarded bounties to around 1,500 researchers from 107 countries, and **transparency** efforts, like our Incident Management program, that includes steps to notify people where appropriate about issues impacting our community, or working with law enforcement or other officials to address what issues we find.
- **Third Party Oversight**, including a third-party privacy assessment process for service providers to assess and mitigate privacy risk at Meta, as well as a formal process for enforcing and offboarding third parties who violate our privacy or security obligations, such as procedures and infrastructure designed to ensure that third-party developers complete the Data Use Checkup, and other technical and procedural mechanisms to monitor compliance.
- **External Data Misuse.** We also have technical mechanisms in place to mitigate and prevent third parties from accessing data from Meta, through proactive and reactive measures like prevention, deterrence, detection and enforcement.



Privacy Product Outcomes. The accountability processes, safeguards, and technical mechanisms we built help ensure that new products and features embed privacy by design. We see these updated processes enable us to improve our privacy approach in new products and features, as we pivot to respond to the world around us. This includes:

- Giving people greater access and control over their information
- Giving people access to their information
- Giving people control to manage their activity
- Offering an ephemeral messaging option
- Creating a secure, private messaging experience
- Providing age-appropriate experiences for youth
- Communicating data privacy practices to people
- Using privacy enhancing technologies for social good



Investments In A Technical Foundation That Supports Privacy, *i.e.*, sustainable technical solutions to meet evolving privacy expectations and ensure consistent application of our privacy requirements across our products and systems, such as:

- A data deletion framework that helps alleviate the risk of potential error through machine-learning automation; and
- Privacy-enhancing technologies based on advanced cryptographic and statistical techniques that minimize the data we collect, process and share — to help protect data at different stages of the data lifecycle.

Our privacy work is never finished, and we understand that our commitment means continuously improving and focusing on this every day.





01. Salient Risks in Privacy

Two highlights of how Meta approaches its salient privacy risks include: protecting users from unlawful or overbroad government demand for user data; and our responsibility to protect user data and use it, share it, store it, and delete it pursuant to privacy and data protection principles and applicable legal requirements.

Meta's approach to both challenges builds on our theory of change: developing systemic, scalable, integrated analytical approaches that provide guardrails informing product development enterprise-wide.

The salient risk of government overreach and abuse in demanding user data from private technology companies was the starting point of Meta’s human rights journey.

I. Protecting Users from Unlawful or Overbroad Government Data Requests

The salient risk of government overreach and abuse in demanding user data from private technology companies was the starting point of Meta’s human rights journey. The salience of this risk has not diminished over time. In compliance with the GNI Principles, and our commitment to the UNGPs, we conducted human rights due diligence to develop an approach for our responses to government demands for user data.

As set out in Facebook’s Data Policy, and guidelines for Law Enforcement, we scrutinize every government request we receive, no matter which government makes the request. There are many legitimate and lawful reasons why governments request information from social media companies: international criminal investigations, anti-terrorism work, prevention of child sexual abuse and other efforts intended to protect the lives, security and rights of people in their countries. We respond to such requests for information that are consistent with internationally recognized standards on human rights, including due process, and the rule of law. When we do

comply, we produce only the information that is narrowly tailored to that request. If we determine that a government request is not consistent with applicable law or our policies, or unlawful (for example, overly broad, or legally deficient in any way), we will push back.

We do not provide governments with “back doors” to people’s information and we would challenge any order that sought to have us redesign our systems to undermine the encryption we provide to protect people’s data, or any attempt to gag us from disclosing the existence of such an order and our efforts to fight it. For example, Facebook and WhatsApp paused — on human rights grounds — all data disclosure to the Hong Kong government after the passage of the National Security Law in 2020.

We “show, not tell,” by regularly publishing transparency reports that set out, by country, the number of government requests we have received and the proportion of requests for which some data was produced.

Reform Government Surveillance Coalition

Along with other major tech companies, Meta is a member of the [Reform Government Surveillance Coalition](#), which urges the world's governments to adopt surveillance laws and practices consistent with established norms of privacy and free expression and the rule of law.

The [six core principles](#) of the coalition are:



1. Limiting governments' authority to collect users' information;



2. Oversight and accountability;



3. Transparency about government demands;



4. Respecting the free flow of information;



5. Avoiding conflicts among governments; and



6. Ensuring security and privacy through strong encryption.





II. Properly Collecting, Using, Storing, and Deleting User Data

We [made progress on our work to give people more control over their data](#), and our broader mission to honor people's privacy in everything we do by building processes, products, and technical mechanisms that laid the foundation for privacy and accountability across the company. The program includes internal education, due diligence, safeguards and controls, and third-party oversight. For example, we provide people with ways to:

- [Manage their information](#). We make it easy for people to access, manage, download or delete the personal data they have provided.
- [Transfer their data](#). We make it simple to move information, such as posts, photos and videos, to another service.
- [Access the Accounts Center](#). People can control the experiences they connect across our apps by using our Accounts Center.
- [Access the Instagram Activity Center](#). People can access and manage content they've commented on, shared or sent on Instagram.
- [Perform a Privacy Checkup](#). This tool helps people control who can see what they share, how their information is used and how to secure their account.
- [See "Why Am I Seeing This Ad?"](#) This tool lets people tap on posts from the friends, Pages, and Groups they follow as well as some of the posts Meta suggests to get more context on why they are appearing in News Feed.
- [Review Off-Facebook Activity](#). We created a tool so people can control — or disconnect — the information businesses send to Meta about their activity on other apps and websites.
- [Review Ads Interests & Preferences](#). This tool gives people more visibility into the actions they've taken to control the topics of the ads you see.
- [Access Cookie Controls](#). In the European Region, we give people a more granular level of control over their cookie choices and more information on how we use different kinds of cookies.
- [Set up Two-Factor Authentication](#). We give people this option in addition to their password to help protect their account from improper access.

C. Civil Rights

Civil rights are human rights, but they are also understood to be personal rights protected by the Constitution of the United States and Federal and State laws. The Meta Civil Rights team focuses on issues of discrimination and exclusion based upon race, ethnicity, religion, national origin, gender, gender identity, sexual orientation and disability.

Meta's commitment to human rights and civil rights are closely interrelated. Our Human Rights Policy lays a strong basis for continuing civil rights and racial justice [work](#) and [standards](#).

The initial framework for our Civil Rights work was the [civil rights audit](#) we commissioned at the behest of the civil rights community to identify instances where the company could take a more comprehensive approach to respecting civil rights. Meta published the findings in full: a total of three reports, the [last of which](#) was published in July 2020. The extensive audit resulted in 117 recommendations and actions with input from over 100 civil rights and social justice organizations.

This audit laid the groundwork for the Civil Rights Team that was created in November 2020 and provided the company with an initial roadmap.

In November of 2021, Meta published a report that detailed [Meta's Progress on Civil Rights Audit Commitments](#), demonstrating success in implementing 65 of the Auditors' recommendations and actions, as well as demonstrating ongoing or in-progress status of 42, and continuing evaluation of the feasibility of eight. This Progress Report explained that Meta's civil rights work focuses on five key pillars:



Law Enforcement & Hate: Addressing harm and creating accountability are critical to enhancing protections for marginalized communities and victims of hate speech, hate incidents and hate crimes are safe on our platforms. It also means working to prevent law enforcement misconduct and surveillance of marginalized communities on our platforms as these activities can violate our policies and chill free expression rights.



Voting & Civic Engagement: Grounding the work in efforts to inform and connect people, communities, and civic institutions to come together and drive real outcomes to improve lives, because when all people have an opportunity to participate and are considered political equals, societies make decisions optimized for the many, not the few.



Technology: It is Meta’s responsibility to help instill civil rights best practices throughout the company with regard to artificial intelligence, machine learning, virtual and augmented reality, and emerging technologies, and lead engagement with external technology experts.



Product: Meta aims to advance our priority of doing good and causing no harm by creating product review approaches and guidance tools aimed at prioritizing protected classes and protecting systemically and historically marginalized communities.



Policy & Enforcement: Meta works to surface civil rights concerns on the front end of policy development and enforcement decisions and ensure a more diverse and inclusive engagement process.

Growing out of the civil rights audit, the Meta Civil Rights Team has grown to include people with expertise spanning hate crimes, voting rights, counterterrorism, immigration, national security, law enforcement, product inclusion, algorithmic fairness and transparency.

As this team continues to advance civil rights and liberties, it does so in partnership with the Human Rights Policy Team. Meta is committed to continuing to build with the values of justice, equity, dignity, and safety in mind.



D. Responsible Innovation and Responsible AI

To help ensure that we build the future of the internet responsibly, Meta developed human rights-based integrated analytical approaches to support our innovation and product development.

Our [Responsible Innovation Dimensions](#) support product development. This framework is evolving over time, but currently includes 10 dimensions: autonomy, civic engagement, constructive discourse, economic security, environmental sustainability, fairness and inclusion, privacy and data protection, safety, voice, and well-being. These dimensions in turn guide our analysis and practice.

Artificial Intelligence (AI) is a core component of the technologies Meta uses to provide value to people and [keep our communities safe](#), from helping [rank posts in News Feed](#) to [tackling hate speech](#) and [misinformation related to the](#)

[COVID-19 pandemic](#). Meta also developed a [dedicated, cross-disciplinary Responsible AI \(RAI\) team](#) within its AI organization to help ensure that AI governance is based on foundational values of respect for human rights, democracy, and the rule of law.

