

FACEBOOK RESPONSE

Indonesia Human Rights Impact Assessment

Overview

Indonesia's citizenry is large, diverse, and highly active digitally; it is one of Facebook's largest markets globally by number of active users.

Although Indonesia has laws protecting freedom of expression, laws criminalizing blasphemy and defamation inhibit the exercise of that freedom. Some 200 people were subjected to criminal investigation under the country's defamation law from 2014-2019. In 2018, some 30 individuals were incarcerated on blasphemy-related charges, and another eight people were jailed in 2019.¹

Indonesia has a large number of media outlets, although journalistic independence and standards vary widely, and misinformation is widespread. It ranks 119th out of 180 countries surveyed in the 2020 Reporters Without Borders [World Press Freedom Index](#), and scores 51/100 ("partly free") in the 2019 Freedom House [Freedom on the Net report](#).

This response to Article One's assessment summarizes the HRIA's methodology and key findings, gives a detailed overview of the mitigations we have implemented, and notes how we'll follow up on other HRIA recommendations.

About This HRIA

We enlisted Article One, a specialized human rights and ethics consulting firm, to independently assess the human rights impacts of our services in Indonesia as part of our shift towards intensified human rights due diligence and regular disclosure. The work they did, the findings they

shared and the recommendations they made helped us better prepare for the country's general election in April 2019 - and will continue to inform our long-term work in Indonesia.

The HRIA's methodology was based on the [UN Guiding Principles on Business and Human Rights](#), and included interviews with 35 organizations. Special attention was paid to ensuring direct engagement with Indonesian civil society stakeholders and international experts.

The Article One research team sought to assess Facebook's positive and negative human rights impacts in Indonesia, examining the time period from early adoption of Facebook by Indonesians until December 2018, when the team conducted in-country engagement. The assessment does not cover events or developments after December 2018.

We are sharing the HRIA's executive summary now, along with this response, to help local and international stakeholders better contextualize Facebook's impacts and investments in Indonesia. It is also part of Facebook's broader commitment to meaningful transparency about our human rights due diligence, and our product integrity work.

We note that while many of the recommendations in the HRIA and our responses are specific to Indonesia, many also have implications for other contexts, and are reflected across multiple HRIAs that Facebook has commissioned. This reflects the universal nature of human rights, the global reach of our products, and the intersectionality of the impacts identified.

What's an HRIA?

A human rights impact assessment (HRIA) is a detailed form of human rights due diligence. An HRIA allows Facebook to identify human rights risks; know its positive and negative human rights impacts; and strengthen positive impacts as well as mitigate human rights harms.

Facebook has committed to do human rights due diligence, including HRIAs, as a member of the [Global Network Initiative](#), which we joined in 2013. The UN Guiding Principles, applicable globally, also state businesses should conduct human rights due diligence, and HRIAs should include meaningful consultation with potentially affected groups and other relevant stakeholders (2.A.18(b)).

Why This Facebook HRIA Response?

There's no standard format for disclosure. Facebook has chosen the unusually transparent step of disclosing HRIA executive summaries, with recommendations. This 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.

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

Indonesia HRIA Findings

The HRIA found that Facebook has had important, positive human rights impacts in Indonesia.

Among the positive impacts, the research found that Facebook had:

- Advanced freedoms of expression, association, and opinion;
- Enabled citizen journalism, facilitated increased awareness of human rights violations, and promoted the right to take part in government;
- Assisted emergency response after the Donggala earthquake and subsequent Tsunami in September 2018 through product features and partnerships, thereby supporting the right to security of person;
- Made significant economic contributions to Indonesia, providing opportunities for small businesses to access new markets, and for traditionally marginalized groups to develop and market small businesses from their homes. These impacts contributed to the right to work and the right to an adequate standard of living; and
- Enabled virtual and real community, creating safe and supportive spaces for vulnerable groups, including women and LGBTQ+ people.

The report also found Facebook was associated with salient human rights risks in Indonesia. Many, but not all, involved behavior by some Facebook users that adversely impacted other users.

