

December 2021

Meta Response: Philippines Human Rights Impact Assessment

TABLE OF CONTENTS

Overview	3
Findings	7
HRIA Recommendations	10
Overview	10
1. Misinformation and Disinformation	11
2. Online Harassment	17
3. Incitement to Violence	21
4. Surveillance	23
5. Online Sexual Exploitation	24
6. Trafficking	29
7. Extremist Activity	30
8. Corporate Accountability	31
9. Updated HRIA Recommendations (from September 2021)	32
Appendix: Detailed Summary of Efforts Undertaken 2017 - 2021	36
I. Digital Citizenship	36
II. Elections Preparedness	37
III. Economic Inclusion	38
IV. Mental Health Resources	39
V. Policy Improvements	39
VI. Scaling Safety and Security	40
VII. Fighting Misinformation	42
VIII. Cybersecurity Enforcement	43
IX. Support for Journalists and Human Rights Defenders	44

Overview

Human rights conditions in the Philippines have long been fragile.¹ Marred by a legacy of brutal authoritarian rule that ended only in 1986, and a history of simmering insurgencies, conditions have worsened in recent years—and deteriorated further in 2020.²

The administration of current president Rodrigo Duterte has been marked by a bloody drug war and extrajudicial killings; the widespread government-sanctioned harassment and “red tagging” of human rights defenders, journalists, and political opponents; and the implementation of overbroad and vague government powers through 2020 anti-terrorism legislation.³ Online misinformation and disinformation are widely reported to be ongoing threats and noted in the executive summary of this assessment by [Article One](#). Trafficking and sexual exploitation, including of children, continue to be major challenges.

In 2021, Freedom House’s influential [Freedom in the World](#) index gave the Philippines a score of 56 out of 100, classifying it as “partly free”, declining from 61 in 2019. In the related [Freedom on the Net](#) report, focused more narrowly on internet freedom issues, the country scored 65 out of 100. Other human rights indices highlight further cause for concern. In 2021, the World Justice Project’s [Rule of Law Index](#) placed the Philippines near the bottom, at 102 out of 139 countries covered, while the Reporters Without Borders 2021 [World Press Freedom Index](#) placed it 138th.

Filipinos are heavy users of social media. Meta’s products are widely used in the country both by individuals and small and medium-sized businesses. They are also extensively used by Filipinos

¹ Meta’s publication of the HRIA’s Executive Summary, and its response thereto, cannot be construed as admission, agreement with, or acceptance of any of the findings, conclusions, opinions or viewpoints identified in the Philippines Human Rights Impact Assessment (“HRIA”) or its Executive Summary, or the methodology that was employed to reach such findings, conclusions, opinions or viewpoints. Likewise, while Meta in its response identifies steps it has taken, or plans to take, which may correlate to points raised or recommendations in the HRIA, these also cannot be deemed an admission, agreement with, or acceptance of any findings, conclusions, opinions or viewpoints in the HRIA or its Executive Summary. The purpose of Meta’s response is to focus on recommendations related to the HRIA, rather than to opine on or contest any specific observations, findings or factual conclusions contained in the HRIA. Any steps taken by Meta in response to the HRIA’s recommendations that may be global in nature, or have impacts on other markets besides the Philippines, should not be interpreted to suggest that such steps are desired or potentially appropriate outside of the Philippines.

² [“Philippines: Events of 2020” Human Rights Watch World Report 2021](#).

³ Ibid; UN A/HRC/44/22 [Situation of Human Rights in the Philippines](#), June 2020.

outside the country. Our family of apps plays an important role in enabling Filipinos to connect with their families and friends, and to exercise their internationally-defined human rights. These include security of person, freedom of expression, freedom of association, the right to an adequate standard of living, the right to information, and myriad economic, social, and cultural rights.

We acknowledge the harms, pain and suffering caused by bullying, harassment and other abuses discussed in the human rights impact assessment (HRIA), which is why we have taken concrete steps to mitigate those abuses. We'll continue to do so to help keep people safe.

This HRIA is an important step forward for Meta and our work in the Philippines. By detailing the ways in which we have sought and plan to implement Article One's recommendations — and where and why we have not — we have attempted to demonstrate accountability for our human rights impact. It's also important to know that our work to address safety, security and human rights is not static. We are constantly working to evolve our products, policies and processes to create better outcomes.

Important as these HRIAs are, we see them as the start of a much bigger process of engagement with our users and those whose rights are impacted by our platform. We hope the insight and details we have shared will enable us to improve, iterate, and respect and support human rights more effectively in the future.

About this HRIA

Against this backdrop of severe human rights challenges, Meta commissioned an independent HRIA into the impacts of its products in the Philippines. We commissioned this independent HRIA to identify and mitigate human rights risks, and to inform our future strategy.

The HRIA was conducted by [Article One](#) in 2020 and 2021. Article One is a specialized human rights and ethics consulting firm. Article One's methodology was based on the [UN Guiding Principles on Business and Human Rights](#) (UNGPs), and included qualitative interviews with 32 stakeholders, including journalists, child protection specialists and civil society activists, as well as a survey of 2,000 Meta users in the Philippines.

We requested the HRIA include quantitative and not just qualitative work to ensure the experiences and feedback of a representative sample of rights holders, including vulnerable rights holders, were directly incorporated in our due diligence. Our quantitative work includes feedback from urban and rural users, journalists, users who identified as LGBTQ+, and other groups.

Given the ongoing challenges of the COVID-19 pandemic throughout 2020, in-country research and interviews were not possible. Article One's assessment thus relied on detailed discussions conducted via video conference, telephone-based surveys, and desk-based research. Article One also independently interviewed more than a dozen Meta staff and multiple teams working on Philippines issues.

We recognize COVID imposed methodological limitations, but chose to move ahead with due diligence nonetheless, given the importance of preventing and mitigating salient human rights risks. The project was updated in mid 2021 to include an analysis of the Anti-Terrorism Law of July 2020 and the ongoing impacts of the COVID-19 pandemic. We hope to deepen in-person human rights engagements in the Philippines if and when it becomes possible to do so safely.

We note that, during the period of HRIA research, local Meta staff members faced security challenges related to offline threats, as well as online harassment campaigns.

We are sharing the HRIA's executive summary, along with this response, to help local and international stakeholders better contextualize Meta's impacts and investments in the Philippines. This disclosure is also part of Meta's [broader commitment](#) to meaningful transparency about our human rights due diligence and our product and policy integrity work.

Why Have an HRIA Response?

There's no standard format for disclosure. To date, Meta has chosen the unusually transparent step of disclosing actual HRIA executive summaries and recommendations. They are written independently by HRIA authors, and accurately represent the issues and exact recommendations made in the full HRIA.

This Meta response is intended as a summary of the process, the findings and a guide to what we have done and will do to follow up. By sharing these documents we're seeking to fulfil the expectations of Principle 21 of the UN Guiding Principles on Business and Human Rights (UNGPs) while minimizing risks to stakeholders and company personnel.

We are publishing this HRIA and response as standalone documents in the interest of a timely disclosure. Going forward, in line with the commitments in our recently-announced [corporate Human Rights Policy](#), we will generally share actions and insights from our due diligence as part of our annual human rights report.

Good human rights due diligence is not just a compliance exercise. We are actively seeking to learn from this HRIA to inform our work in the Philippines, to mitigate risks in other settings and to serve users around the world. This HRIA, along with accompanying processes, has helped change Meta policies, operations and products. The impact has been timely and real.

Acknowledgements

We are deeply grateful to the dozens of Filipino and international human rights defenders, civil society organizations and others who provided input to this assessment. You have enabled us to implement the very important mitigations described in this response and chart a course forward on the critical issues this HRIA identified, benefiting people in the Philippines and around the world.