Those foundational values are at the root of the wide range of principles statements that have been released around responsible AI development, most especially the European Commission's [High-Level Expert Group's Ethics Guidelines for Trustworthy AI](#) and the [Organization for Economic Cooperation and Development's Principles on Artificial Intelligence](#), which Meta helped develop.

Meta, in turn, organized its Responsible AI efforts around [five key pillars](#):



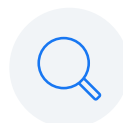
Privacy & Security



Fairness & Inclusion



Robustness & Safety



Transparency & Control



Accountability & Governance



We also invest in research and open-sourcing datasets and tools to help facilitate responsible use of AI, such as [privacy-preserving machine learning](#), [AI explainability](#), and [fairness](#). For example, in 2021 we released our [Casual Conversations data set](#), composed of over 45,000 videos designed to similarly help researchers evaluate computer vision and audio models for accuracy across a diverse set of ages, genders, apparent skin tones, and ambient lighting conditions.

We are also improving transparency by piloting simple, standardized documentation of our models and [using interpretability software such as Captum](#). Although work in this area is still in its infancy, [our hope](#) is that ultimately we will be able to build an integrated transparency solution that can automatically feed information from internal documentation efforts — like model cards⁸ — into new transparency features and controls for the people using our products.

In addition to our technical research and product-focused work, we are actively participating in efforts to establish clear AI principles and best practices, including [collaborating with the OECD on the AI Observatory project](#) to study and

disseminate emerging best practices that are in line with its [AI Principles](#).

Through our [partnership with Open Loop](#), we are building innovative “policy prototyping” projects for testing new potential AI policy requirements with regulators and startups before they become law, to ensure that they are both practical and impactful. [We launched](#) projects in Europe and the Asia Pacific region.

Finally, we are [funding a global effort](#) to solicit diverse academic research on AI ethics and governance topics, supporting the publication of academic papers in [Asia](#), [Africa](#), and [Latin America](#) and providing foundational support for an [independent Institute for Ethics in Artificial Intelligence at the Technical University of Munich](#).

08. Model Cards are short documents accompanying trained machine learning models that provide benchmarked evaluation in a variety of conditions, such as across different cultural, demographic, or phenotypic groups (e.g., race, geographic location, sex, Fitzpatrick skin type) and intersectional groups (e.g., age and race, or sex and Fitzpatrick skin type) that are relevant to the intended application domains. Model cards also disclose the context in which models are intended to be used, details of the performance evaluation procedures, and other relevant information.

E. Accessibility



The [UN Convention on the Rights of People with Disabilities](#) sets out eight fundamental principles that underpin the protection of the rights of people with disabilities. These include non-discrimination, full and effective participation and inclusion in society and accessibility. Given the importance of the digital world for people with disabilities to be able to participate fully in exercising their rights, Meta sought to apply cutting-edge technologies and make other investments to create more accessible features and services for users with a wide range of disabilities. Such measures include:

- [Applying the latest AI techniques to increase object and concept recognition in Automatic Alt Text](#) or “AAT,” our award-winning automatic photo-description technology that works with screen readers to describe photos to people who are blind or have low vision, by more than 10X.
- Re-engineering facebook.com to [integrate greater accessibility into the website](#), including technology that helps people using screen readers better understand what’s on a page and find what interests them more quickly.
- [WhatsApp has various accessibility features](#), including a voice-assisted camera feature, which makes it easier for people who are blind or have low vision to take photos on iOS devices, and a new WhatsApp email accessibility support channel for users to share suggestions, ask questions, or report problems.

- We introduced new accessibility-specific features in Quest 2, including color filters to assist people with color-sensitivity and “adjust height” to enable better VR experiences for people who are reclining, seated, or lying down.
- Automatically generated captioning available for [Facebook Live](#), [Workplace Live](#), [Instagram TV](#), and [Live Audio Rooms](#), as well as for Facebook ads, Pages, and Groups.

In addition, we engaged in several initiatives to help promote accessibility beyond Meta. We work collaboratively with leaders in the accessibility community globally, and are working within our industry and through the [W3C](#) to establish an interoperability standard for web and assistive technology.

We also co-founded and remain active in [Teach Access](#), a collaboration of industry, academia, and disability advocates to advance the teaching of accessibility in education. Students learn state-of-the-art best practices in accessible software design, preparing them to make future technologies that are “born accessible.” We are investing early in making the metaverse accessible, including by creating XR Accessibility guidelines for developers in partnership with the XR Association and additional guidelines specifically for third-party developers of the Quest VR platform.

F. Anti-Slavery and Human Trafficking

Meta is opposed to all forms of human trafficking, slavery, servitude, forced or compulsory labor, and all other trafficking-related activities. [We are committed](#) to respecting all applicable international human rights standards, labor and employment laws, rules, and regulations, and to working to mitigate the risks of modern slavery and human trafficking in our business operations and supply chains. We manage labor, human rights, and environmental risks in our supply chain through our [Responsible Supply Chain](#) (RSC) program. We collaborate and share ideas with the broader industry and global community to update and develop solutions that promote best practices for a responsible supply chain, and to that end we are members of the [Responsible Business Alliance](#) and the [Responsible Labor Initiative](#).

We require all personnel be trained on and comply with our [Code of Conduct](#), which affirms our commitments to human rights and prohibits violations of law, including labor and employment laws. Meta requires that priority suppliers, which are identified using a risk-based methodology, conform with all of the standards on labor, health and safety, the environment, business ethics, and the establishment of management systems

in accordance with the [Responsible Business Alliance \(RBA\) Code of Conduct](#), which includes standards related to indicators of modern slavery and human trafficking.

We [verify supplier conformance](#) with RSC policies and standards through continuous dialogue, independent audits and assessments, corrective action plans, worker surveys, and other forms of assurance. Any non-conformances identified are actively addressed through our corrective action and key performance indicator programs. We regularly assess and track the effectiveness of our actions taken in mitigating the associated risks of modern slavery and human trafficking. We strive to improve year-on-year; every year we publish a refreshed/new [statement](#) to demonstrate progress and strengthening of our program efforts.

As noted, Meta [strongly opposes the abuse of our family of apps and platforms](#) to facilitate any form of human exploitation. At the core of our strategy, we continually update and refine our policies to ensure we are keeping up with salient human rights risks related to how abuse tactics change over time.

Strategic partners include:

[National Center for Missing and Exploited Children](#), [International Center for Missing and Exploited Children](#), [Internet Watch Foundation](#), [ECPAT](#), [Polaris](#), [Stop the Traffik](#), and Meta's Expert Advisory Group

We implemented new in-product features to raise awareness of trafficking, deter violating behavior, and offer support to victims. One example is our search redirection that detects when people search for potentially violating keywords and delivers a custom deterrence message. We are increasing our investments and our ability to identify the illicit actors, networks, organizations, and businesses that perpetrate these activities and disrupt them accordingly.

We collaborate with global NGOs to create educational campaigns, provide resources and assist victims and survivors of human trafficking. We organize internal training, and events to raise awareness of human exploitation amongst our employees.

We continually seek to expand and refine our understanding of emerging salient risks related to human exploitation as we amend and update our enforcement guidelines.

G. Election Integrity

We are [committed to protecting elections](#), to increasing authenticity, transparency and accountability for advertisers, to giving everyone a voice on our technologies and to empowering people to vote. Between 2020 and 2021, Meta made significant investments in teams and technologies to better protect elections, empower people to vote and create safe technologies to share diverse beliefs.



01. United States Presidential Election 2020

For the United States Presidential Election in 2020, we started preparations two years in advance. These enabled us to identify emerging threats and put systems in place to mitigate anticipated risks.

Among the threats we expected to see in the lead-up to the November election were: perception hacking, the shift from larger-scale operations to narrower campaigns that try to slip under the radar and leverage unwitting authentic people, and also blurring lines between authentic public discourse and manipulation by co-opting domestic groups to amplify and join influence operations. [Actions taken included:](#)

- We removed almost a dozen foreign operations backed by Russia, China and Iran that used fake accounts to deceive users and undermine trust in the United States under our policy against coordinated inauthentic behavior, announced each of these takedowns publicly, and shared information with independent researchers so that they could review and draw independent findings.
- We removed more than 265,000 pieces of content on Facebook and Instagram in the US between March 1 and Election Day for violating our voter interference policies.
- We attached informational labels to content that discussed the legitimacy of the election or claimed that lawful methods of voting like mail-in ballots would lead to fraud.
- We ran the largest voting information campaign in American history, connecting people with reliable information about voting from state and local election authorities as well as nonpartisan civic partners.
- We [worked](#) with law enforcement in the days and weeks after January 6, 2021 and continue to do so with the goal of ensuring that information linking the people responsible for their crimes is available.