Risks and adverse impacts included:

- The proliferation of misinformation and disinformation, locally described as “hoaxes”. Among the hoaxes that have gained traction are those related to local and national elections, kidnapping and health-related information. The report indicated that low rates of search engine use and of digital literacy, as well as product features enabling virality, may have contributed to the spread of such disinformation and misinformation;
- Freedom of expression risks related to government actions seeking to enforce blasphemy and defamation laws against individuals based on their social media usage. Fear of such actions resulted in self censorship by users. To the extent that Facebook may have restricted access to content on the basis of these laws, the restrictions may have infringed users' rights to freedom of expression. (Facebook notes it has strong processes in place to protect freedom of expression and the right to privacy as defined by articles 17 and 19 of the ICCPR, the result of commitments made a member of the Global Network Initiative.²);
- Privacy risks related to users inappropriately disclosing the personal information of others (doxing), as well as potential Cambridge Analytica privacy violations and generalized concerns that the government might arbitrarily seek access to user data. Facebook reports actions taken in response to government user data requests in its transparency reporting;
- Cyberbullying and harassment of women who were outspoken online, including the non-consensual public sharing of intimate images (NCII), and the use of social media, including Facebook and Instagram, as a potential means of connecting customers to sex workers.³ In any cases where Facebook was used to connect customers to sex workers who had been trafficked, the action of those users would infringe the right to be free from slavery;
- Bullying of children and evidence of child sexual exploitation, including online grooming;⁴ and
- Harassment, bullying, and involuntary “outing” of LGBTQ+ users, even though Facebook has also provided a safe space in which LGBTQ+ individuals communicated and organized around LGBTQ+ rights.

2. As a member of the [Global Network Initiative](#), Facebook has strict processes in place to respect freedom of expression and privacy as defined under articles 17 and 19 of the ICCPR. We are externally assessed against these commitments every two years. The 2018-2019 public assessment report is available here: <https://globalnetworkinitiative.org/wp-content/uploads/2020/04/2018-2019-PAR.pdf>

3. Human rights reporting indicates each of Indonesia's 34 provinces is a source and destination of trafficking. United States Department of State Trafficking in Persons Report 2019 <https://www.state.gov/reports/2019-trafficking-in-persons-report-2/indonesia/>

4. The [US Department of State 2019 Trafficking in Persons Report](#)

Facebook Response to Findings and Recommendations

The report also found Facebook was associated with salient human rights risks in Indonesia. Many, but not all, involved behavior by some Facebook users that adversely impacted other users.

Research for the Indonesia HRIA took place at a time when Facebook was carrying out parallel efforts to strengthen its understanding of, and work in, contexts with elevated social or political risks.

Since that time, we have set up a dedicated, multi-disciplinary team with subject matter experts focused on countries at risk of conflict. Over the last two years, the team has formalized an approach for defining and prioritizing countries at risk of conflict, and tailoring policy and product solutions to account for the unique needs of individual countries. We have invested in changes to platform architecture, rolled out a variety of product interventions to limit the spread of problematic content; and improved proactive detection of potentially violating or inflammatory content to mitigate human rights risks.

We have further intensified our human rights due diligence, and significantly expanded risk mitigation efforts prior to elections—including prior to Indonesia's 2019 national elections.

A. WHAT WE'VE IMPLEMENTED

Facebook has made significant changes as a result of the Indonesia HRIA and related processes. These are grouped according to the seven recommendation areas in the HRIA report.

Improve Corporate-Level Accountability

We have:

- Continued to comply with our commitments to privacy and freedom of expression as a member of the Global Network Initiative (GNI). Earlier this year, GNI completed its bi-annual assessment of Facebook, finding that we “strengthened its systematic review of both privacy and freedom of expression” and “is making good-faith efforts to implement the GNI Principles with improvement over time”;
- Formalized an approach to help us determine which countries require high priority conflict prevention interventions, such as product changes, UX research, or other risk mitigation steps;

- Incorporated human rights principles into the Community Standards Values in September 2019;
- Created a new senior role to lead company work on human rights, and are expanding related roles and resources;
- Built clear processes for expert human rights input into product and content policy development; and
- Significantly improved mechanisms for user control over their privacy and information, with new features like Privacy Checkup, Privacy Shortcuts, Off-Facebook Activity, and others, and have worked to increase awareness of these features and tools through partnerships with local CSOs and to conduct consumer awareness initiatives on privacy features in Jakarta and Bandung.