Findings

The HRIA found the Meta family of apps — in addition to being widely used and very popular — have had positive human rights impacts in the Philippines, as well as adverse impacts.

Article One’s assessment, based on qualitative and statistically significant quantitative research focused on a variety of at-risk rights holders, found that using the Facebook app:

- Had expanded opportunities for human rights defenders, journalists, government critics, and marginalized communities to express themselves, share information, and build communities
- Was critical for communication with family and friends in the Philippines and abroad, especially during the COVID-19 pandemic
- Was likewise essential for monitoring and defending human rights during COVID
- Was an important platform for accessing economic opportunities. This was especially the case for women and other groups who historically faced greater restrictions in accessing formal employment; 84% of female survey respondents indicated Facebook was somewhat or very useful for work or business, compared to 78% of men
- Provided a tool for remote teacher training, thereby supporting childrens’ access to education during COVID

These impacts promoted the realization of rights to work and an adequate standard of living. Survey results indicated users had also experienced benefits related to freedom of expression and association; the right to participate in government; rights of access to information and education; and freedom from discrimination.

The HRIA also noted Facebook was associated with a range of salient human rights risks. These include concerns about:

- Misuse of the platform to spread misinformation and disinformation, which may contribute to adverse impacts for a range of rights, including the rights to information, and to security of person
- Online harassment of a variety of groups including: journalists (particularly female journalists), individuals focused on government accountability, political dissidents, LGBTQ+ users; and ethnic Chinese users
- Incitement of violence, including involvement of government officials, against human rights defenders and alleged drug users and dealers. This included the practice of “red-tagging” of human rights defenders, which placed targeted individuals at significantly heightened risk of unlawful killing.⁴
- Government use of the platform to surveil human rights defenders and other users. The HRIA noted that government requests for user data were reviewed for legal sufficiency and accordance with Facebook’s global policies, including the [Global Network Initiative \(GNI\) Principles](#).
- Online sexual exploitation, which in the Philippines takes multiple forms, and includes predators' use of multiple platforms to groom or abuse children, including Meta products. In addition, women in the Philippines continued to experience cases of sharing of non-consensual intimate images (NCII).
- Trafficking, which remains a challenge in the Philippines, and is increasingly facilitated online. Trafficking risks included commercial adoption, labor trafficking, and organ sales.
- Highly localized extremist organizing, despite strong policies against such activity

⁴ <https://www.hrw.org/news/2021/02/10/philippine-general-should-answer-red-tagging#>

Importantly, despite instances of red-tagging, the HRIA **did not** find misuse of Facebook’s platform to plan or organize extrajudicial killings against alleged drug users or dealers, though stakeholders were concerned online discourse was used to dox, harass, or threaten alleged drug users or dealers, and may have contributed to a dehumanizing tone of conversations about drug use on Meta platforms.

The HRIA noted there was guidance from human rights standards and related legislation to help inform how the company should manage some salient risks, but not all. Official guidance related to balancing of rights related to misinformation and disinformation was less clear. It is important to be honest about this challenging gap, and Meta notes the [work of](#) the UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, among others, in developing related standards.

In addition, stakeholder feedback indicated Meta had improved its proactive risk management practices since 2018, but suspicion of the platform from civil society stakeholders — including as a result of the bullying and harassment of journalists and human rights defenders — remained high, even while Philippines users were avid users of the platform.

Meta Response to Findings and Recommendations

We greatly appreciate the work and insights of this HRIA. The section below is focused on HRIA recommendations, and our responses. But it’s also important to note that the company has taken multiple steps since 2017 to improve our Community Standards and the way they’re enforced in the Philippines. We’ve also undertaken extensive efforts aimed at improving digital citizenship, economic empowerment, support for journalists, and other key challenges. Please refer to the Appendix for a detailed overview of these efforts.

HRIA Recommendations

Overview

The HRIA made 40 recommendations. At the time of writing, Meta had committed to implement 24 recommendations, partly implement 7 recommendations and was assessing the feasibility of another 9.

It is also important to be transparent, and note we may not implement all the recommendations we're assessing feasibility for. We will provide an update in our future annual human rights reporting.

For ease of reading, we have categorized our response to HRIA recommendations to match the order in which they appear in the HRIA executive summary. We are categorizing our responses as follows:

Implementing: We have implemented or are implementing steps that are consistent with, or have otherwise satisfied, the recommendation.

Implementing in part: We have implemented or are implementing steps that encompass, or have satisfied, certain of the recommended actions.

Assessing feasibility: We are assessing the feasibility and impact of the recommendation and will provide further updates in the future.

No further action: We will not implement the recommendation, either due to a lack of feasibility or disagreement about how to reach the desired outcome.

As noted above, for a detailed summary of our Philippines program and policy efforts from 2017 to the present, please see the Appendix.

1. Misinformation and Disinformation

- a) Develop a risk-mitigation plan for the 2022 Presidential elections, focusing especially on the organized use of disinformation campaigns.

Implementing

Meta has made it a high priority to invest in risk mitigation and preparation efforts for the 2022 Philippine presidential election. This work has been underway since 2019. Our efforts will build from our experience in other elections, such as the 2020 elections in the US and Myanmar, as well as our work on [social media and conflict](#).

We have already initiated a cross-functional working group dedicated to election preparations. We have also begun briefing various stakeholders on digital advocacy, community standards, online safety and political advertising. In addition, we are also strengthening relevant civil society, fact-checking and media partnerships.

We have also expanded our support to Philippines fact checking groups. In 2021, Internews, supported by Meta, launched the Philippine Fact-Checker Incubator Program to support six new fact-checking organizations ahead of the 2022 Philippine general elections. Likewise, we began labelling media outlets determined to be state controlled, including media entities in the Philippines, in 2020.

Meta also launched a Global Accelerator [amongst our fact-checking community](#). It brings fact-checking organizations together to innovate, learn and collaborate on new strategies to improve their business — on and off Facebook.

- b) Expand efforts to track and remove coordinated disinformation campaigns that rely on real profiles.

Implementing

We launched a [policy](#) against certain mass harassment/brigading in August 2021 — strongly inspired by the experiences documented in the HRIA — after consultation with key human rights stakeholders in the Philippines and around the world.

We have also strengthened our ability to take action against coordinated social harms by authentic users.

- c) Implement the recommendations of Facebook’s Civil Rights Audit related to voter interference in the Philippines.

Implementing

In response to feedback from our global stakeholders and the important work of the Civil Rights Audit in the US, we have significantly expanded our global policies addressing voter interference.

Our Community Standards now [prohibit](#) offers to buy or sell votes; content advocating or providing instructions for voter fraud; misrepresentation of dates, locations, times, and methods for voting or voter registration; misrepresentation of who can vote and requirements for voting; and content implying that people will be infected with COVID-19 (or other diseases) if they vote. These policies apply globally, including in the Philippines.

We’ve also [introduced](#) measures to reduce the risk that potentially harmful Groups are recommended to people, and we let people know when they’re about to join a Group that has Community Standards violations. We limit invitation notifications for these groups, so people are less likely to join.

For full details on Meta’s follow up to the Civil Rights Audit, [see here](#).

- d) Scale Facebook’s commitment to the [EU Code of Practice on Disinformation](#) to include the Philippines.

Assessing feasibility.

We support multi-stakeholder, collaborative and iterative approaches that self-regulatory initiatives — like the EU Code of Practice on Disinformation and the Australian Code of Practice for Disinformation and Misinformation — facilitate. Such approaches strike a reasonable balance between freedom of expression, safety and security, and provide a forum for ongoing dialogue between stakeholders to find answers and solutions as the digital ecosystem evolves.

