

2018

**ASSESSING THE
HUMAN RIGHTS IMPACT
OF THE FACEBOOK PLATFORM
IN SRI LANKA**

ARTICLE ONE

EXECUTIVE SUMMARY FOR FACEBOOK INC.

PHOTO BY THAROUSHAN KANDARAJAH

EXECUTIVE SUMMARY

The Facebook platform has broken down barriers across the world. It has made neighbors of strangers and has, as its mission states, “*brought the world closer together.*” It has fundamentally altered the ability for people to freely express themselves by democratizing access to powerful channels of communication.

At the same time, in some cases, the platform has been misused by bad actors to infringe on human rights. This includes using the platform to amplify hate speech, spread misinformation and disinformation, target and harass vulnerable groups, and infringe on the right to privacy.

Sri Lanka represents one of the most critical countries when it comes to potential human rights infringements on the platform. Recently emerging from a decades-long civil war, Sri Lanka continues to struggle with ethnic and religious tensions that play out both offline and online. For example, harmful rumors about Muslims in Sri Lanka spread across the Facebook platform in March of 2018. These rumors fanned existing ethnoreligious tensions and may have contributed to offline violence against Muslims.¹

To determine the degree to which Facebook’s platform may or may not have contributed to adverse human rights impacts, and to mitigate the risk of further impacts, Facebook partnered with Article One in 2018 to conduct a country-level human rights impact assessment (HRIA) of its platform in Sri Lanka.

The HRIA is designed to pursue the following objectives:

1. **Assess the role Facebook plays in Sri Lanka through a human rights lens.¹**
2. **Proactively address risks to avoid harm to rightsholders, including the platform’s users.**
3. **Determine how best to evolve and apply Facebook’s global policies and standards in high-risk contexts.**
4. **Enable Facebook to maximize opportunities for positive human rights impacts.**
5. **Inform internal risk assessments moving forward.**

This Executive Summary outlines the findings of the HRIA and presents recommendations for Facebook to mitigate actual and potential impacts that are in line with the expectations of the UN Guiding Principles on Business and Human Rights (UNGPs).

It is important to note that the findings from this assessment do not include potential impacts after September 2018, when Article One conducted its two field trips to Sri Lanka. While the situation in Sri Lanka has changed significantly since that time, this assessment does not provide an overview of how those changes may have played out on the Facebook platform. Nor does it incorporate any Facebook interventions that may have taken place, on or off-platform, since September 2018.

¹ New York Times (2018) “Where Countries Are Tinderboxes and Facebook Is a Match” and The Conversation (2018) “Violent Buddhist extremists are targeting Muslims in Sri Lanka.”

² Informed by the [United Nations Guiding Principles on Business and Human Rights \(UNGPs\)](#) and the [Global Network Initiative \(GNI\) Principles](#)

METHODOLOGY

The HRIA methodology was informed by guidance from the UNGPs and the GNI Principles as well as Article One’s award-winning methodology for, and experience in, conducting human rights impact assessments around the world. The assessment included six key steps, as shown on the right.

The methodology emphasized the voices of rightsholders potentially impacted by Facebook and their legitimate representatives. As such, special attention was paid to ensuring direct engagement with Sri Lankan civil society organizations, Facebook users and international experts. Engagements included interviews with 29 civil society organizations and other experts. In addition, Article One observed Facebook-led focus groups with 150 participants who regularly engage with the Facebook platform.

In-country engagements were conducted the weeks of August 13, 2018 and September 3, 2018.

Article One applied the methodological approach outlined in the UNGPs to determine the saliency of and Facebook’s responsibility for the potential impacts surfaced during the assessment. To determine relative priority of salient human rights impacts, Article One assessed the scale, scope and remediability of impacts. Facebook’s relationship to impacts was assessed using guidance from UN Human Rights, including whether the company: a) incentivized harm, b) facilitated harm, c) failed to conduct adequate human rights due diligence, and d) the extent to which it knew or should have known about adverse impacts.³

Article One then developed a series of recommendations to support Facebook in maximizing its positive human rights impacts and mitigating its adverse impacts in Sri Lanka.

CONTEXTUAL ANALYSIS

Ethnic conflict between the Sinhalese Buddhist majority and the Tamil minority has dominated Sri Lanka’s political landscape since the nation first declared independence in 1948.