02. Myanmar Election 2020

Having identified Myanmar as high priority and high risk in advance of its November 2020 election, Meta created a multi-disciplinary planning team with a human rights presence to help identify, prevent or mitigate key salient human rights risks. Meta also [took steps to prepare](#) for the election, including:

- Stood up a dedicated Integrity Product Operations Center with 24/7 coverage during the election to respond to risks and on/offline developments in real-time.
- Expanded our misinformation policy in Myanmar in an effort to help combat voter suppression, and help support the integrity of the electoral process, including working with local partners to remove verifiable misinformation and unverifiable rumors (e.g., posts falsely claiming a candidate is a Bengali, not a Myanmar citizen, and thus ineligible).
- Required all electoral and political ads in Myanmar have a “Paid for by” disclaimer attached to them to show the organization or person behind the ad.
- Worked with partners in Myanmar to verify official Facebook Pages of political parties and candidates.
- Worked with civil society partners to surface and rapidly review potentially harmful content, including misinformation that could lead to offline harm, veiled threats, and hate speech.
- Introduced an Image Context reshare product which provided people using the platform in Myanmar with additional context before they shared images that were more than a year old and could be potentially harmful or misleading.
- Introduced a new feature that limits the number of times a message can be forwarded to five. This safety feature was made available on Messenger in Myanmar in 2020.
- Utilized hate speech classifiers to proactively detect probable hate speech in Burmese and either demote, route for human review, or automatically delete depending on confidence level.
- Implemented proactive detection to surface content containing certain keywords/phrases associated with hate, misinformation, and other harms for further review.
- Introduced a third-party fact-checking program in Myanmar as part of our ongoing integrity efforts to reduce the spread of misinformation and improve the quality of the news people find online.
- Directed people to authoritative sources of election information, where they could learn how to check voter lists, as well as voting times and locations, including Union Election Commission, Vote MM and First Time Youth Voters for 2020.
- Trained civil society organizations and reporters on journalist safety, media and digital literacy, as well as Meta’s policies and third-party fact-checking programs. As part of this effort, we held a monthly television talk-show on digital literacy called Tea Talks, that focused on issues like online bullying and account security.
- Introduced tools to newsrooms in Myanmar such as CrowdTangle, a public insights tool that makes it easy to follow, analyze and report on what’s happening with public content on social media.
- Identified and disrupted six networks engaging in Coordinated Inauthentic Behavior in Myanmar. These networks of accounts, Pages and Groups were masking their identities to mislead people about who they were and what they were doing by manipulating public discourse and misleading people about the origins of content.

03. Preparation for Philippines Presidential Election 2022

In line with our approach to other major global elections, Meta made it a high priority to invest in risk mitigation and preparation efforts for the 2022 presidential election in the Philippines. These efforts were informed and guided by recommendations of the [Philippines Human Rights Impact Assessment](#), and were a demonstration of our [commitment to follow up](#) on that work. We built new products and developed stronger policies in collaboration with the Commission on Elections, election watchdogs, independent fact checkers and civil society organizations. [In addition](#):

- In partnership with the Philippine Commission on Elections and various civic organizations including the Legal Network for Truthful Elections (LENTE), we launched a civic education campaign, conducted training on Meta’s policies, and held roundtables with civil society.
- We’ve expanded our flagship digital literacy program, Digital Tayo to reach over 6.5 million people in the Philippines. Digital Tayo covers topics such as online safety, privacy, digital citizenship, news and media literacy, and launching civic campaigns.
- We supported Internews Philippines to launch the Philippine Fact Checker Incubator, to increase the capacity of local organizations for fact checking.
- To help people better understand who is behind the news they see on Facebook and Instagram, we applied labels to media outlets that are determined to be wholly or partially under the editorial control of their government, including those in the Philippines.
- As part of tackling other emerging harms, we also removed a network of over 400 accounts, Pages, and Groups in the Philippines that worked together to systematically violate our Community Standards and evade enforcement.



H. COVID-19

The world has suffered immense health and human rights consequences since COVID-19 was [declared by the WHO](#) to be a Public Health Emergency of International Concern (PHEIC) in January of 2020.

Throughout the reporting period, Meta mobilized to support public health, amplify authoritative information, connect users to essential services, as well as assist relevant agencies in their life-saving work. Three areas of Meta's COVID-19 response are particularly relevant to this report: (1) emergency due diligence, (2) content policy due diligence, and (3) additional actions to support the right to health.⁹

In 2021, we also produced a [detailed white paper](#) summarizing actions Meta took to support effective COVID-19 response while protecting privacy.

09. Further details of Meta's COVID response are available at our [Responding to COVID 19 site](#), and on the [newsroom](#)



01. Emergency Due Diligence

As the pandemic broke out, the Human Rights Policy Team worked on an emergency footing with other internal teams to align emergency response actions with UNGP guidance and prioritization frameworks. Their [actions included](#): real time decision support and advice related to the rights to health, information, and freedom of

expression; as well as rapid review of emergency product measures, adjustments to automated systems, adjustments to the at-risk countries workstream, and frameworks for prioritizing content for human review by a temporarily reduced pool of content moderators.

02. Content Policy Due Diligence

As the pandemic spread, Meta's content policy teams worked to respect and incorporate authoritative guidance related to the right to health, the right to information, and [permissible limits on freedom of expression](#).

Informed by this and other work, Meta took aggressive steps to limit the spread of misinformation about the virus and its effective vaccines and treatments, as well as to connect people with reliable health and service information. We also updated other relevant content policies to mitigate COVID-19-related risks, including (among others) our hate speech policies; coordinating harm policies; restricted goods policies; and bullying and harassment policies.

The full set of our relevant policies are available in: the [Misinformation section of the Facebook Community Standards](#), the [COVID-19 and Vaccine Policy Updates & Protections section of the Facebook Help Center](#), by the [Instagram Community Guidelines](#), and the Instagram-specific page relevant to [COVID-19 and Vaccine Policy Updates and Protections](#).

Under these policies, we remove content containing claims related to COVID-19 and vaccines or treatments that, according to public health authorities, are (a) false, and (b) likely to contribute to imminent physical harm. Imminent physical harm examples include: increasing the likelihood of exposure to or transmission of the virus, or having adverse effects on the public health system's ability to cope with the pandemic.

Based on input from experts in health communication and related fields, we took additional steps to reduce the distribution of content that does not violate Meta policies but may present misleading or sensationalized information about vaccines in a way that would be likely to discourage vaccinations.

In an important step for enforcement at scale, Meta [developed and deployed](#) AI tools to scale fact-checker's work and detect copies of rated content when someone tries to share them. We also built new computer vision classifiers to enforce bans on violating ads and commerce listings for certain medical products.



of the time people who saw the warning label did not click past to view

These efforts were effective. In April 2020, for example, [we put warning labels on about 50 million pieces of content based on around 7,500 fact-checks from partners](#). During the month of March 2020, we displayed warnings on about 40 million posts related to COVID-19 on Facebook, based on around 4,000 articles by our

independent fact-checking partners. 95% of the time people who saw the label did not click past to view. From the beginning of the pandemic to July 2021, we put these warning labels on [167 million pieces](#) of content.

We also made significant efforts to serve public health organizations and emergency responder organizations, globally and locally. We connected over 2 billion people from 189 countries to

resources from health authorities through our COVID-19 Information Center and pop-ups on Facebook and Instagram. We provided [ad credits](#) to governments and health agencies while WhatsApp enabled national health ministries around the world to communicate at scale with their citizens by onboarding them to the WhatsApp Business API, allowing them to provide timely information and respond automatically to frequently asked questions.

03. Privacy Due Diligence and Data for Good

Our [Data for Good](#) Program has been sharing privacy-protective tools and data with its 600+ partners in more than 75 countries, many of whom are working to track the spread of the virus; measure the effectiveness of prevention measures; and better understand attitudes towards vaccines.

Meta sought to address the urgent needs of public health authorities while maintaining a strong commitment to privacy and other fundamental rights. After committing to be guided by public health experts in our response, Meta decided to prioritize areas where our existing strengths made our contributions particularly impactful.

Meta understood from epidemiologists that surveys, in particular, could assist their efforts in three ways: (1) forecasting the spread of the virus; (2) assessing peoples' understanding of preventing behaviors, such as mask wearing; and (3) understanding attitudes and behaviors related to vaccinations as vaccines became available.

As a result, we launched a number of mobility datasets that were used across dozens of countries to better understand the effectiveness of lockdown orders and partnered with select academic and public health partners to create a

large scale COVID-19 Trends and Impact Survey, to which over 100 million people in 130 countries and territories responded. We also hosted public visualizations of this publicly available information to help communicate to a wide range of audiences.

While Meta was developing its strengths-based approach, the human rights and privacy teams undertook detailed human rights due diligence to guide Meta positions on responding to governmental COVID-relevant data sharing requests, the development and promotion of contact-tracing interventions, and the collection of sensitive data via chatbots in encrypted messaging services. These complex frameworks confirmed the importance of the strength-based approach. Meta did not support the promotion of automated contact tracing apps and did not seek to develop such technology, focusing its efforts instead on sharing population-level insights on changes in mobility and how this might affect transmission and the effectiveness of stay at home orders around the world.

More detailed information on these initiatives, along with key discussion questions, is available in our April 2021 white paper: [A Retrospective — Protecting Privacy in Our Covid-19 Response](#).

Conducting Human Rights Due Diligence and Disclosure

Due diligence is the second core pillar of Meta's Human Rights Policy. Pursuant to the UNGPs, we endeavor to (a) assess actual and potential human rights impacts; (b) take appropriate action and implement findings; (c) monitor implementation; and (d) communicate how impacts are being addressed. We seek to identify vulnerable or marginalized groups and engage, meaningfully, to listen to concerns, and create solutions. As part of our "show, not tell" ethos, we have sought to demonstrate meaningful transparency in our due diligence exercises.