Meta's approach to addressing misinformation and disinformation is global. It includes our partnerships with third-party fact checkers, as well as our work to support and elevate sources of quality information, to address misrepresentation and the coordinated use of fake accounts, and to prohibit misinformation in advertisements in accordance with our Advertising Policies and Community Standards. We are actively exploring ways to provide additional reporting on misinformation and disinformation.

- e) Expand Ads Library to the Philippines to ensure transparency on political advertisements.

Implementing

We have made the [Facebook Ad Library](#) available in the Philippines to capture all live ads and all ads known to be related to politics and elections.

As part of our ads transparency efforts, we require people who want to run an ad targeting the Philippines related to elections or politics to complete an identity confirmation process and demonstrate they are located within the country. These advertisers are then required to place a disclaimer on their ads about elections or politics. These ads will be archived in the Ad Library for 7 years where anyone can find out more information on the advertiser. We are intending to require authorization for social issue ads from early 2022.

- f) Update policies to forbid political ads that include disinformation on Meta's platforms.⁵ This could include having political ads fact-checked by IFCN certified fact checkers before they are approved for dissemination on Meta's platform.

Implementing in part

We require advertisers to follow our [Advertising Policies](#), *in addition to* our [Community Standards](#).

⁵ The Philippines HRIA defines as disinformation "content that gives the appearance of being factual and news-like but actually relies on deliberate falsehoods to mislead and/or promote a political agenda. Objective falsehood is a key component of the term." NYU Stern (2017) "Harmful Content: The Role of Internet Platform Companies in Fighting Terrorist Incitement and Politically Motivated Disinformation."

We [prohibit](#) ads containing claims that have been debunked by third-party fact checkers, unless they consist of direct and unedited speech by politicians. (It is important to note that our policy prohibiting misinformation that leads to imminent physical harm **is still applicable** to politicians).

We do not allow third party fact checking on the direct and unedited speech of politicians because we believe that by censoring the speech of politicians, we'd leave people less informed about what their candidates or elected officials say, and leave politicians less accountable for their words.

There will be some instances, however, where a rating from our fact-checking partners will affect politicians. Specifically, when a politician shares a specific piece of content — e.g. a link to an article, video or photo created by someone else that has been previously debunked on Facebook — we will reject its inclusion in ads. This is different from a politician's own claim or statement.

Our policies also prohibit [voter interference](#) in elections, including in ads. Examples of prohibited content include misrepresentations of: who can vote; dates, times, and locations of elections; and qualifications or materials required for voting.

We also prohibit ads that contain certain COVID-19 and vaccine claims that have been debunked by global health organizations such as the CDC and World Health Organization, or [claims](#) from any advertisers that discourage people from vaccination or advocate against vaccines.

Given the volume of advertising, and the difficulty of assessing often opinion-based claims included in political and social issues advertising, it is not a viable or a desirable use of inherently limited fact-checking resources to have each political ad systematically assessed by a third-party fact checker for accuracy prior to approval.

Such an approach may also limit the journalistic independence of our fact-checking partners, who are currently free to choose the claims they wish to assess. In practice, it may also create an unintended bias against advertisers speaking about certain types of issues or based in certain areas where fact-checking capacity is relatively limited.

- g) Include state actors in the existing dangerous individuals and organizations policy.

Assessing feasibility

We will consider exploring related policy options as part of our global policy development process.

- h) Assess opportunities to expand fact-checking of political leaders beyond misinformation that may result in imminent harm.

Implementing in part

We assess such opportunities on a continuous basis. For example, in 2020-21, we expanded prohibitions against certain COVID-19, vaccine and electoral content, as noted in the paragraphs above. We also pioneered new methods of enforcing misinformation policies and claim checking at scale in US, Myanmar and other elections. Our teams preparing for the Philippines election will build from this work.

We are not considering additional community standards changes in this area at the time of writing. Our approach to fact-checking is grounded in Meta's fundamental belief in free expression, respect for the democratic process and the belief that political speech is the most scrutinized speech there is. Just as critically, by limiting political speech we would leave people less informed about what their elected officials are saying — and leave politicians less accountable for their words.

- i) Provide financial support to cover the cost of security and mental health support for fact checkers at partner organizations.

Assessing feasibility

Meta has [contributed more than \\$84 million](#) to programs supporting our fact-checking efforts since 2016, including [support for debunking COVID-19 content](#), a [mentorship program](#), a [Climate Misinformation Grant](#) and an [Accelerator](#) to support fact-checkers' businesses. We are always looking for additional ways to support fact-checking partners based on their needs, and will consider options for mental health support as we plan for the future. Fact checking partners who meet the relevant criteria also will be eligible to apply for our broader human rights defender and journalism safety initiatives.

- j) Expand existing misinformation and harm policy to include further categories of health misinformation.

Implementing

Since the primary research for this HRIA was conducted, we have taken important steps to expand and refine our policies around health misinformation.

In December 2020, we began [removing misinformation about COVID-19 vaccines](#), and [sending notifications](#) directly to anyone who has interacted with health misinformation that was removed for violating our Community Standards. In February 2021, we also expanded the list of COVID-19 misinformation claims we would take action on, in line with guidance from global health authorities. In April 2021, we started removing Pages, Groups and Instagram accounts that repeatedly violate our COVID-19 and vaccine policies.

Our Community Standards are a living document, meaning that we are constantly evolving our policies to keep pace with changing online behaviours. Our teams will actively explore additional policy changes related to health misinformation, taking into account important freedom of expression, access to information, proportionality considerations, and guidance from the Oversight Board.

- k) Expand investment in digital literacy training and provide funding for journalism schools in the Philippines.

Implementing

First, Meta has launched an expanded version of [We Think Digital](#)/Digital Tayo in the Philippines. This includes interactive content on safety and privacy, and new modules on news verification, community building and civic engagement. Digital Tayo continues to grow its partner roster adding more government agencies, youth organizations, non-government organizations and community-based groups.

Second, there is also ongoing work on integration of digital literacy into the learning curriculum of pre-service tertiary and basic education teachers. We are also developing a

Facebook and Messenger School Guide for students, teachers and home-based learning facilitators that will provide key information on our Terms of Service and Community Standards, coupled with best practices and safety and privacy tips to aid online learning, in partnership with the Philippine Department of Education and Stairway Foundation.

Third, we are also supporting a project with [EYT](#), a community-oriented independent media outfit, to train local/campus journalists on digital literacy, visual storytelling, and fact-checking in preparation for the 2022 general elections.

The [Meta Journalism Project](#) has also partnered with Reuters to launch a free online digital journalism foundations training program. This includes modules on digital news gathering, news verification and reporting, publishing on social media, and also wellness and resilience training while reporting. More information about this partnership can be found [here](#).

In June 2021, we published the Community Standards in Filipino, enabling Filipino users to read the Community Standards in their native language.

2. Online Harassment

- a) Enable and empower users to take control of their security and privacy by expanding and increasing awareness of platform-based reporting functions, including through more visible product tools.

Implementing

We have an extensive suite of product tools designed to empower people to address the potentially negative impacts of unwanted contact. These vary across different apps, but include the ability to unfriend, unlike, ignore, block, restrict and report. See our [abuse resources](#) for more information.

Some examples include: on Instagram you can restrict comments from certain people on your posts; the comments will only be visible to that person, or you can approve them to be visible also to others. On Messenger you can ignore a conversation, turn off

notifications for a conversation, delete a conversation or block the person from sending you messages. On Facebook, you can add a friend to a restricted list, and they won't see posts you share only with friends.

We are committed to continually improving these tools. In March 2021 we also introduced "[Choose Who Can Comment](#)", a product tool that allows Facebook app users to limit commenting on profiles and pages to select audiences on their public posts. In August 2021, Instagram announced Limits, [a set of features](#) to help users more easily prevent unwanted comments and DMs. These features are particularly relevant for creators and public figures who experience or may anticipate a rush of abusive comments or DMs.