While the civil war officially ended in 2009, ethnic divides continue to shape Sri Lankan life and politics. Both news reports and civil society have highlighted the role of Sinhalese Buddhist nationalism in politics, outlining a “majority-minority mind frame” that fuels tensions and “a belief that Sri Lanka is a Buddhist nation that must be protected from foreign elements, violently if necessary.”⁴ The rise of Sinhalese Buddhist nationalism has deepened sectarian divides and contributed to new waves of violence.

At the time of writing in late 2018, human rights risks remain high, with discrimination against minorities, arbitrary arrests, surveillance and harassment of civil society activists and journalists among some of the most serious risks in-country.⁵

Tamils maintain that widespread discrimination continues to take place in university education, government employment, housing, health services, language laws, and procedures for the naturalization of noncitizens. Furthermore, growing fears by the Sinhalese majority that the international Islamic community may someday take over Sri Lanka have accelerated Sinhalese-Muslim tensions.⁶ These tensions have played out most tragically on the ground, but also on Facebook, where the platform has 6.85 million monthly active users and 4.4 million daily active users.



Step 1: In-Country Engagement



Step 2: Desk Review



Step 3: Stakeholder Mapping



Step 4: In-Country Engagement



Step 5: Facebook & International Expert Engagement



Step 6: Analysis & Internal Report

³ OHCHR (2018) “OHCHR response to request from BankTrack for advice regarding the application of the UN Guiding Principles on Business and Human Rights in the context of the banking sector” and OECD (2018) “OECD Due Diligence Guidance for Responsible Business Conduct” ⁴ The Conversation (2018) “Violent Buddhist extremists are targeting Muslims in Sri Lanka.” ⁵ Human Rights Watch (2018) “World Report: Sri Lanka” ⁶ Minority Rights Group International (2018) “Sri Lanka Country Profile”

POSITIVE HUMAN RIGHTS IMPACTS

Facebook's economic contribution to Sri Lanka has been significant. Facebook has provided opportunities for small businesses to access new markets, both domestically and internationally. It has provided opportunities for groups who have traditionally been excluded from formal employment to develop and market small businesses from their homes. For historically disadvantaged groups, these opportunities are essential to realizing their human right to work and to an adequate standard of living.

Article One repeatedly heard that Facebook had played a powerful role when it came to free expression and civic engagement in Sri Lanka. Civil society groups, women, LGBTQ+ individuals, human rights activists, and other groups have historically found limited platforms for their activism. Facebook changed that by allowing these groups to communicate and connect with a broader network in a fast and cost effective manner.

Despite frequent public criticism of Facebook, the majority of stakeholders engaged, including Sri Lanka-based human rights advocates, defended the platform and argued that its value to the country—if managed appropriately—was significant.

Finally, Facebook's efforts around digital literacy and emergency response are cited as important contributions in the country.

SALIENT HUMAN RIGHTS RISKS

Salient human rights risks are divided based on impacts to vulnerable groups, aligning with the expectation of the UNGPs to pay special attention to the experience of vulnerable populations. This decision was also informed by Article One's in-country

engagements which highlighted that the most severe issues were borne by vulnerable groups, including religious minorities, women, LGBTQ communities, children, and human rights defenders. It is important to note that intersectionality presents heightened challenges for vulnerable groups.

Salient issues are informed by the relevant articles in human rights instruments, including the Universal Declaration of Human Rights (UDHR) and the Convention on the Rights of the Child (CRC).⁷

A. IMPACTS ON RELIGIOUS MINORITIES

One major source of salient human rights risk related to the Facebook platform in Sri Lanka is the spread of rumors and hate speech to incite violence against Muslim minorities. The March 2018 violence in Kandy is, at the time of writing, the most publicized incident of such violence, but not the only one.

Article One's assessment showed that the Facebook platform contributed to spreading rumors and hate speech, which may have led to "offline" violence. Indeed, the assessment found that the proliferation of hate speech (e.g., *"Kill all Muslims, don't even save an infant; they are dogs"*) and misinformation (e.g., that a Muslim restaurateur was adding sterilization pills to his customers' food) may have contributed to unrest and in the case of the restaurateur and others, physical harm.