In this section, we detail examples of the innovative approaches to diligence that we have operationalized, as described earlier:

1. Human rights impact assessments of specific countries, product initiatives, or strategic initiatives;
2. Integrated analytical approaches embedding UNGP prioritization criteria; and
3. Rapid salience analyses to support crisis situations, real-time content policy feedback, and product decision-making.

As the UNGPs were only approved by the UN Human Rights Council in 2011, the field of human rights due diligence is nascent, with rapidly evolving methodologies and best practices. For social media companies like Meta, whose products, scale of operations and impact are also evolving at a dramatic pace, the challenge of iterating methodologies capable of identifying salient risks and developing meaningful recommendations is considerable.





A. Human Rights Impact Assessments

A human rights impact assessment (HRIA) is a detailed, direct form of human rights due diligence that allows companies like Meta to identify potential human rights risks and impacts, to promote human rights and seek to prevent and mitigate risks. This work is in line with both our commitments as members of the GNI, and our responsibility under UNGP 18.¹⁰

In 2018, Meta commissioned and published the full text of an independent HRIA for Myanmar: [An Independent Assessment of the Human Rights Impact of Facebook in Myanmar](#).

In May 2020, we [published executive summaries of and recommendations from three more HRIAs for: Indonesia, Sri Lanka and Cambodia](#), along with our responses ([Indonesia](#), [Sri Lanka](#), [Cambodia](#)), at which time we also directly briefed interested and affected rights-holders, with the exception of Cambodia where COVID-19 and the security environment prevented us from doing so.

Rights-holders, in this context, included not only users of our platforms and services, but the many others whose rights were potentially impacted by online activity and conduct.

Countries targeted for human rights due diligence exercises have stemmed from Oversight Board recommendations, recommendations from prior diligence exercises, and other ad-hoc determinations based on stakeholder feedback. We are moving to a more systematic approach informed by decision frameworks and stakeholder engagement.

10. UNGP 18 requires businesses, “in order to gauge human rights risks” to “identify and assess any actual or potential adverse human rights impacts with which they may be involved either through their own activities or as a result of their business relationships” while drawing on “internal or external expertise and “meaningful consultation with potentially affected groups and other stakeholders.”

01. Philippines HRIA



Salient Risks:

The [Philippines HRIA](#) was conducted by an independent third party from February to July 2020 and published in December 2021. The assessment found that Meta technologies are widely used and have positive human rights impacts in the Philippines, as well as salient human rights risks. For example, the assessment found that Meta technologies play an important and positive role in providing access to

economic opportunities, giving voice to people and being essential tools for monitoring and defending human rights during the COVID-19 pandemic. It also highlighted salient human rights risks, including concerns about the misuse of our technologies for misinformation and disinformation, online harassment, incitement of violence, surveillance, online sexual exploitation, human organs trafficking, and extremist activities.

Recommendations and Meta's Response:

The HRIA made 40 recommendations. At the end of 2021, Meta [committed to implement 24 recommendations](#), partly implement seven recommendations, and was assessing the feasibility of another nine. We will be tracking our implementation regularly through 2022 and beyond. In response to the HRIA recommendations,

we made substantial commitments to prepare for the 2022 Philippine presidential election and strengthened our partnerships and programs with Philippine civil society organizations and government entities. Meta's full commitments are available at [Meta Response: Philippines Human Rights Impact Assessment](#).

02. Summary Disclosure - India Human Rights Impact Assessment¹¹



Purpose

In late 2019, Meta commissioned an independent human rights impact assessment (HRIA) on potential human rights risks in India related to its platforms. The project was undertaken by Foley Hoag LLP¹². The project launched in March 2020. It experienced limitations caused by COVID-19, with a research and content end date of June 30, 2021.

The assessment was conducted independently of Meta. Meta provides this summary of the report's insights and actions consistent with commitments made in its Corporate Human Rights Policy, and the United Nations Guiding Principles on Business and Human Rights, including Principle 21(c).

This summary cannot be construed to endorse, accept or adopt the independent assessor's conclusions, recommendations or analysis.¹³

11. This summary was developed after consulting The Danish Institute for Human Rights Guidance on HRIA of Digital Activities Reporting and Evaluation

12. Foley Hoag LLP is a US law firm with a human rights practice.

13. Meta's publication of this summary, and its response thereto, cannot be construed as admission, agreement with, or acceptance of any of the findings, conclusions, opinions or viewpoints identified by Foley Hoag, or the methodology that was

employed to reach such findings, conclusions, opinions or viewpoints. Likewise, while Meta in its response references steps it has taken, or plans to take, which may correlate to points Foley Hoag raised or recommendations it made, these also cannot be deemed an admission, agreement with, or acceptance of any findings, conclusions, opinions or viewpoints.



Methodology

The HRIA involved interviews with 40 civil society stakeholders, academics, and journalists. Stakeholder identities were kept anonymous and not shared with Facebook/Meta. The due diligence also involved the review of content policies, certain relevant content, as well as a survey, designed across multiple dimensions of diversity, of over 2000 individual rights holders.

The report refers extensively to international standards. These include the United Nations Guiding Principles on Business and Human Rights; the U.N. International Bill of Human Rights (including the Universal Declaration of Human Rights (“UDHR”), the International Covenant on Civil & Political Rights (“ICCPR”), and the International Covenant on Economic, Social and Cultural Rights (“ICESCR”); the U.N. Rabat Plan of Action, including freedom of expression, information, and opinion; the right to non-discrimination; and standards relating to advocacy of national, racial, or religious hatred.



Limitations

The project experienced certain limitations caused by COVID-19. HRIAs conducted on the operations and activities of digital products exist in an emerging and constantly adapting space.

Based on the guidance of human rights experts, we have produced this synthesis in order to mitigate security risks as per UNGP 21(c).

While commissioned by Meta, the independent Assessment is not adopted by Meta or a statement on its behalf. Instead, the Assessment will ideally serve as a baseline to help identify and allow Meta to address the most salient platform-related human rights risks, and help guide additional human rights due diligence.

This document is intended to be a summary of the Assessment.



Context

In 2019, civil society groups published several reports criticizing Facebook content policy rules and content moderation processes in India. Using the guidance of the UNGPs, Meta initiated a human rights due diligence project to identify and mitigate potential human rights risks.



Insights

The HRIA was an independent report by a third party.

It found Meta platforms had provided an invaluable space for civil society to organize and gain momentum, provided users with essential information and facts on voting, and also enabled important public health updates. It also noted that Meta had committed to platform safety and integrity in India, and had devoted considerable energy in recent years to promoting respect for its users, and to protecting user safety.

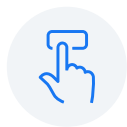
These efforts included: implementing an industry-leading set of content rules, developed in close consultation with the human rights community; substantial investments in new resources to detect and mitigate hateful and discriminatory speech; new content moderation resources, and recruitment of expert global human rights personnel.



Risks

The HRIA noted the potential for Meta's platforms to be connected to salient human rights risks caused by third parties, including: restrictions of freedom of expression and information; third party advocacy of hatred that incites hostility, discrimination, or violence; rights to non-discrimination; as well as violations of rights to privacy and security of person. It stated Meta faced criticism and potential reputational risks related to risks of hateful or discriminatory speech by end users.

The Assessment also noted a difference between company and external stakeholder understandings of content policies. It noted persistent challenges relating to user education; difficulties of reporting and reviewing content; and challenges of enforcing content policies across different languages. In addition, the assessors noted that civil society stakeholders raised several allegations of bias in content moderation. The assessors did not assess or reach conclusions about whether such bias existed.



Actions

The HRIA developed recommendations covering implementation and oversight; content moderation; and product interventions; and other areas.

As previously stated, the conclusions and recommendations made in the Assessment are not statements made on behalf of or adopted by Meta. Meta is studying these findings, however, and will consider them as a baseline to identify and guide related actions. We continue to work towards providing a safe and open platform to all users.

On implementation and oversight, Meta has taken steps to expand its human rights team; aims to ensure there is appropriate regional expertise; and will seek to build awareness of its diverse slate recruitment approach globally. We'll deepen our stakeholder engagement in a comprehensive manner.

On content moderation, Meta already actively seeks to align our content moderation and enforcement policies with international human rights law.

In 2021—22, we made content policy changes related to a) groups at risk of outing under our [coordinating harm](#) policies and b) context-specific content attacking concepts and practices associated with protected characteristics. We will also develop a prototype test for application of the Rabat Principles, and will implement it for use as a policy tool of last resort.

In addition, in 2020—2022 Meta significantly increased its India-related content moderation workforce and language support. As of the time of writing, Meta had reviewers across 21 Indian languages, with both language and cultural expertise, as well as multiple classifiers. Translating widely used languages has substantial practical benefit for our users.

With regard to transparency on content trends, Meta began publishing detailed India transparency reporting in June 2021. The [India monthly report](#) is available at our [Transparency Center](#) and covers multiple policy areas for Facebook, Instagram, and [WhatsApp](#). In addition, we are actively seeking ways to improve our transparency reporting around government requests to remove or restrict content, as reported to the Oversight Board.

We will expand participation in our [Resiliency Initiative](#), which empowers local communities with digital tools to combat hate, violence, and conflict within and beyond their networks. We will also seek to expand our counterspeech [Search Redirect Program](#), which redirects users who search using hate and violence-related search terms towards resources, education, and outreach groups that can help.