We have matched these product-related steps with policy steps, outlined in section d below.

- b) In line with guidance from [Amnesty International](#), "offer personalized information and advice based on personal activity on the platform" (e.g., share guidance on privacy and security settings when users make a report of violence and abuse). This guidance should not contain private information, but rather direct users to resources and tools that can help mitigate future risks.

Implementing

We agree that providing tailored guidance within our apps is one of the best ways to help people understand our policies and settings. In line with this recommendation, we provide contextually-aware information on privacy, safety, and security features throughout the Facebook app.

For example, when reporting a post for harassment, we also share one-tap options to block or unfollow the user or page that made the post, as well as details on the relevant portions of our Community Standards. When we consider a user may be particularly vulnerable to security risks, we provide notifications encouraging them to turn on two-factor authentication and other advanced protections.

We develop technology to fight behavior that threatens to keep women and vulnerable users offline. We offer a number of tools to help people control their experience on

Facebook, such as ignoring unwanted messages and blocking other people without them being notified. Victims and their allies can also report violating behavior. Other tools available in the Philippines include:

- Comment controls: You [can control who can comment on your public posts](#) by choosing from a menu of options ranging from anyone who can see the post to only the people and Pages you tag.
- Profanity filter on Facebook: You can block different degrees of profanity from appearing on your Page. We determine what to block by using the most commonly reported words and phrases marked offensive by the community.

In consultation with stakeholders, we'll continue to explore additional opportunities to provide such information, including to human rights defenders in the Philippines.

- c) Strengthen policy measures against harassment, particularly the organization, incitement, or coordination of harassment, against users.

Implementing

We updated our Community Standards to prevent calls for coordinating harassment in 2020. We also launched a [mass harassment/anti-brigading policy](#) in August 2021, a significant step in addressing this complex challenge.

To reduce the reach of those looking to use our platform for harmful purposes, we now also [make it harder](#) to find and join Groups that violate our policies. In addition, when Group members have repeated violations, we now block them from being able to post or comment for a period of time in any group, and prevent them from inviting others to any groups, and creating new groups.

- d) Adopt a “public interest” consideration, in place of a “public figures” consideration and provide policy exceptions and product protections for at-risk public figures, including human rights defenders (HRDs) and journalists.

Implementing

We launched multiple [changes to our bullying and harassment policy](#) in 2020-2021.

These included:

- Clarifying that most human rights defenders and journalists are “limited public figures” with stricter policies prohibiting bullying and harassment than public figures;
- stronger protections against gender-based harassment and sexualization, including of public figures; and
- a new policy prohibiting certain mass harassment (“brigading”), including against human rights defenders as defined by the [United Nations](#).

We also launched highly relevant product protections, such as Facebook Protect and the News Page Index and Journalist Registration.

- [Facebook Protect](#) is an in-app security measure that provides eligible users more security as they login and reduce the chance of unauthorized access to their accounts.
- The [News Page Index](#) helps us identify Facebook Pages that primarily publish news content. Publishers that are included [may be eligible](#) for products, policies and programs that benefit news publishers. To be included, they must abide by guidelines we developed with news organizations, academics, and industry groups. The News Page Index was launched in Filipino in November 2021.
- [Journalist Registration](#) offers account security protections that help further protect against online abuse. It is the beginning of Meta's efforts to better identify and support journalists on our platform.

We have also funded the [Human Rights Defender Fund and Journalist Safety Initiative](#), a pilot fund to support at-risk journalists and human rights defenders in the Asia-Pacific region.

- e) Consider partnering with UN agencies and civil society organizations to develop a program to raise awareness of HRDs.

Implementing

Meta conducted briefings for Human Rights defenders on policies around data sharing and Facebook's Community Standards, as well as training on using Facebook for digital advocacy.

In August 2021, we launched the Online Security and Resiliency Training for human rights defenders with [The Asia Foundation](#). The series aims to equip HRDs with knowledge on how to stay safe and be more resilient online. It includes digital security, mental wellness, and community engagement. As of October 2021, we have trained 61 organizations (961 individuals trained). The series will continue in 2022.

We financially support several global response/advocacy mechanisms for HRDs, and will actively explore global campaign options. There is also additional information on HRDs in section 4 below.

3. Incitement to Violence

- a) Formally adopt the [Rabat Plan](#) "six-part test" into Facebook's Community Standards implementation, specifically in relation to Violence and Incitement.

Implementing in part

Meta is actively exploring ways to incorporate learnings from the Rabat threshold test into our policies. While the nuance of the test makes it operationally infeasible to directly apply at scale, we have adapted them into two active content policy tools, including an escalation-based framework on hate speech attacking concepts, not people (adopted 2021), as well as an updated framework accounting for civil and human rights principles to evaluate discussions of state use of force by governments against civil and human rights principles.

The Human Rights Team also uses the test as a basis to advise on future policy development.

In 2020, Meta worked with the UN Office of the High Commissioner for Human Rights to have the test translated and [made available in 26 additional languages](#).

- b) Explore opportunities to share platform insights (not specific user or content data) with accountability or preventive bodies via [Data for Good](#) or other mechanisms.

Implementing

We note this HRIA response is one such method, as is our voluntary lawful disclosure to the Independent Investigative Mechanism for Myanmar, active from late 2019.

Our commitment to report human rights insights and actions, contained in our March 2021 [corporate human rights policy](#), is a potentially significant opportunity to do so.

In addition, our Human Rights, Stakeholder Engagement and Public Policy teams routinely informally discuss insights on broad content trends — including those related to social media and conflict and incitement to violence — with partners in the human rights community, academia, and international organisations. In doing so, we seek to inform policy development and enforcement, support rapid response, develop best practices and support accountability for human rights violations.

Our [Transparency Center](#), launched in 2021, is a significant resource. Our [Community Standards Enforcement Report](#), updated quarterly, includes information on the prevalence of various types of harmful content over time and volumes of such content removed. A range of public Facebook content is made available to academic, human rights, and other stakeholders via [Social Science One](#) and [CrowdTangle](#).

We financially supported, and participated in, related human rights multi-stakeholder conversations related to documentation and disclosure in 2020 - 21.

We'll actively explore additional ways to share additional insights and data with our partners, consistent with our privacy and other legal obligations.

4. Surveillance

- a) Collaborate with civil society to advocate for the passing of the Human Rights Defenders Bill.

Assessing feasibility

We are collaborating with civil society to encourage protections for HRDs, but passage of this bill is stalled.

- b) Develop a standalone HRD policy and implement recommendations developed as part of a Meta-commissioned HRD research assessment.

Implementing

We commissioned HRD-specific due diligence in 2020. As a result, our corporate [human rights policy](#) (launched March 2021) specifically references the needs and risks of human rights defenders. This will provide a strong policy basis for future work.

In addition, we have:

- significantly revised our [bullying and harassment policies](#) (details in section 2 above)
- launched new restrictions on calling for or engaging in certain [mass harassment](#)
- created an HRD-specific cross-functional working group
- integrated defenders into our due diligence systems focused on vulnerable/targeted groups, improving our ability to identify and launch relevant product solutions, such as [locked profile](#) in Afghanistan and Myanmar;
- engaged more systematically with relevant UN special procedures
- launched the Human Rights Defenders' and Journalists' Fund. This is a two year pilot fund in APAC, which, due to high digital penetration and other factors, is a region with concentrated risk.
- launched the Online Security and Resiliency Training for Human Rights Defenders with The Asia Foundation in 2021 to equip HRDs with knowledge on how to stay safe and be more resilient online

- c) Co-develop technical tools with HRDs that work to protect their rights while on Meta platforms.

Assessing feasibility

The human rights team is actively seeking to develop a more consistent set of co-design opportunities.