The actions of Facebook users may infringe upon a minority's right to:

- be treated with dignity (UDHR 1);
- non-discrimination (UDHR 2);
- life and security of person (UDHR 3);
- property (UDHR 17), and;
- freedom of religion (UDHR 18).

⁷ United Nations (1948) [Universal Declaration of Human Rights](#) and OHCHR (1989) [The Convention on the Rights of the Child](#).

B. IMPACTS ON WOMEN

Women face similar risks online as they do offline in Sri Lanka.

However, cyberviolence has the added effects of significant reputational impacts on women, which in turn can have follow-on impacts related to family and work. Article One's assessment found that Facebook's platform was used for gender-based hate speech and harassment ("cyberbullying"), non-consensual sharing of images in public ("creepshots" and "upskirting"), and non-consensual sharing of intimate images ("revenge porn").

The actions of Facebook users may infringe upon a woman's right to:

- be treated with dignity (UDHR 1);
- non-discrimination (UDHR 2);
- security of person (UDHR 3) if the threat of posting images results in women being placed in harm's way;
- privacy (UDHR 12);
- freedom of expression and assembly (UDHR 19 and UDHR 20) if cyberbullying results in self-censorship or disengagement from the platform, and;
- work (UDHR 23) and to an adequate standard of living (UDHR 25) as follow-on impacts if images shared on the platform impact the ability to maintain or secure employment.

C. IMPACTS ON CHILDREN

Children face unique vulnerabilities online, including on the Facebook platform. Article One found evidence of child sexual exploitation related to the posting of sexually explicit photos and comments related to children, as well as online grooming of children. These included comments on images of children such

as *"sexy fuck her hard"* and a #HeForShe grooming tactic that was found to have taken place on Facebook Messenger.

The actions of Facebook users may affect a child's right to:

- be treated with dignity (UDHR 1);
- security of person (UDHR 2) if the sharing of these images results in self-harm or sexual exploitation offline;
- privacy (UDHR 12);
- be free from attacks on a child's reputation (CRC 6), and;
- be protected from sexual exploitation and abuse (CRC 34).

D. IMPACTS ON THE LGBTQ+ COMMUNITY

The Facebook platform simultaneously provides a safe space in which LGBTQ+ individuals can communicate and organize around LGBTQ+ rights, and also a space in which anti-LGBTQ+ hate groups can form and spread viral and vicious campaigns against both individuals and groups. LGBTQ+ Facebook users in Sri Lanka have experienced harassment and bullying as well as experiences of being "outed" on the platform.

The actions of Facebook users may infringe upon an LGBTQ+ individual's right to:

- be treated with dignity (UDHR 1);
- non-discrimination (UDHR 2);
- security of person (UDHR 3) if online harassment leads to offline harm;
- privacy (UDHR 12) if individuals are "outed" by other users or by the platform architecture;
- free expression (UDHR 19) and free assembly (UDHR 20) if harassment leads to self-censorship or a reduction in

engagement on the Facebook platform, and;

- work (UDHR 23) and to an adequate standard of living (UDHR 25) as follow-on impacts if images shared on the platform impact the ability to maintain or secure employment.

D. IMPACTS ON HUMAN RIGHTS DEFENDERS

The Facebook platform has been a powerful tool for activism, but aspects of its use present ongoing risks to human rights defenders who may face harassment and surveillance of their online activity, including in relation to Facebook. This includes online harassment of human rights defenders by other users, as well as a potential overreach by government agencies seeking to monitor defender activity online.

The following rights of human rights defenders could be infringed:

- be treated with dignity (UDHR 1);
- security of person (UDHR 3) in so much as harassment has included threats;
- privacy (UDHR 12) if there are illegitimate requests for user data by government authorities, and;
- free expression (UDHR 19) and assembly (UDHR 20) if harassment and the threat of surveillance result in self-censorship or a reduction in engagement on the Facebook platform.

FACEBOOK'S RESPONSE & RECOMMENDATIONS

Under the UNGPs, Facebook should exercise due diligence to mitigate adverse human rights impacts of its business. Article One found, however, that Facebook's response was slow and at times insufficient to meet that standard, which potentially exacerbated impacts. Indeed, Facebook's lack of formal human rights due diligence in Sri Lanka prior to this HRIA and the limited cultural and language expertise among Facebook staff at the time of the May 2018 Kandy incident may have contributed to offline harm stemming from online engagement. This was potentially exacerbated by now phased out algorithms designed to drive engagement on the platform, regardless of the veracity or intention of the content.