With regard to product Interventions and other measures:

In 2021—2022, Meta designed and deployed multiple product features to try to prevent or mitigate acts of abuse on WhatsApp. Meta also disclosed a detailed independent human rights impact assessment of its expansion of End-to-End Encryption in April 2022. The recommendations will help guide our [approach to safer private messaging](#): helping to prevent abuse and to safeguard people's privacy, and giving people controls to help them stay safe.

We are developing consultation processes around our annual human rights disclosures at the time of writing. We are implementing mergers and acquisitions human rights due diligence, to be incorporated into our standardized due diligence processes. We are launching our human rights website in July 2022. As per our corporate Human Rights Policy, we are sharing insights and actions from this due diligence in our annual human rights report.

B. Product-Focused Human Rights Impact Assessments

Meta commissions detailed human rights impact assessments in instances of ground-breaking new products, or systemic product changes to

current platforms. Product-focused HRIAs seek to provide guidance and identify considerations for product development.



01. End-to-End Encryption (“E2EE”)

E2EE keeps people and their personal communications safe from hackers, criminals and authoritarian regimes. In 2016, we implemented E2EE by default on WhatsApp and as an option on Messenger. In March 2019, we announced plans to extend this protection [by default across our messaging apps](#).

In October 2019, Meta commissioned Business for Social Responsibility (BSR) to undertake an HRIA of the expansion of E2EE. The subsequent two years of work took place during the reporting period. The [HRIA was published in full](#) in April 2022, accompanied by a [Meta response](#).

Salient Risks:

BSR noted a number of positive opportunities that E2EE would enable, including further protecting a diverse range of human rights such as: privacy, freedom of expression, protection against cybercrime threats, physical safety, freedom of belief and religious practices, and freedom from state-sponsored surveillance

and espionage. BSR also explained how human rights risks — such as child sexual abuse and exploitation, mis/disinformation, violent extremism, human trafficking — all operate independently of E2EE, and should be addressed separately, without undermining E2EE and focus on key marginalized communities.

Recommendations and Meta's Response:

The report [made a number of recommendations](#) to address risks and maximize the benefits of E2EE, divided among four categories of action: product, process, product policy and public policy; to wit:

- **Recommendations about specific products and features:** user reporting; UX and user testing; user education; languages; friction; opt-in account linking.
- **Recommendations for how Meta can detect and mitigate human rights risk:** process harm prevention strategies, such as metadata analysis; Machine Learning (ML) techniques for proactive detection; holistic child rights strategy; investigate client-side scanning techniques; assess impacts of cross-app communications.
- **Recommendations for product policy changes:** consistent privacy policies and improved transparency; define content standards; ML explainability; improved user reporting; user appeals transparency; grievance mechanisms.
- **Recommendations for how Meta should engage external stakeholders:** advocate for end-to-end encryption; engage policy makers; engage stakeholders; collaborate with researchers; quantify harms; collaborate with industry; train law enforcement.

Meta's [response](#) details our commitment to implementing 34 of the recommendations, partly implementing four, assessing the feasibility of another six and taking no further action on one — the pursuit of client-side scanning technologies.



The report concludes that client-side scanning technologies as they exist today would undermine the integrity of E2EE and disproportionately restrict people's privacy and other human rights.

Meta is committed to implementing the vast majority of the recommendations and is working diligently toward our plans for expanding E2EE as a means to help protect people and support their human rights.

C. Due Diligence Exercises

Meta uses rapid, decision-useful due diligence exercises to provide salient risk assessments for moments of crisis, immediate product development needs, and ongoing content policy decisions.

This work can build on the diligence exercise designating a country high priority, or be responsive to emerging needs. It can look back at moments of crisis to provide insights on improved processes going forward; or

run simultaneously to the crisis or product development to provide immediate, real-time feedback to inform decision-making. These assessments can be conducted in-house, or with a third-party. They are designed to run for shorter periods of time than our formal HRIAs (*i.e.*, weeks/months compared to years, on average). While their form can differ, their function is uniform: actionable, decision-useful, and immediate insights to guide the management of salient risks.



01. Ray-Ban Stories

Salient Risks:

In September 2021, in partnership with EssilorLuxottica, Meta introduced [Ray-Ban Stories](#), a first-generation pair of smart glasses that have cameras in the frames that can take photos and videos, which you can download to a companion app and then upload to social media.

Prior to launch, to mitigate human rights issues posed by people who may misuse such a device, Meta undertook an internal analysis

of the possible human rights implications of this product. Salient risks identified included: informed consent from bystanders, the safety of people wearing the device, as well as its effect on vulnerable groups that might be at increased risk of adverse effects from its operation, like women and children, human rights defenders, or minority groups. It also examined privacy risks and future, ongoing risks posed by any new technology, from visual search to the cloud storage of data.

Recommendations and Meta's Response:

The saliency assessment began by noting some positive human rights impact of hands-free and voice-controlled photo and video capabilities, including help for people with disabilities, or aiding accountability for human rights violators. It then noted if and where use of the device could implicate various human rights treaties, including things like: whether it could be seen by people with visual impairments, and if not whether that could violate the rights to dignity and privacy, or non-discrimination.

Its recommendations included other acts of human rights due diligence, including expanding user research to focus on vulnerable bystanders, like homeless communities, minority groups and human rights defenders, as well as recommendations to conduct further field-level focus groups to understand the product's impact *in situ*.

To mitigate these issues, the due diligence exercise recommended exploring several courses of action in consideration of bystanders, such as: developing an acceptable use policy, a do-not-disturb function, other signals for bystanders or the option to tag content as being from the device when it is shared.



D. Integrated Analytical Approaches

As part of our broader human rights due diligence efforts, we put into practice various integrated, analytical approaches that help us comprehend our most salient risks.



01. Countries At Risk

As we note in Section 02 of our [Human Rights Policy](#),

“We pay particular attention to the rights and needs of users from groups or populations that may be at heightened risk of becoming vulnerable or marginalized.”

In identified countries (and priority elections), we take a more customized comprehensive due diligence, risk mitigation and response approach — acting quickly to initiate the process to remove content that violates our policies and taking protective measures, including deploying country-specific support. Since 2018, this work has been led by teams with expertise in human rights, hate speech, women and child protection, misinformation, and polarization. Many have lived or worked in high-risk countries and speak relevant languages.

Beginning in 2019, we developed [a human rights due diligence process for reviewing and](#)

[prioritizing countries with the highest risk of offline harm and violence](#), every six months. To assess long-term conditions on the ground, a cross-functional team of data scientists, political scientists and regional experts review quantitative and qualitative data twice a year drawn from over 60 sources including [Varieties of Democracy \(V-Dem\)](#), [Uppsala Conflict Data Program](#), the [United States Holocaust Memorial Museum’s Early Warning Project](#), the [Armed Conflict Location & Event Data Project](#), and the World Bank. This data can include information regarding civic participation and human rights, societal tensions and violence, and the quality of relevant information ecosystems.

UNGP-Informed Factors to Determining Countries at Risk:

- **Long-term conditions and historical context:** We rely on regional experts, platform data and data from more than 60 sources to assess the long-term conditions on the ground.
- **How much the use of our products could potentially impact a context:** We prioritize countries based on a number of factors, including: where our apps have become most central to society, such as in countries where a larger share of people use our products; where there is potential for an increase in offline harms; and where social media adoption has grown.
- **Current events on the ground:** We give special consideration to discrete events that might magnify current societal problems, such as local risk or occurrence of atrocity crimes, polarizing elections, episodes of violence, and COVID-19 vaccination and transmission rates.

Guided by the above, we provide specialized support and focus teams on countries most at risk by working to drive human rights to the front of our product, policy, operations, and corporate decision-making. Our focus is threefold: remove content that violates our policies, respect people's right to free expression, and help to keep people safe both online and offline.

As part of our work, we [took the following actions](#):

- **Understand and engage with local contexts and communities** to ensure our enforcement accounts for local context. We expanded our global network of third-party fact-checkers, invested significant resources in more than 30 countries with active conflict or societal unrest, and together with UN partners and dozens of local and global NGOs, we developed programming to make online engagement safer.
- **Develop and revise policies to prohibit harmful content and behavior** by constantly evaluating and refining our policies to address evolving nuances of hate speech, identify groups at heightened risk of violence and [human rights abusers](#), or the potential for rumors and misinformation to contribute to physical harm, particularly in countries where ethnic and religious tensions are present.
- **Rapidly enhance our technology and enforcement to help keep our community safe** when the risk of impact is greater, *i.e.* in advance of elections and during periods of heightened unrest. For example, we may reduce or restrict distribution of content from accounts previously found to post content violating our policies, while our teams investigate. Also, we may offer additional features, *e.g.* providing people in Afghanistan following the Taliban takeover, the ability to lock their profile as an extra layer of privacy, security and protection.
- **Improve our moderation across languages** by adding more expertise.