5. Online Sexual Exploitation

- a) Use leverage with the government of the Philippines to expand protections against online sexual exploitation of children.

Implementing

In line with the UN Guiding Principles on Business and Human Rights, we have actively sought to support outreach, understanding and consultation around the drafting of various online child safety bills which aim to address the issue of online sexual exploitation of children and further define the duties of platforms and tech companies in addressing related issues. This work has entailed, among other engagements, attendance at various Technical Working Group meetings.

We also established active, ongoing coordination mechanisms with the [Inter-Agency Council Against Trafficking](#), Department of Justice, Department of ICT, and the Department of Education as well as NGO partners [UNICEF](#), [Plan International](#), [Stairway Foundation](#) and [Child Rights Network](#). This allows us to quickly respond to reports and continuously gather local insights and trends that will help us improve our policies, tools and resources to better protect children online.

We review legal requests carefully and respond in accordance with applicable law. We train and work extensively with law enforcement, including the Philippine National Police's (PNP) [Anti-Cybercrime Group](#), the [Women and Children's Protection Center](#) and the National Bureau of Investigation, including to provide training on how to efficiently submit legal requests. We combat child online exploitation including working with the Philippine National Police's (PNP) Anti-Cybercrime Group, the Women and Children's Protection Center and the National Bureau of Investigation, along with their foreign law

enforcement and NGO partners such as Homeland Security Investigations, the [International Justice Mission](#) and others.

We report apparent instances of child sexual exploitation appearing on our site from anywhere in the world to the [National Center for Missing and Exploited Children](#) (NCMEC), which then refers cases out to global law enforcement agencies: More details can be found [here](#).

In August 2019, we provided funds and engineering expertise to create the updated NCMEC Case Management Tool. The tool is now available to law enforcement around the world — including the Philippine Department of Justice — making it easier for them to access, prioritize and respond to reports of online child sexual exploitation.

Aside from the proactive work we do, we respond to formal legal compulsory requests from Philippine law enforcement agencies for user data requests related to child exploitation cases. We understand, however, that Philippine legal restrictions require local law enforcement to first obtain a search warrant from a court to make a compulsory request that would qualify for disclosure under Facebook’s Data Policy and US law. Given that law enforcement must first initiate and complete court proceedings to be in a position to make a legal compulsory request for data, there are a correspondingly low number of requests for Facebook to process in this area.

In particular, if we have reason to believe a child is in immediate/imminent danger, we may proactively refer a case to local law enforcement to help make sure the child is immediately safeguarded.

We have also provided multiple training sessions to law enforcement agencies over the years in the Philippines organized by the [International Centre for Missing and Exploited Children](#) and NCMEC.

We have a dedicated global team that supports law enforcement agencies in accordance with our commitments as a member of the [Global Network Initiative](#). The team operates 24/7, 365 days a year, from 15 sites around the world.

- b) Commission a child rights impact assessment of Messenger Kids.

Assessing feasibility

Messenger Kids was developed specifically to meet child safety needs. Product development included extensive consultation with a team of global advisors and independent experts in the fields of online safety, child development, and media from organizations such as the [Family Online Safety Institute](#), [Center on Media and Child Health](#), [MediaSmarts](#), the [Yale Center for Emotional Intelligence](#) and the [Fred Rogers Center](#). In the Philippines, we consulted the Stairway Foundation, the Child Rights Network, UNICEF and Plan International.

We will actively consider this recommendation while also focusing on future child-related product development due diligence.

- c) Evolve default privacy settings for minors (anyone under the age of 18) (multiple details).

Implementing in part

To protect minors, we have already put special safeguards in place, and will continue to do so.

We:

- require everyone to be at least 13 years old before they can create an account
- set their accounts to share with ‘friends only’ by default
- take steps to prevent default audience options for posts from including ‘public’
- [restrict](#) the visibility of minors' profiles in public search, prevent public sharing of content (including location, contact information, school or birthday and friend lists)
- block adult-only features (e.g., Facebook Dating)
- remind minors that they should only accept friend requests from people they know, and we do not allow unconnected adults to message minors
- use photo-matching technologies that help us detect, remove, and report the sharing of images and videos that exploit children
 - These photo-matching technologies create a unique digital signature of an image or video (known as a “hash”) which is then compared against a

database containing signatures (hashes) of previously identified illegal images to find copies of the same image. We use these technologies across our public surfaces, as well as on unencrypted information available to us on our private-messaging services, including user reports.

- We also run these technologies on links from other internet sites shared on our apps and their associated content to detect known child exploitation elsewhere on the internet. All violating content found is reported to the National Center for Missing and Exploited Children, which works with law enforcement around the world.
- use artificial intelligence and machine learning to proactively detect child nudity and previously unknown and new child-exploitative content, as well as inappropriate interactions with children, sometimes referred to as “grooming”
- allow anyone — whether they use our services or not — to report suspected underage people on Instagram and Facebook. We have dedicated channels where we review these accounts based on these or other reports.

In addition:

- minors and adults can also turn on Timeline Review to review posts they're tagged in before the posts appear on their timeline. If a minor or anyone else is tagged in something they don't like, they can remove the tag or ask the person to take the post down.
 - to protect children’s image privacy rights, on Facebook, a parent or legal guardian can request that we remove an image of their child under the age of 13 by filling [out this form](#). A young person between the ages of 13 and 18 can request we remove their image using this form.
- d) Consider requiring children to review a child-friendly version of the Community Standards when signing up for a Facebook account and develop child-friendly safeguarding tools, including guidance on privacy settings, data collection and reporting of grooming.

Implementing in part

We are exploring the feasibility of child-friendly versions of the Community Standards and/or related rules. In addition, we are working with the Department of Education and

the Stairway Foundation to develop a Facebook Safety Guide for Schools to support school personnel, home-based learning facilitators, and learners in safely navigating Facebook as a communication platform to share important school updates, information, and announcements.

We have already developed child-friendly safeguarding tools, and will continue to do so.

For example: in May 2020 [we introduced safety prompts](#) in Messenger, which are surfaced to minors before interacting with an adult who has been exhibiting potentially suspicious behavior. For example, if an adult is sending a large amount of friend or message requests to people under 18, this tool alerts the recipients to be cautious when interacting with an adult they may not know, educates them on harassment and empowers them to ignore, block or report someone when something doesn't seem right.

We [launched a Youth Portal in our Safety Center in 2018](#), a central place for teens that includes education on staying safe and first person accounts from teens around the world about how they are using technology in new and creative ways. It features tips on things like security, reporting content, and deciding who can see what they share and some guidelines for how teens can get the most out of the internet

- e) Consider launching a child online safety month or comparable effort to raise awareness of risks and tools through Meta's platforms.

Implementing

As part of 2021 Safer Internet Day, we ran an [online safety education campaign](#) on Facebook and Instagram from February to March 2021 that has been viewed over 18 million times. We worked in partnership with the Child Rights Network, Stairway Foundation, Department of ICT and [Called To Rescue](#) to educate users on Facebook privacy and security tips. Meta annually participates in Safer Internet Day activities.

We also continue to improve our in-app education about tools and resources to prevent online child exploitation.

Building on our current participation in other national child safety campaigns, at the time of writing we were running the [#ReportItDontShareIt campaign](#) in partnership with the Inter-Agency Council Against Trafficking (IACAT), Child Rights Network and the Stairway Foundation. The campaign raises awareness of the risks of sharing child sexual exploitation material, promotes a shift in local attitudes, and encourages people to report such content.

6. Trafficking

- a) Take efforts to exclude from Facebook search results content related to the sale of human organs.

Implementing

We have updated 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](#).

- b) Update human exploitation policy to prohibit all sale of children, even in cases where the sale may be legally permissible.

Implementing

We have extensive language prohibiting multiple forms of human exploitation in our existing policies. This has been [updated](#) to unambiguously prohibit all sales of children.