The assessment also suggests that, prior to the 2018 internet disruption, Facebook failed to effectively implement its Community Standards in Sri Lanka by failing to take down content that violated these Standards in a timely manner. This resulted in hate speech and other forms of harassment remaining and even spreading on Facebook.

Finally, the majority of civil society organizations engaged by Article One stated that they had tried to engage Facebook regarding the misuse of its platform dating back to 2009. This reportedly included direct calls to the company, and public reports linking the platform

to offline violence. Article One repeatedly heard allegations that Facebook did not respond to these calls promptly, and that the company was largely unresponsive until the government shut down of social media in March of 2018.⁸

Between March and September 2018, Facebook instituted several mitigation tactics, including:

- added policy expertise in a Policy Manager for Sri Lanka;
- added language expertise, including 57 content moderators covering Sri Lanka;
- improved automated detection for the Sinhala language hate speech;
- improved reporting channels through Messenger;
- onboarded additional Trusted Partners, and;
- stakeholder engagement with civil society organizations and human rights defenders.

At the time of writing this report, Article One encouraged Facebook to implement key recommendations, all summarized in the following table.

⁸ New York Times (2018) ["In Sri Lanka, Facebook Contends With Shutdown After Mob Violence"](#)

RECOMMENDATIONS

IMPROVE CORPORATE-LEVEL
ACCOUNTABILITY

- Develop and operationalize a Human Rights Policy
- Formalize approach to developing “At-Risk Country” list
- Explore opportunities for remediation, including for example public apologies for impacts the company may have contributed to
- Increase accountability on human rights through an independent board member responsible for human rights and/or the establishment of a Human Rights Advisory Board
- Respect user’s right to privacy by following guidance from Ranking Digital Rights to “clearly disclose that targeted advertising is on by default,” and improve mechanisms for user control over their information⁹

EVOLVE COMMUNITY STANDARDS

- Fill content gaps in Community Standards
- Make Community Standards sufficiently accessible and understandable, including by ensuring they are translated into all languages spoken by significant numbers of Facebook users in Sri Lanka
- Secure public comments when updating the Community Standards
- Develop a process to apply cultural context to Content Moderation

INVEST IN CHANGES TO
PLATFORM ARCHITECTURE

- Create friction to sharing problematic content that does not violate Community Standards
- Ensure typehead suggestions on the Facebook search function that do not surface sexualized content, especially as it relates to children
- Raise awareness among users of privacy settings
- Raise awareness of mitigation approaches

ADDRESS CHALLENGES TO THE
PLATFORM-LEVEL GRIEVANCE
MECHANISM

- Improve ease of reporting
- Provide appropriate resources for content moderation and audit moderators
- Assess risks of reliance on automation for content moderation
- Expand appeals process in line with the Santa Clara Principles
- Expand the Trusted Partner Network in Sri Lanka

ONGOING DUE DILIGENCE

- Expand Country-Level Human Rights Impact Assessments to other critical countries and conduct follow-up assessments on existing countries whenever material changes occur at the country level
- Conduct product and policy-level assessments
- Develop AI to predict when online hate speech and misinformation may trigger offline unrest and violence

TRANSPARENCY

- Increase transparency when it comes to content decisions
- Disaggregate data based on the source of flagging, including governments, users, trusted flaggers, and automated systems
- Provide users with the ability to opt out of Facebook-driven curation through an easily accessible function on the platform
- Publish data on content moderators, including the number of content moderators disaggregated by language expertise, country of origin, age, gender, and location
- Develop and publish criteria for determining which organizations would qualify as a Trusted Partner, their responsibilities and steps the company will take to ensure they continue to serve the role effectively

USE LEVERAGE TO ADDRESS
ROOT CAUSE CHALLENGES

- Support civil society through financial or in-kind support
- Promote the expansion of privacy laws in Sri Lanka

⁹ Ranking Digital Rights (2018) “[Privacy Failures](#)”

ABOUT ARTICLE ONE

Article One is a specialized ethics and human rights consulting firm. We envision a world in which business succeeds and all people are free and equal in dignity and rights.

In support of that vision, Article One works with companies and institutions to implement strategies to promote corporate respect for human rights, responsible innovation and ethical management.