In 2021, we hired content moderators in 12 new languages, including Haitian Creole, Kirundi, Tswana and Kinyarwanda.

i. Our Program In Action: Case Studies

Recent examples of our work in this area include our efforts in [Ethiopia](#), [Mexico](#), [Myanmar](#), [Afghanistan](#), and Haiti. We highlight two case studies which exemplify this work in action: (1) Myanmar which held elections in 2020 and

experienced a military coup in early 2021; and (2) Ethiopia which, since November 2020, has experienced a brutal civil war characterized by severe human rights violations and massive humanitarian needs.

a. Myanmar

Evolving Challenge:

Ahead of the November 2020 elections in Myanmar, and given the Myanmar Military's (the "Tatmadaw") history of atrocities against the Rohingya people, Myanmar was identified as high priority and at high risk for human rights violations. Though Meta's efforts to protect platform integrity during the election were successful, the political climate substantially deteriorated, culminating in the 2021 coup d'état by the Tatmadaw that unleashed a wave of repression, attacks on peaceful demonstrations, and the silencing of dissenting voices.

Diligence:

A cross-functional, multi-disciplinary team came together to prepare for the elections; later, worked to identify the salient risks due to the coup, and took steps to help prevent and mitigate those risks.

Action:

Following the coup, the military government unleashed a wave of human rights violations against the people of Myanmar, protestors, human rights defenders, journalists, and others. We treated the situation as an emergency. Our Integrity Operations Center ran around the clock, bringing together subject matter experts from across the company, including Myanmar nationals with native language skills, to monitor and respond to threats in real time.

As a result, Meta took a series of [strong actions](#):

- Meta banned the Tatmadaw and military-controlled state and media entities from Facebook and Instagram, as well as ads from

military-linked commercial entities, and removed Pages, Groups and accounts representing military-controlled businesses. We based our actions on:

- Extensive documentation by the international community of these entities' direct role in funding the Tatmadaw's ongoing violence and human rights abuses, which also formed the basis for escalating sanctions imposed by the US, EU, and other governments.
- The UN Fact-Finding Mission on Myanmar's 2019 report on the economic interests of the Tatmadaw, in line with the UN Guiding Principles on Business and Human Rights.
- We reaffirmed Meta's ban on 20 military-linked individuals and organizations in 2018, including Commander-in-Chief Min Aung Hlaing, for their role in severe human rights violations; and the decision to remove at least six Coordinated Inauthentic Behavior networks run by the Tatmadaw from 2018 to 2020.
- Meta implemented a Myanmar-specific policy removing praise, support and advocacy of violence by Myanmar security forces against civilians and/or arrest of civilians by the military and security forces on our platform, under our [coordinating harm and publicizing crime policy](#).
- We offered a "Lock Your Profile" feature for Facebook, providing additional safety, privacy and security for our on-the-ground users.
- Since the coup, we disabled the Tatmadaw True News Information Team Page, and MRTV and MRTV Live Pages for continuing to violate our policies. We also reduced the distribution of content on at least 23 pages and profiles controlled and/or operated by the Tatmadaw so fewer people see them.

b. Ethiopia

Evolving Challenge:

As of this report's publication, Ethiopia is undergoing a conflict that began in November 2020 between the government of Ethiopia and forces in the northern Tigray region. As the turmoil spreads, the situation remains volatile. Ethiopia is an especially challenging environment to address these issues, in part because there are multiple languages spoken in the country. Meta and its platforms are key avenues for communication in Ethiopia.

Diligence:

Since early 2020, human rights due diligence for Ethiopia has included field-based information gathering and research; focused stakeholder engagement; and formal due diligence. In 2021, Meta commissioned forward-looking human rights due diligence to help identify existing and prospective human rights risks and inform internal decision-making. We engaged activists, human rights defenders, journalists, civil society groups throughout Ethiopia and the diaspora, adding to the insights gained regularly from human rights groups, NGOs, local civil society organizations, and other regional experts.

Action:

Using learnings from our due diligence, Meta increased efforts to tackle harmful content (e.g., violence and incitement, hate speech, coordinated inauthentic behavior, misinformation and harm), and addressed elections-related risks and content involving dangerous organizations and individuals. [We also:](#)

- **Improved and simplified our reporting tools** to make it easier for Ethiopian users to report possible violating content, so we can investigate.
- **Rolled out a [media literacy campaign](#)**, aimed at educating and informing people on how to detect potential false news, and ran billboard advertising campaigns across Addis Ababa, the first of its kind in Africa, focused on informing and educating people on how to stay safe online and use social media responsibly.



- **Partnered with the Center for African Leadership Studies** to implement “[My Digital World](#),” a series of live webinars on topics such as online safety, privacy, digital citizenship, news and media literacy.
- Established **dedicated reporting channels** for specialized international and local human rights and civil society organizations to quickly review problematic content they identify for possible violations.
- **Launched a [Lock Your Profile](#) feature (as in Myanmar) for Facebook**, providing additional safety, privacy and security for our on-the-ground users.
- **Increased our investment in proactive detection technology** that helps identify violating content such as hate speech before it is reported, including adding in Amharic and Oromo capabilities.
- **Hired more Ethiopian nationals** with native language skills to monitor and respond to threats in real time and onboarded more Trusted Partners to provide contextualized advice.

02. Security Policies

Meta strives to combat coordinated inauthentic behavior, defined as coordinated efforts to manipulate public debate for a strategic goal where fake accounts are central to the operation. Such content on Meta’s platforms could lead to salient human rights risks advanced through coordinated posts that:

- attack people on the basis of their opinion, beliefs, or protected characteristics;
- are connected to harm;
- depict violence;
- bully or harass users;
- encourage self-injury, misinformation or disinformation; or
- exacerbate conflict, corruption, and instability in conflict affected and high risk areas.

Meta performs ongoing diligence to review, assess, and ultimately help remove deceptive campaigns around the world. When we find campaigns that seek to mislead people about who they are and what they are doing using fake

accounts, we remove both the inauthentic and authentic accounts, Pages, and Groups directly involved. Using both automated and manual detection, Meta continues to remove accounts and Pages connected to networks that we took down in the past.

Under our UNGP and GNI commitments, this work is informed by stakeholders and performed with transparency. Since 2017, Meta has reported on over 150 influence operations and published information on threats we see from nation states, commercial firms or unattributed groups. We share information across the world with investigative journalists, government officials, independent researchers, and industry peers to better understand and expose internet-wide security risks, including ahead of critical elections. In 2021 alone, we removed 52 networks that originated in over 30 countries and expanded our beta research platform (containing over 100 data sets) to researchers studying influence operations.



i. Combating Emerging Threat Actors

The global threats we tackle have significantly evolved since we started sharing our findings about [Coordinated Inauthentic Behavior](#) in 2017. Adversarial networks are always looking for new ways to evade our policies. To account for this constantly shifting environment, we build our defenses to adapt and incorporate new layers of defense. Our goal over time is to make these behaviors more costly and difficult to hide, and less effective.

- **New Defenses against Coordinated Mass Harassment (“Brigading”), Abusive Reporting and Coordinated Violations of our Community Standards**

In December 2021, [we shared](#) our first report that brought together multiple network disruptions for distinct violations of our security policies: Coordinated Inauthentic Behavior and two new protocols — Brigading and Mass Reporting.



Brigading:

We remove any adversarial networks we find where people work together to mass comment, mass post or engage in other types of repetitive mass behaviors to harass others or silence them.

As detailed in the [December, 2021 Adversarial Threat Report](#), we removed a network of accounts that originated in Italy and France and targeted medical professionals, journalists, and elected officials with mass harassment. We also removed a network of accounts in Vietnam for violating our Inauthentic Behavior policy against mass

reporting by coordinating false reports against activists and others who publicly criticized the Vietnamese government in an attempt to have these users removed from Facebook.



Mass Reporting:

We remove any adversarial networks we find where people work together to mass-report an account or content to get it incorrectly taken down from our platform.

We are starting to remove coordinated violating networks when we find people — whether they use authentic or fake accounts — working together to violate or evade our [Community Standards](#).



ii. Protecting Users Against Espionage

Cyber espionage actors typically target people across the internet to collect intelligence, manipulate them into revealing information, and compromise their devices and accounts. When

we disrupt these operations, we take down their accounts, block their domains from being shared on our platform, and notify people who we believe were targeted by these malicious groups.



Our Campaign Against Surveillance for Hire

In December 2021, [we announced our action](#) against surveillance-for-hire entities operating on our platform, as part of a broader campaign effort that also included a call for further regulation and industry collaboration to counter this threat. Surveillance-for-hire entails commercial players duplicitously engaging in reconnaissance, engagement and exploitation of users on social media platforms, which can lead to salient human rights risks. During the reporting period, we took action against seven different surveillance-for-hire entities who indiscriminately targeted people in over 100 countries on behalf of their clients. These providers are based in China, Israel, India, and North Macedonia, [among others](#).



Broader Response to Abuse by Surveillance-For-Hire Groups

While cyber mercenaries claim that their services and surveillanceware are meant to focus only on criminals and terrorists, our own investigation, [independent researchers](#), our [industry peers](#) and [governments](#) have [demonstrated](#) that targeting includes journalists, dissidents, critics of authoritarian regimes, families of opposition and human rights activists. In fact, for platforms like ours, there is no scalable way to discern the purpose or legitimacy of such targeting. This is why we focus on enforcing against this behavior, regardless of who is behind it or who the target might be.

To support the work of law enforcement, we already have authorized [channels](#) where government agencies can submit lawful requests for information, rather than resorting to the surveillance-for-hire industry. These channels are designed to safeguard due process and we [report](#) the number and the origin of these requests publicly.