- c) Conduct on-platform awareness-raising training related to the risks of illegal recruitment and how to detect potentially exploitative content.

Implementing

In partnership with [Stop The Traffik](#), we launched a [Labor Trafficking/Domestic Servitude Awareness Campaign](#) in March 2021 to provide information to potential domestic workers and low-skilled laborers about their legal rights, local recruitment guidelines for overseas job applications and available helplines that would help them avoid illegal recruitment and abuse.

At the time of writing, we were building out an organ trafficking potential victim awareness project to help educate potential victims of organ trafficking about risks and provide resources, with a target launch in early 2022.

Our content review team and proactive detection technology uses English and Filipino keywords to identify and remove content that may lead to human exploitation, including specifically domestic servitude.

We developed these keywords based on consultations with expert organisations in the Philippines, including the Inter-Agency Council Against Trafficking, the Department of Health, the National Kidney and Transplant Institute, the [Overseas Workers Welfare Administration](#), the [Blas Ople Policy Center](#), the Department of Foreign Affairs and the [Philippine Overseas Employment Administration](#).

7. Extremist Activity

- a) Further invest in Moro language capabilities, particularly highlighting key terms in Maranao, Maguindanaoan, and Tausug, to allow for more accurate, proactive review of potential terrorist content.

Assessing feasibility

At the time of writing, Meta has content moderation capacity in Cebuano and Filipino. We'll evaluate how to prioritize additional capacity based on our criteria for global language prioritization and the UN Guiding Principles on Business and Human Rights.

In addition:

- We've launched [The Resiliency Initiative](#), a partnership between Meta and The Asia Foundation, which aims to promote tolerance, strengthen interfaith and inter-ethnic understanding, and counter violent extremism by helping to build resilient communities across APAC. The initiative delivers a range of capacity-building programs for civil society organizations to help develop localized positive narrative campaigns online that aim to counter intolerance and discrimination. In 2020 the program reached over 60 organisations across 10 countries in Asia (Indonesia, Thailand, Philippines, Malaysia, Pakistan, Sri Lanka,

Maldives, Bangladesh, Nepal and India). In 2021, an [online portal](#) was launched to make the information resources accessible to more communities and as of this writing, a new batch of organizations was lined up for training.

- We have implemented the Boost Your Business training program in the Bangsamoro Autonomous Region of Muslim Mindanao for digital readiness and opening online economic opportunities for the local community, with [Bayan Academy](#).

8. Corporate Accountability

- a) Develop a standalone human rights policy to guide Meta's approach to managing human rights

Implementing

[Meta's corporate human rights policy](#) was launched March 2021. It is also embedded in [Keep Building Better](#), the revised Meta code of conduct launched June 2021.

- b) Consider adding an independent board member responsible for human rights and/or establishing a human rights advisory board, similar to the Safety Advisory Board.

Assessing feasibility

Under our [corporate human rights policy](#), salient human rights issues are periodically reported to the company Board of Directors. Meta has benchmarked business practice related human rights stakeholder advisory and consultation mechanisms, and will seek to implement such a mechanism.

- c) Build on transitional justice work to contribute to broader recognition of human rights harms in countries with legacies of human rights abuses.

Implementing

Meta is actively incorporating transitional justice and principles into its policy, communications and human rights work. Transitional justice experiences are highly relevant to Meta and other social media companies because they often confront

systems-level challenges, and struggle with issues of scale, speed, and remedy.

- d) Continue efforts to align with the Santa Clara Principles.

Implementing in part

The principles give excellent guidance. We will continue to work to align with the principles, but note content volumes are so large we cannot commit to human review for all cases. Instead, we seek to prioritize based on severity of potential harm in a framework aligned to the UN Guiding Principles on Business and Human Rights.

- e) Formally integrate human rights expectations into the Responsible Innovation process, requiring that all “high risk products” undergo a human rights review.

Implementing in part

We have aligned Responsible Innovation (RI) principles with human rights guidance, but the process mentioned above has been deprecated. We will take further steps to integrate human rights guidance into RI product risk management processes.

- f) Assess which specific aspects of Meta’s business model are more or less likely to impact human rights adversely and develop mitigation measures based on the assessment results.

Assessing feasibility

Our human rights policy requires us to know our salient human rights risks. We are developing our salient risk framework across the company, which is an important precursor to this work.

9. Updated HRIA Recommendations (from September 2021)

- a) Initiate/expand information sharing on user engagement with COVID misinformation. This data can help independent researchers understand how misinformation is spread and the efficacy of mitigation tactics.

Implementing

We are convening two dedicated workshop series with leading misinformation, safety and expression experts to advise us on best practice research and thinking on effectiveness of misinformation enforcement, one of them focused on the Asia-Pacific region. We will consider opportunities to share related knowledge.

- b) Increase efforts to combat disinformation and misinformation in the Philippines, including:
- onboarding new third-party fact checkers to ensure sufficient resourcing to address disinformation in advance of the upcoming elections
 - applying Meta’s heightened penalties for public figures who violate Facebook’s Community Standards

Implementing

As part of our preparation for the Philippines election, we are already expanding our support and training for the local fact checking community, in line with the [IFCN Principles](#).

We have also created additional means to strengthen enforcement of misinformation policies (including, among others, voter interference and misinformation that leads to real world harm).

Our policy of heightened penalties against public figures who violate the Community Standards during periods of civil unrest is global, and thus will be applied when policy criteria are met. More [details here](#).

Locally, the following programs have also been put in place:

- We continue to implement our [Digital Tayo](#) training to youth and community-based groups, government agencies, and civic organizations nationwide. In 2021, we have nine new partners — building up our partner roster to 52 government and non-government organizations. Additionally, we have conducted training on Meta Community Standards to civic and elections-related organizations.

- We are supporting a project with FYT, a community-oriented independent media outfit, to train local and student journalists on digital literacy, visual storytelling and fact-checking in preparation for the 2022 general elections.
 - In partnership with local civic organizations, we will implement a series of information activities and campaigns on media literacy and digital discernment in the run up to the elections.
 - In 2021, we launched the [Youth Leaders Incubators Program](#) in partnership with [Out of The Box Media Literacy](#) and [Mano Amiga](#) to build a civic learning opportunity for Filipino youth that aspire to spark social impact in their communities and to encourage youth participation in the socio-political discourse in the lead up to the elections. Ten incubated youth organizations are taking part in an eight-month program that includes social media training, mentoring and technical assistance.
- c) Increase capacity to address red-tagging, including: building awareness among local civil society of Meta’s red-tagging policies, expanding the trusted partner program, and taking steps to review red-tagged content in a timely manner, consistent with the threat of life.

Implementing

We seek to continue our strong work prohibiting content that outs specific at risk groups, including red-tagging. The latter is an adversarial practice with constantly evolving dynamics, present across the entire media ecosystem. We will continue to adapt our handling practices as content dynamics evolve.

We are working to expand our Trusted Partner network prior to the 2022 national elections, and to expand their awareness of relevant policies.

We also continue to engage with the [Philippine Commission on Human Rights](#), as well as other local rights organizations. The Online Security and Resilience Training for Human Rights Defenders, a training program with The Asia Foundation, is also currently being implemented.

- d) Increase the capacity to mitigate the prevalence of online sexual exploitation and human trafficking on Meta platforms, including:

- investing greater resources in the prevention of child exploitation and sharing of child sexual abuse material (CSAM) in unencrypted spaces
- investigating new methods and technologies to address the spread of CSAM in encrypted spaces, while still respecting the right to privacy and other human rights
- expanding approaches to identify and limit dangerous engagements that could result in exploitation

Implementing

We lead the industry in our work to prevent and disrupt the sharing of CSAM on both our encrypted and unencrypted services. We will continue to invest deeply in this space, including by expanding our work on identifying and limiting potentially exploitative interactions with children. We are already deploying industry-leading methods and are seeking to pioneer new ones, to help address the potential spread of CSAM in encrypted spaces.