Evolve Community Standards

We have:

- More than doubled the number of local language speaking content reviewers, including full time staff and scaled support, to provide 24/7 content moderation coverage in Bahasa Indonesia. All content moderators and outsourced operations are audited;
- Created and hired new staff positions dedicated to Indonesia policy and programs, enabling significantly better local consultation with CSOs and experts on the evolution of the Community Standards;⁵
- Established and intensified a regular cadence of engagement with key stakeholders on the ground, including listening sessions, gathering input on FB policy, under the hood session on Community Standards, and co-design workshops and social media training together with CSOs;
- Created a new policy to remove verified misinformation and unverifiable rumors that may contribute to the risk of imminent offline physical harm. This policy is highly relevant to conflict-affected and other challenging settings;⁶
- Updated existing policies to protect vulnerable users, including the protection of users whose “outing” might increase risks of offline harm (e.g. involuntary outing of veiled women, LGBTQ+ individuals or human rights activists);
- Expanded our bullying policies to increase protections provided to all individuals, including public figures such as human rights

5. For a detailed overview of our global stakeholder engagement process and how it contributes to policy making, see https://www.facebook.com/communitystandards/stakeholder_engagement. Minutes of relevant policy meetings at <https://about.fb.com/news/2018/11/content-standards-forum-minutes/>

6. For a detailed overview of our global stakeholder engagement process and how it contributes to policy making, see https://www.facebook.com/communitystandards/stakeholder_engagement. Minutes of relevant policy meetings at <https://about.fb.com/news/2018/11/content-standards-forum-minutes/>

defenders and journalists (we have, for example, updated our policies to explicitly prohibit female-gendered cursing and attacks on the basis of derogatory terms related to sexual activity). The updates we've made to our bullying and harassment policies also cover dehumanizing speech, which is referenced among Article One's recommendations on hate speech; and

- Expanded our policies against voter interference to prohibit misrepresentations about how to vote, and statements about whether a vote will be counted, which enabled us to remove election-related misinformation ahead of Indonesia's 2019 presidential election.

Invest in Changes to Platform Architecture

We have:

- Developed proactive hate speech detection technology in Bahasa Indonesia to help us more quickly and effectively identify potentially violating content, reducing Facebook's dependence on user reporting;
- Increased friction for sharing problematic content across the platform. We have made product interventions on WhatsApp to limit the spread of disinformation — clearly labelling forwarded messages, capping forward limits for all messages to five, and lowering the limit for highly forwarded messages to just one, with the latter cutting the virality of highly forwarded messages by 70%;
- Invested in new technologies to proactively detect child nudity and child exploitative content when it's uploaded. This investment enables us to enhance our reporting to law enforcement. It also means we can identify accounts seeking to engage in inappropriate interactions with children more quickly, so we can remove them and mitigate risk of harm. We have open sourced this technology to encourage use by others;
- Deployed new tools to fight sharing of non-consensual intimate images. We're using machine learning and artificial intelligence to proactively detect near nude images or videos that are shared without permission on Facebook and Instagram. This enables us to limit harms by finding this content before anyone reports it; and
- Raised user awareness of mitigation approaches with the global NCII victim-support hub in our Safety Center, Not Without My Consent, available in 50 languages.

Address Challenges to the Platform-Level Grievance Mechanism

We have:

- Launched an improved reporting/blocking tool for Facebook Messenger, greatly increasing the ease of reporting potentially violating content and of blocking unwanted interactions;
- Introduced appeals of content moderation decisions in 2018, and expanded them significantly since to include almost all policy areas/abuse types, seeking alignment with the Santa Clara Principles;
- Significantly improved content moderation support and resources, with increased staffing, enhanced wellness and resiliency resources and improved tools that allow content reviewers to customize some of the ways they do their work; and
- Launched an ambitious independent operational grievance mechanism, the Facebook Oversight Board, whose creation was accompanied by a detailed human rights review.

Ongoing Due Diligence

We have:

- Implemented Indonesia-specific risk mitigation and monitoring efforts prior to the 2019 national elections, including by engaging in regular, extensive CSO consultations; launching multiple user information and awareness products; by putting into place political advertising authorization and transparency tools and setting up a dedicated election operations center to monitor and respond to real-time risks;
- Commissioned PwC Indonesia and INDEF to conduct a robust study to measure Social and Economic Impact of Facebook and Family of Apps in Indonesia to individual, community, organizations, businesses, and government;⁷
- Expanded and intensified human rights due diligence at the country, product, and product intervention level, with follow up due diligence underway; and
- Increased capacity to identify and mitigate human rights risks in real time, as a result of the combined effect of multiple other workstreams.

Transparency

We have:

- Published the internal guidelines we use to enforce our Community Standards such that our public-facing policies include the granular details of what is and is not allowed on Facebook;
- Invited external guests to the cross-functional policy development meeting, the Product Policy Forum, at which we discuss and debate changes to our Community Standards, ads policies and major News Feed ranking changes. We also make minutes from the meeting available to the public;
- Continued to make metrics available that track how well we're doing at enforcing our Community Standards in a bi-annual Community Standards Enforcement Report. As of November 2019, the report includes numbers on how much content people appealed; and
- Developed clear criteria to inform the partnerships we develop at the local level.