To [protect people](#) against cyber mercenaries operating across many platforms and national boundaries, we engage in a collective effort with other platforms, policymakers and civil society to counter the underlying market and its incentive structure.

Providing Access to Remedy

The third pillar of Meta's [Human Rights Policy](#) concerns the provision of remedies for human rights impacts.

Access to Remedy is a key component of the [UNGPs](#), with regard to both State and non-State-based grievance mechanisms. Affected rights-holders should be able to claim remedy without fear of victimization.

At Meta, we strive to provide pathways for stakeholders to report concerns, for Meta to review them, and for Meta to provide remedy and remediation consistent with UNGP 31. Meta maintains multiple grievance pathways, identified in the [Code of Conduct](#), in the

[Help Center](#) on platforms and apps, including an appeals process to the first-of-its-kind [Oversight Board](#). Anonymous reporting for Meta personnel is available through the [Whistleblower and Complaint Policy](#).

We seek to provide meaningful transparency, with routine publication of policies with human rights implications, information related to how content determinations can be appealed, and decisions of, and responses to, the Oversight Board. When content is removed, we notify users and clearly identify pathways for content removal decisions to be appealed. Data regarding removal decisions and other aspects of policy enforcement are published through the [Transparency Center](#).



A. The Oversight Board

After extensive global consultation, Meta founded the [Oversight Board](#) in 2020. The Oversight Board was born, in part, to respond to concerns that Meta needed an independent third party to adjudicate violations of its rules, provide policy recommendations and address access to remedy.

The Board [consists](#) of experts from a variety of cultural and professional backgrounds, chosen because they are “experienced at deliberating thoughtfully and collegially, skilled at making and explaining decisions based on a set of policies or principles, and familiar with digital content and governance.”¹⁴

As its [charter makes clear](#), the purpose of the board is to: “[p]rotect free expression by making principled, independent decisions about important pieces of content and by issuing policy advisory opinions on [Meta’s] content policies.”

The independence of the Oversight Board is strongly protected by its Charter, its Bylaws, its members, and its funding.

Prior to its launch, Business for Social Responsibility completed a [Human Rights Review](#) of the Oversight Board. This review made a series of recommendations, including prioritizing cases that present the most severe human rights harms, using the UNGP’s scope, scale and remediability criteria. The review also recommended that the human rights impacts of each case be reviewed, and that the human rights principles of legitimacy, necessity, and proportionality be used. One year later, a follow up [human rights review](#) by Business for Social Responsibility found that “the Oversight Board’s governing documents provide an increasingly valuable framework for [taking a] human rights-based approach to content decisions.”

The Oversight Board plays a critical, industry-leading role in the fulfillment of Meta’s responsibility under the UNGPs to provide “Access to Remedy,” which requires business

enterprises to “establish or participate in effective operational-level grievance mechanisms for individuals and communities who may be adversely impacted [by their operations].” This not only makes it possible to “identify adverse human rights impacts as a part of an enterprise’s ongoing human rights due diligence,” it also “makes it possible for grievances, once identified, to be addressed and for adverse impacts to be remediated early and directly by the business enterprise, thereby preventing harms from compounding and grievances from escalating.”

For its part, the Meta Human Rights Policy Team supported this project by working with colleagues on initial due diligence, the Charter, member selection, initial referrals, board education, training of Meta staff for interacting with the board, due diligence on transparency and follow up to board recommendations.

Following Oversight Board decisions, Meta started the process of reinstating [identical content with parallel context](#) in the following cases: [Uyghur Muslims](#), [Hydroxychloroquine](#), [Azithromycin and COVID-19](#), and a [Nazi Quote](#). These actions affect not only content previously posted on Facebook and Instagram but also future content. For cases where the board upholds our judgment, we continue to ensure identical content with parallel context remains either up or down, in line with the board’s decision.

The Oversight Board has [directly called for transparency](#) when circumstances require it, and [brought in for discussion](#) people critical of the company.

14. <https://oversightboard.com/meet-the-board/>

In another example of its focus on transparency, when Meta referred to the Board a question regarding its restriction of then-US President Donald Trump’s access to posting content on his Facebook page and Instagram account, the Board observed that Meta failed to mention a “cross-check” system, and that, “Given that the referral included a specific policy question about account-level enforcement for political leaders, many of whom the Board believes were covered by cross-check, this omission is not acceptable.”

The Board, via its case decisions and policy recommendations, is defining what a human rights-based approach to content means in practice, to a level of detail, practical application and intellectual rigor, that did not previously exist.

The Oversight Board publishes a quarterly [transparency report](#) highlighting its extensive efforts adjudicating the most difficult and pressing speech issues facing Meta. Cases before the board include those submitted by Meta, users or others.



B. Code of Conduct

[Meta’s Code of Conduct](#) sets out guidelines for business conduct required of all Meta personnel, and is accompanied by multiple channels to report concerns or violations, and potential remedies. It helps embed human rights into our overall governance by including respect for human rights

as a core part of our expectations for all Meta personnel.

The ability to make anonymous reports is covered by Meta’s [Whistleblower and Complaint Policy](#). Retaliation or intimidation against individuals reporting issues in good faith is strictly forbidden.

Protecting Human Rights Defenders



The fourth pillar of our [Human Rights Policy](#) is dedicated to protecting Human Rights Defenders, as defined by the [UN Declaration on Human Rights Defenders](#), recognizing them as a priority group for support and protection.

We commissioned internal due diligence in 2019–20, the goal of which was to take a first step toward designing a human rights defender program, mapping existing work, identifying gaps and opportunities, and suggesting possible next steps. We then established an internal Human Rights Defender working group, which has significantly improved communication and information-sharing across teams and helped to identify gaps, priorities, and potential interventions to better support human rights defenders. These interventions include steps related to safety, security, research, policy development, and product or operational support.

We're building on [our existing work to protect defenders' accounts](#) — efforts that include [combatting advanced threat actors](#) targeting them, protecting them from incorrect content removals using [Cross-Check](#), offering advanced [security options](#), taking steps to thwart unauthorized access to the accounts of defenders who are arrested or detained, and partnering with human rights organizations on outreach and training.

One important workstream has been our work on civic actors, who may face higher risks due to their contributions to civic affairs, either online or offline. These risks include threats to privacy, reputation, well-being or physical safety. Interventions applied for civic actors include harassment monitoring, enhanced account security and impersonation takedowns.

WhatsApp's end-to-end encryption is [widely recognized](#) to be essential for the protection of human rights defenders and journalists — as well for racial and religious minorities, LGBTQ+ communities and other disadvantaged groups.

Engagement with human rights defenders (and related UN actors) has increased significantly. We are driving regular engagement at the local and global levels to inform our understanding of related risks, mitigations, and challenges. Our engagements have included [digital security training for activists](#) in a number of countries.



A good example of innovative work in the Africa region has been the creation of the Sub-Saharan Africa (SSA) Women’s Working Group, which was convened in 2021 for women activists across SSA working on issues that intersect with gender equality and digital spaces. Participants represented grassroots, regional and international organisations covering issues ranging from online gender-based-violence, feminist rapid-response services, online safety for the LGBTQ+ community, tech-feminism and tech law & governance.

The working group helped us improve policies and products affecting women users on Facebook platforms and include these voices in ongoing work that requires engagement for content policies and other platform developments and products affecting this constituency.

When Meta launched our [Human Rights Policy](#) in March of 2021, we also committed to [create a fund](#) that will give offline assistance to human rights defenders facing critical threats and support new digital security efforts.



This fund was formally launched in [November 2021](#), as the Human Rights Defender Fund and Journalist Safety Initiative, in partnership with [Civil Rights Defenders](#) and [International Center for Journalists](#) and is being piloted in the Asia-Pacific region.

The fund supports devices and security technologies, temporary relocation, and emergency legal and safety assistance, as well as mental health support to human rights defenders and journalists. It also enables the development of peer-to-peer support programs and relevant mental health practitioner networks. The fund is being distributed over a two-year period, and managed independently.

We have made good progress in integrating the protection of human rights defenders into our responses in countries at-risk and will continue to look to develop and roll out new policies and products in more regions — especially those identified as most critical for the well-being of human rights defenders and other civic actors.

Governance, Oversight, Accountability

The fifth core pillar of our Human Rights Policy reflects our commitment to UNGP 16 to embed respect for human rights throughout our enterprise.



A. Governance Model

Meta's [Human Rights Policy](#) and its commitments to human rights are enterprise-wide. Nick Clegg (President, Global Affairs) and Jennifer Newstead (Chief Legal Officer) oversee policy implementation. Salient risks are reported periodically to the Audit and Risk Oversight Committee of the Meta Board of Directors.

The global Human Rights Policy Team drives application of the Policy, setting the company's human rights agenda, identifying priorities, shaping strategy, inspiring and informing colleagues, and engaging with human rights activists and experts. This team was created in 2019, with the appointment of Miranda Sissons as its inaugural director and by the end of 2021 had four full-time staff, with several more positions onboarding in 2022.¹⁵

In a company of tens of thousands of employees and contingent workers, and with billions of people using our apps and services, a key challenge for the Human Rights Policy program is to ensure that we leverage our expertise and resources for maximum impact.

We seek to ensure that salient human rights issues are directed to the most relevant decision-making forum. The Human Rights Policy Team functions in an umbrella governance structure: consulting with and convening the functional units implementing human rights-informed policy and product changes.