APPENDIX

Detailed Summary of Efforts Undertaken 2017 - 2021

We appreciate the observations in this HRIA, and have used its recommendations to strengthen our human rights risk management.

This section, which adds to the action steps outlined in the main body of this document, gives an overview of the work we've done since 2017 to support users in and from the Philippines. We hope it will be useful in showing the development, over time, of key Meta initiatives, policies and programs.

These efforts were focused on digital citizenship, elections preparedness, economic inclusion, support for journalists and human rights defenders, mental health resources, policy improvements, scaling safety and security, fighting misinformation and cybersecurity enforcement.

Importantly, from March 2021 this work is now underpinned by [Meta's Corporate Human Rights Policy](#).

I. Digital Citizenship

Building up people's ability to use online services like Facebook responsibly helps ensure that the internet can be both safe and open. Of course, it remains our responsibility to help people use our products and platforms in a way that benefits them and their communities.

Some of our digital citizenship efforts:

- Annually hosting the first [Philippines Digital Youth Summit](#), a three-day youth leadership that helps young people with skills to build online campaigns about digital literacy, empathy, and critical thinking.
- Partnering with [Globe's Digital Thumbprint](#) program to help students and teachers build news literacy.

- Working with civil society and government partners to launch the [Digital Tayo program](#) — a digital literacy and digital parenting training program for outbound overseas Filipino workers, students, parents, public school teachers, HRDs and LGBTQ+ communities. Since its launch, we've trained over 400,000 Filipinos in-person. Throughout the COVID-19 pandemic, Digital Tayo shifted to online training and resources, which have been viewed and accessed by over 5 million Filipinos across various social channels from March 2020 through April 2021. This includes LGBTQ+ organizations that we trained on digital literacy, and on modules on privacy and safety, responsibility and empathy. In addition, we also trained civil society organizations, including LGBTQ+ organizations, on how to use Facebook to develop their advocacy and information campaigns.
- Implementing digital advocacy training and digital safety training seminars for non-governmental organizations, including human rights defenders.
- Launching the [Meta Government Digital Transformation Summit](#), which trained governmental public information officers and relevant staff on maximizing use of Facebook for public service advisories, and on Meta's Community Standards (2019 and 2020).

II. Elections Preparedness

We provided support in the 2019 midterm elections, and are currently preparing for the presidential elections in 2022, building from our local and global experience.

- In 2019, we provided support for the local elections through voter registration and Election Day reminders, partnerships with the [Commission on Elections](#) and elections organisations, conducted briefings as well as page management and community standards training for the Commission on Elections, National Movement for Free Elections (NAMFREL), [Parish Pastoral Council for Responsible Voting](#) (PPCRV), and the [Legal Network for Truthful Elections](#) (LENTE).

We also hosted more than 60 members of the Filipino media for an election workshop to increase the quality of related content on-platform.

Our integrity work in removing coordinated inauthentic behavior networks, promoting digital literacy and tackling misinformation on the platform (please see respective sections) were part of the broader approach to the Philippines elections.

Building on this, we have doubled our efforts for the 2022 general elections. In addition to implementing and improving on what we have done in 2019, we have launched various programs that included new modules in Digital Tayo focusing on civic education, an expanded Digital Tayo and elections Trusted Partner network, as well as the Philippine Fact Checker Incubator program.

We are also implementing the Ads Transparency Library and have conducted briefings to civil society organizations, media and the Commission on Elections.

We'll make further announcements in the run up to 2022, which include local engagements, campaigns and partnerships.

III. Economic Inclusion

We're helping people use our platforms to build and sustain businesses and livelihoods. Meta and its platforms are powerful tools for small businesses who don't have access to the same kinds of advertising tools and resources that large businesses do. Small businesses use Facebook and Instagram to start their businesses, build brand awareness and promote new products. These apps also help them attract and build relationships with customers. By helping small businesses succeed, we can help entire communities improve their access to other resources, as well.

Our approach to date:

- We've launched Facebook Community Boost, a training program in digital marketing training for micro, small and medium-sized enterprises (MSMEs) and digital advocacy training for NGOs in three cities (Manila, Cebu, Dumaguete). This training included modules on digital literacy and the Meta Community Standards. In total, Community Boost, alongside #SheMeansBusiness trained a total of 18,112 MSMEs in 2019.

- We're continuously providing digital marketing and financial literacy upskilling training for small businesses in the Philippines in partnership with Bayaan Academy, [Connected Women](#) and selected government agencies. Together, we've launched [She Means Business](#), a digital marketing training program aimed at women and [Boost with Facebook](#), a similar program designed for small businesses. We have been running these programs since 2017.

As the COVID-19 pandemic spread, the importance of our ongoing efforts of helping small businesses increased as well. And thus, Boost with Facebook (formerly Boost your Business) was updated to include modules on business resiliency, financial planning and mental health. [She Means Business](#) was updated to include financial education, resiliency, and access to capital for women (promoting inclusion and economic opportunities). Despite the limitations put in place by the pandemic, we were able to train over 8,000 local MSMEs in 2020.

IV. Mental Health Resources

Over the past few years, we've undertaken a series of efforts to improve access to mental health resources, particularly for the LGBTQ+ community in the Philippines. We partnered with the tech-worker advocacy group [Pride@Tech](#) — as well as other tech companies and the mental health tech startup [MindNation](#) — to offer free psychological consultations, reaching over 2 million people who identify as LGBTQ+. In October 2020, we held a community chat streamed on Facebook Live, together with [#MentalHealthPH](#), to raise awareness about the mental health resources and safety tools available on Facebook.

V. Policy Improvements

Meta has an extensive [policy development and improvement process](#) that helps ensure we continue to take local context into consideration when writing and enforcing our policies. In the Philippines, we have paid careful attention to local context and worked with local experts and NGOs to refine our global rules, such as academics from the University of Philippines, the Legal Network for Truthful Elections (LENTE), and the [Foundation for Media Alternatives](#) (FMA).

Some improvements to date:

- Creating context-specific content policies in 2017 aimed at mitigating the risk of ‘outing’ people based on certain individual characteristics. This was an early effort to combat what would come to be known as red-tagging. In 2019, improving our policies against outing of individuals based on additional characteristics.
- Strengthening our policies against [gender-based bullying](#) and harassment, and updating our hate speech policies to prohibit gendered [cursing](#).
- Updating our policies to address challenges posed by the proliferation of empty accounts with prolonged dormancy. We updated the [Account Integrity](#) and Authentic Identity portions of our Community Standards, and are developing product tools to continue to identify and remove such accounts.
- Launching important new policies against [bullying and harassment](#), including a policy against mass harassment, as well as increased protection for human rights defenders, journalists, and female public figures.
- Launching the Meta Community Standards in Filipino.

VI. Scaling Safety and Security

In order to effectively enforce our policies, we have undertaken extensive efforts over the past few years to increase the number of people working on trust, safety and security globally, and for the Philippines, as well as scale up algorithmic efforts. Today, we have more than 40,000 people working on safety and security globally.

Some of our efforts to date:

- In 2018, we expanded staffing in the Asia-Pacific region, recruited a local public policy and communications team in the Philippines and informed our Community Standards development process by working closely with local stakeholders.
- We formalized partnerships with local, non-governmental and civil society organizations, known as Trusted Partners, that can alert us to and provide us with additional context about emerging issues. In 2019, we expanded our Philippines-related content review workforce, adding capacity in Cebuano.
- We invested in Filipino-specific artificial intelligence tools related to hate speech and child safety, enabling us to better detect and prevent platform abuse.