Use Leverage to Address Root Cause Challenges

We have created strong processes for participating in industry and multi-stakeholder efforts to respect and advance human rights. For example:

- We have supported local civil society in a number of ways, among them we have:
 - Significantly strengthening of our local third-party fact checking program to include six Indonesian fact-checking organizations, launched in early 2018, to combat misinformation;
 - Trained more than 24,000 individuals and 60 civil society organizations in Indonesia on effective and safely use of social media for cause related work, including on privacy, digital rights, safety, and women's rights; and
 - Expanded on the ground partnerships for safety and other programs. Local organizations have specific training and direct channels for escalating and acting on harmful content, are regularly consulted on issues related to the Community Standards, Facebook's content moderation policies;

- On misinformation and disinformation specifically:
 - We now work closely with industry partners, meeting on a bi-weekly basis to share learnings and work together to combat disinformation and influence operations;
 - When conducting our Coordinated Inauthentic Behavior takedowns, Facebook shares information and often works together with its industry partners. Facebook also collaborates with civil society investigators, including at the Atlantic Council's Digital Forensics Research Lab, Graphika, Stanford's Internet Observatory, and others around the world; and
 - We are a partner in the multi-stakeholder Carnegie Partnership to Counter Influence Operations, and regularly participate in multi-stakeholder efforts in Indonesia and in the US to study and collaborate on combatting disinformation.

Together, all these actions have enhanced Facebook's capacity to act to mitigate crises, and to respect human rights.

In 2019, for example, Facebook was able to respond in a timely and proactive manner to platform and expression risks during Indonesian elections and subsequent riots, as well as during West Papua-related demonstrations. It also initiated a related takedown for coordinated inauthentic behavior.

B. WHAT'S IN PROCESS

Facebook recognizes human rights due diligence is an ongoing process, and we have not yet implemented all the HRIA's recommendations.

Work in progress includes:

- Building on our human rights accountability and governance strategy, which is evolving rapidly;
- Designing the first iteration of a human rights defenders program, based on the findings of the this and other HRIAs, and specific due diligence on the needs of human rights defenders;
- Increasing the practical tools we can use to educate users about our Community Standards in diverse local languages;

- Working together with Civil Society Organizations, business associations, and government to promote regulations that more align with international human rights norms; and
- Promoting the right to work and the right to an adequate standard of living, we empower individuals, women and young entrepreneurs, and local businesses, through our #ShemeansBusiness, *Laju Digital*, and Akademi Instagram. At the time of writing in May 2020, we had engaged more than 40,000 entrepreneurs and SMBs, 1,300 young entrepreneurs, and 28,000 women entrepreneurs.

Likewise, Facebook has significantly expanded its contact with CSOs and other local stakeholders, but there is much more to be done, including in working on the ground with multi-stakeholders to ensure our follow up to this report is thoughtful, effective, and relevant.

C. WHAT WE HAVEN'T IMPLEMENTED

There are further recommendations we will continue to explore with HRIA authors and Indonesian stakeholders, including the development of additional tools to empower users. We also, separately, note the need to educate users more broadly on our transparency reporting, and on relevant GNI-related commitments to protect freedom of expression and privacy.

It is also important to be transparent, and note that some recommendations may not currently be technically or operationally feasible. The recommendation to develop AI to predict when online hate speech may trigger offline violence, for example, is possibly beyond the bounds of current global technical and academic capacity.

Similarly, we are unlikely to create a process involving ongoing open global calls for comment on changes to our Community Standards, given the scale involved, although we do conduct extensive stakeholder engagement as part of our policy development process.

We'll continue to identify potential solutions to human rights challenges in partnership with other stakeholders, and to implement them as they're developed.

Final Note

Internet ecosystem risks in Indonesia remain very real. Even though we have worked very hard to minimize risks, we do not want to suggest our response to this HRIA means we've solved all those risks, or can prevent all future harms.

Still, this human rights impact assessment is an important step forward for Facebook, combined with the simultaneous disclosure of HRIAs on Sri Lanka and Cambodia.

By detailing the ways in which we have sought to implement Article One's recommendations—as well as where we have yet to act—we have attempted to demonstrate accountability for our human rights impact.

Acknowledgments

We are deeply grateful to the many human rights defenders, civil society organisations, and others who took the time and thought to participate in this assessment. We believe the process and findings of this HRIA have helped make a significant difference to Facebook's work in Indonesia, and around the world.