15. As of March 2022 the global Meta Human Rights Policy Team consists of 9 full time employees.

Teams working on community operations, law enforcement and security, responsible innovation, product integrity and public policy (among others) regularly identify, prevent, or mitigate salient human rights risks. Similarly, teams focused on civil rights, Africa, the Middle East and Turkey, sustainability, International Organizations, and Stakeholder Engagement also work to implement policy goals.

Key activities to ensure proper governance over human rights responsibilities include:

- Convening and driving the work of various cross-functional, inter-departmental working groups dedicated to the implementation of the Human Rights Policy;
- Joining cross-functional, inter-departmental working groups to provide human rights perspective and principles to their work;

- Driving human rights due diligence processes with key product and public policy colleagues;
- Participating actively in the at-risk country work, ensuring that human rights considerations are effectively prioritized;
- Coordinating work on human rights defender protection;
- Representing Meta at various human rights fora including the GNI, [B-Tech](#), and with key UN and NGO human rights agencies.

[Meta's Board of Directors' Audit & Risk Oversight Committee's charter](#) provides for the committee's oversight of programs, policies and risks related to social responsibility, which includes oversight of human rights-related risks. Beginning in 2021, the Director of the Human Rights Policy Team updates the Audit and Risk Oversight Committee at least annually.

In the course of the last two years, Meta's teams have worked across a wide range of human rights issues, including but not limited to:

- Business and Human Rights;
- Freedom of Expression (particularly proportionality);
- Privacy;
- Freedom of Assembly;
- Political participation;
- Civic engagement;
- Non-discrimination;
- International humanitarian law;
- International criminal law;
- Transitional justice processes, particularly process design for truth seeking and reparations initiatives;
- Agile software development processes;
- Atrocity crime prevention;
- Child rights;
- Regional human rights practices and instruments;
- Modern slavery;
- Lawful use of lethal force and law enforcement principles;
- Right to health;
- Machine learning techniques;
- The San Remo Principles;
- Open source evidence investigation;
- Quantitative and qualitative research processes and indicators;
- Comparative law;
- Know your customer (KYC) processes; and
- Adult learning practices.

B. Stakeholder Engagement



Our [Human Rights Policy](#) is founded on deep engagement with rights-holders and other stakeholders.¹⁶ Stakeholder Engagement is thus relevant to every pillar of the policy.

Stakeholder engagement is a company-wide effort, integrated into several company processes. When developing the Human Rights Policy, for example, the Human Rights Policy Team engaged extensively with human rights experts and activists, including activist investors. Team members are in regular contact with human rights organizations, activists, academics and others to hear concerns and criticisms, seek advice and brief on ongoing developments.

As described in earlier sections of this report, such engagement includes consultations and co-design with human rights defenders on measures to protect account security, engagement on coordinated inauthentic behavior and espionage, roundtables with regional civil society actors ahead of elections, and much more.

Stakeholder engagement is also an important part of Meta's human rights due diligence process. It enables us to identify and understand the human rights impact of our products and services as

well as develop appropriate mitigation strategies, including fair and transparent policies. During the reporting period, we also sought to pioneer new methodologies of direct rights holder consultation. For example, we combined a large scale quantitative survey of rights holders with traditional methods of stakeholder consultation in our human rights due diligence in the [Philippines](#).

In another important example, the Content Policy Stakeholder Engagement team strengthens our content policies by bringing global knowledge and feedback into the policy development process. This team builds relationships with NGOs, academics and other thought leaders, guided by our [core principles](#) of inclusivity, expertise, and transparency.

In all, the team covers the following policy areas: Community Standards, advertising policies, product policies, misinformation, content ranking, and news integrity.

[Work with stakeholders crafting our policies](#) includes:

¹⁶. By "stakeholders" we mean rights-holders, all organizations, and individuals who are impacted by, and therefore have a stake in, Meta's policies

Engagement on Stereotypes

We created a policy on harmful stereotypes. As part of our approach to [hate speech and harmful stereotypes](#), we convened several academic roundtables and connected with external stakeholders who helped us identify the risks associated with certain stereotypes in their communities.

Understanding State Media

To inform our policy definition and criteria for [state media](#), we consulted with more than 65 global experts specializing in media, governance and human rights.

This input was crucial to helping us understand the different ways and degrees to which governments exert editorial control over certain media entities.



Guidance on “Public Figures”

We adopted a definition of “public figure” for our [bullying and harassment](#) policies based on engagement with academics, civil society organizations, free expression experts, human rights defenders, journalists and creators/influencers.

Recommendation Guidelines

In developing our Recommendations Guidelines for the [Facebook app](#) and [Instagram](#), we consulted with 50 leading experts who specialize in recommender systems, expression, safety and digital rights.

C. Program Spotlight: Trusted Partners

Meta's Trusted Partner (TP) program is an escalation channel made up of over 400 key non-governmental, not-for-profit, humanitarian, and international organizations in 113 countries who report content, accounts, and behavior that we review in context.

Trusted Partners are critical allies in our efforts to understand the local impact of our services, stay abreast of emerging trends, correct mistakes made in scaled review, and improve policies and enforcement in specific policy areas such as misinformation and harm. Their reports provide valuable insights, data points, and edge cases, which help inform policy development processes and help keep our users safe both on and offline.

In the past, TPs have alerted us to developments and risks that our internal monitoring efforts had missed. In Myanmar, for example, TPs provided us with additional context about the recent coup and online risks that emerged in its aftermath. In Ethiopia, our TPs helped us combat harmful misinformation and hate speech stoking inter-ethnic violence. In Afghanistan, our civil society partners helped us to mitigate the targeting of international aid workers and other vulnerable persons in our platforms.

The Trusted Partner program also provides a vehicle for amplifying the voices of marginalized communities, incorporating relevant information about the risks they face on our platforms, and providing support to under-represented and disproportionately-impacted groups.



D. Engagement with International Organizations

Part of our accountability process under our governance pillar is transparent, collaborative, and reciprocal engagement with the UN. Our International Institutions and Relations team supports engagement across the UN and intergovernmental organization system, including dialogue with UN human rights actors, which continues to broaden and deepen both at global and country level.

As long-running human rights issues continue to manifest online, we see this collaborative discourse with the UN human rights system as particularly valuable — this includes thematically focused actors such as UN Human Rights Council-mandated Special Rapporteurs, but also broader dialogue on tech and human rights across the UN Secretariat and with specialized agencies such as UNESCO.

- The International Institutions and Relations and Human Rights Policy Teams have met regularly with the UN Office of the High Commissioner for Human Rights since 2018, and we have been active members of the [B-Tech Project](#).
- In addition, we meet with various entities of the UN system including the Office of the Special Advisor for the Prevention of Genocide, as well as UN Special Rapporteurs on: Freedom of Expression, Human Rights Defenders, Protection of Minorities, Myanmar, Racism and Racial Discrimination and Freedom of Religion and Belief to discuss the intersection of human rights and technology, and its relevance to the actions of the UN system across peace and security, human rights, and humanitarian agendas.
- We also work with UN entities as trusted partners at both global and country level — including Haiti, Bosnia, and Ethiopia. We work with them to mitigate content risks and to better understand the links between on-platform conduct and off-line harms. We are collaborating on issues of racial justice, the implementation of the UNGPs and the critical importance of end-to-end encryption. We are contributing our expertise and experience to a variety of human rights reporting and norm

development processes; and have briefed and shared insights from our at risk countries with multiple relevant groups.

- Meta briefed the UN Security Council’s October 28, 2021 [Arria Formula Meeting](#) on hate speech and social media.
- A key part of our [engagement](#) with the UN system and other multilateral actors has been to encourage a strong defense of an open and unified internet, with an emphasis on the many negative human rights impacts of internet fragmentation. Meta is also part of the [Internet Fragmentation](#) group.
- Meta has been represented in several human rights and hate speech related panels, including by senior executives, to further discourse and thinking around tackling hate speech and misinformation.
- As part of UNESCO’s [Social Media 4 Peace](#) project, we have worked with UN representatives in Bosnia and Kenya to discuss with civil society and government stakeholders, ideas of how social media platforms can promote peace.
- Since June 2021, we have voluntarily disclosed 16 productions with organic content from hundreds of Pages, and Basic Subscriber information for hundreds of Pages and Profiles, to the Independent Investigative Mechanism for Myanmar (IIMM) which is seeking accountability for the most serious international crimes and violations of international law committed in Myanmar.
- In September 2021, Meta joined the [UN Global Compact](#), the world’s largest corporate sustainability initiative. This marked an important milestone in Meta’s commitment to human rights, labor, environment and anti-corruption. To take actions that advance societal goals, Meta has committed to the Global Compact’s [10 Principles](#) on Human Rights, Labor, Environment, and Anti-Corruption.

A Final Note

“The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”

ISAAC ASIMOV



Isaac Asimov once wrote, “The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”

While this report was being written, a war in Ukraine was fiercely waged, precipitated by a Russian invasion. As with the Arab Spring of 2011, content on social media is seen to have a profound role in shaping how this conflict is perceived, and potentially, its dynamics.

This conflict, along with other crises in countries across the world, underscores how important it is for Meta and other companies to live up to their human rights commitments — for wisdom to always accompany knowledge.

We hope this report gives detailed insight into the work Meta has done to identify, mitigate, and prevent human rights risks from 2020 — 2021. We look forward to more work, and more reporting, to come.