- In one specific enforcement action, we removed [a network](#) of 95 Pages and 39 accounts for violating our spam and authenticity policies. These accounts had names like Duterte Media, Duterte sa Pagbabago BUKAS, DDS, Duterte Phenomenon, DU30 Trending News, Hot Babes and others. 4.8 million people followed at least one of these pages, all of which linked to the same adfarm.
- We significantly expanded our specialized global teams working on cybersecurity and misinformation. And we established a team — spanning product, engineering, policy, research, and operations — to proactively prevent platform abuse in countries at risk of [conflict](#). These efforts improved our ability to minimize abuses on the platform while also respecting people’s fundamental rights to free expression and association.
- We enabled a comment warning on Instagram that encouraged people to pause before posting potentially hurtful comments.
- We strengthened practical support for our child safety efforts, including by funding development of and supporting access to National Center for Missing & Exploited Children’s (NCMEC) case management tool for local law enforcement agencies, making it easier for them to prioritize and respond to reports of child sexual exploitation, including in the Philippines.
- We updated our artificial intelligence tools to better detect content using local language terms relating to the sexual exploitation of children. This change helped us more rapidly review content.
- We improved safety and privacy for children when we launched [Messenger Kids](#), an app to allow kids to use video and chat functions in a parent-controlled environment. To prevent unwanted contacts in messenger we also created Messenger Safety Tips and launched campaigns and partnerships with the Child Rights Network and Stairway Foundation, including developing child safety GIF storybooks.
- We sought to increase action, awareness and enforcement on other human exploitation issues, including organ trafficking and domestic servitude. Our efforts included roundtable sessions with the Inter-Agency Council Against Trafficking, the Department of Health, the National Kidney and Transplant Institute, the Overseas Workers Welfare Administration, the Blas Ople Policy Center, the Department of Foreign Affairs and the Philippine Overseas Employment Administration.

VII. Fighting Misinformation

During the period under assessment, Meta took multiple actions to fight misinformation, both globally and in the Philippines. These included skills-focused efforts aimed at helping the public identify misinformation when they see it, as well as efforts to label, demote, and remove misinformation related to COVID-19. We also set up [third party fact-checking programs](#), and implemented capacity building programs for the fact checking community in the Philippines, and globally. This is also a focus area of our digital literacy training program, Digital Tayo.

Some specifics on our efforts:

- Adopted (and later expanded) a global policy to remove verifiable misinformation at risk of contributing to imminent physical harm. This policy applies to all content, including the direct and unedited speech of politicians.
- Established a Philippines [third party fact checking program](#) in partnership with global news agency Agence France Press and local media groups Rappler and Vera Files. The program, certified by the non-partisan [International Fact-Checking Network](#), enables independent fact checkers to rate and review the accuracy of content on the platform, thus helping us combat viral misinformation.
- Partnered with The Asia Foundation to launch [The Resiliency Initiative](#), a counterspeech effort against terrorism, violent extremism, and hate.
- Sought to combat COVID-19 misinformation, with the onset of the global pandemic in 2020, while continuing our long term work. In particular, we:
 - launched the Misleading Manipulated Video policy, which prohibits “deep fake” videos that distort reality and are highly misleading
 - published [guidelines](#) to apply our Misinformation and Harm policy to remove COVID-19 and vaccine related misinformation
 - established a global COVID-19 Information Hub that included local partners and sources of information as well as a Philippines specific [COVID-19 Messenger Bot](#) with the Department of Health (known as “KIRA”)
 - provided digital advocacy training for the Department of Health to supplement their online efforts in raising public awareness about COVID-19
 - provided ad credits to help promote the agency’s campaign and announcements

- provided ad credits to the [Philippine Red Cross](#), the [Philippine Disaster Resilience Foundation](#) and [UNICEF Philippines](#) for their own COVID-19 campaigns
- Globally, we slowed down the sharing of frequently forwarded messages in [Messenger](#) and [WhatsApp](#), an effective way to slow down the spread of viral misinformation and content that has the potential to cause offline harm.
- We also strengthened penalties against users who [repeatedly share](#) misinformation.
- We have also expanded our support to Philippines fact checking groups. [Internews](#), supported by Meta, launched the Philippine Fact-Checker Incubator Program in 2021 to support six new fact-checking organizations ahead of the 2022 Philippine general elections.
- Meta also launched a global accelerator [amongst our fact-checking community](#). It brings fact-checking organizations together to innovate, learn, and collaborate on new strategies to improve their business — on and off Facebook.

VIII. Cybersecurity Enforcement

We've taken a number of actions to enforce our cybersecurity policies, including taking specific enforcement actions, and removing networks violating our policy against coordinated inauthentic behavior (CIB).

The first action was in January 2019, when we [banned Twinmark Enterprises](#), a digital marketing group in the Philippines that repeatedly violated our misrepresentation and spam policies — including through CIB, the use of fake accounts, leading people to ad farms and selling access to Facebook Pages to artificially increase distribution and generate profit.

The second took place in [March 2019](#), when we removed 200 Pages, Groups, and accounts linked to a network organized by former Duterte campaign social media director, Nicanor Gabunada Jr.

In [September 2020](#), we removed two separate networks (among others) for violating our policy against CIB. One originated in China, focused on audiences in the Philippines, Southeast Asia and the United States; the second originated from politically affiliated actors in the Philippines, and focused on domestic audiences.

For more insight into our global cybersecurity enforcement and learning, please see Meta's recent cybersecurity reporting including our May 2021 [Threat Report: The State of Influence Operations 2017-2020](#).

IX. Support for Journalists and Human Rights Defenders

Meta recognizes human rights defenders, including journalists, are a high risk user group. This recognition is part of our [corporate human rights policy](#), adopted in March 2021.

Some of our efforts to date:

- In 2019, we launched the [Facebook Journalism Project](#). This is an initiative with publishers around the world to strengthen the connection between journalists and the communities they serve. In the Philippines in particular, the Facebook Journalism Project supported [Tsek.Ph](#) in partnership with local organizations, fact-checkers and universities.
- In 2020, we continued our ongoing work with journalists and news organizations to train them in best practices for optimizing digital skill sets and business practices. Offerings included [bootcamps](#) in Community Standards, business analytics and strategies, an executive leadership program, as well as scaled training in mobile journalism.
- In 2021, we launched the Online Security and Resiliency Training for Human Rights Defenders, a training program with The Asia Foundation focusing on digital security, mental health and community engagement.
- We also sought to improve journalist and human rights defender safety by:
 - introducing stronger security features via the [Journalist Registration](#) program
 - briefing Human Rights defenders on Meta's Community Standards and data sharing policies
 - training defenders on using Facebook for digital advocacy
- We have also committed to expanding our protections for journalists in the Asia-Pacific region who are often at risk because of their work. This includes support of a new online learning program with the [International Center for Journalists](#) (IFCJ). This program will be available in the coming months.
- This support is complemented by our pilot [Human Rights Defender Fund and Journalist Safety Initiative](#), launched in partnership with [Civil Rights Defenders](#) and the

International Center for Journalists. The project will give offline assistance to human rights defenders, including journalists, facing critical threats and support new digital security efforts, beginning in Asia. The fund may be used for a range of support including but not limited to technology and devices, safety and protection software, emergency legal and safety support — to those individuals targeted as a result of their activities in defense of human rights. This project also includes support for a new e-learning program with the ICFJ and online safety resources specifically for journalists who are at risk. These resources will be available in early 2022.

- Other support for journalists includes programs such as the [Reuters Digital Journalism Foundations](#) course and the [Facebook Fundamentals for Newsrooms](#), that includes training on staying safe across social media and on our platforms. Recently, we also announced [our partnership with the Rory Peck Trust](#) to help freelance journalists globally.