

DECEMBER 2020 (REVISED MARCH 2021)

info@democracy-reporting.org
www.democracy-reporting.org

DEMOCRACY
REPORTING
INTERNATIONAL

REGULATING SOCIAL MEDIA IN SRI LANKA

An Analysis of the Legal and Non-Legal Regulatory Frameworks
in the Context of Hate Speech and Disinformation

CONTENTS

LIST OF ABBREVIATIONS	2
INTRODUCTION	3
1. FORMAL REGULATORY FRAMEWORK	5
1.1 The international legal framework on the freedom of expression and the prohibition of the advocacy of hatred	6
1.2 The domestic legal framework	6
1.2.1 International Covenant on Civil and Political Rights (ICCPR) Act, No. 56 of 2007	7
1.2.2 Penal Code, No. 2 of 1883	12
1.2.3 Police Ordinance, No. 16 of 1865	14
1.2.4 Public Security Ordinance, No. 25 of 1947	14
1.2.5 Sri Lanka Telecommunications Act (SLTA), No. 25 of 1991	16
1.2.6 Prevention of Terrorism (Temporary Provisions) Act (PTA), No. 48 of 1979	18
1.2.7 Computer Crimes Act, No. 24 of 2007	19
1.3 Other regulatory mechanisms: Institutions	20
1.3.1 SLCERT – Sri Lanka Computer Emergency Readiness Team / Coordination Centre	20
1.3.2 Cyber Crimes Division of the Criminal Investigation Department	20
1.4 Overview of the legal framework and regulatory bodies	21
2. ALTERNATIVE NON-LEGAL FRAMEWORKS	23
2.1 Voluntary self-regulation by social media companies	23
2.2 Independent international regulation	28
2.3 Counter-messaging by civil society and private institution initiatives	29
CONCLUSIONS AND OBSERVATIONS	33

LIST OF ABBREVIATIONS

ADN	Asia Democracy Network
AFP	Agence France-Presse
AI	Artificial Intelligence
BBS	Bodu Bala Sena
CCA	Computer Crimes Act
CID	Criminal Investigation Division
CMEV	Centre for Monitoring Election Violence
CSO	Civil Society Organisation
DRI	Democracy Reporting International
ER	Emergency Regulation
EU	European Union
HIRA	Human Rights Impact Assessment
HRCSL	Human Rights Commission of Sri Lanka
ICCPR Act	International Covenant on Civil and Political Rights Act, No. 56 of 2007
ICCPR	International Covenant on Civil and Political Rights
ICERD	International Convention on the Elimination of All Forms of Racial Discrimination
ICTA	Information Communication Technology Agency of Sri Lanka
IGP	Inspector General of Police
IHRL	International Human Rights Law
NetzDG	Network Enforcement Act
NPC	National Police Commission
PAFFREL	People's Action for Free and Fair Elections
PCCSL	Press Complaints Commission of Sri Lanka
PSO	Public Security Ordinance, No. 25 of 1947
PTA	Prevention of Terrorism (Temporary Provisions) Act, No. 48 of 1979
SFCG	Search for Common Ground
SLCERT	Sri Lanka Computer Emergency Readiness Team / Coordination Centre
SLPI	Sri Lanka Press Institute
SLTA	Sri Lanka Telecommunications Act, No. 25 of 1991
STF	Special Task Force
TRCSL	Telecommunications Regulatory Commission of Sri Lanka
UDHR	Universal Declaration of Human Rights
UNICEF	United Nations International Children's Emergency Fund
WHO	World Health Organisation

REGULATING SOCIAL MEDIA IN SRI LANKA

An Analysis of the Legal and Non-Legal Regulatory Frameworks in the Context of Hate Speech and Disinformation

INTRODUCTION

Over the past few years, there has been an increasing number of incidents of ethnic and religious violence in Sri Lanka. Many of these incidents involved the spread of hate speech and disinformation on social media, before and after they occurred.¹ Even after the outbreak of COVID-19 in March 2020, the spread of prejudicial and hateful statements has continued to manifest on social media – particularly targeting the Muslim community.²

1. Specifically, violent events in Aluthgama, Gintota (2017) Digana (2018) and communal clashes in the aftermath of the Easter Sunday attacks indicate the frequency and degree of ethno-religious violence in Sri Lanka, which have been associated with spread of hate speech and disinformation in social media. See Events in Aluthgama: Dhari-sha Bastians, 'Death toll rises to 4 from Aluthgama riot', Daily FT 2014, at <http://www.ft.lk/article/308988/Death-toll-rises-to-4-from-Aluthgama-riots> [last accessed 2 December 2019]; Gintota (2017): Dhari-sha Bastians, 'Gintota and the shadows of extremism', Daily FT 2017, at <http://www.ft.lk/opinion/Gintota-and-the-shadows-of-extremism/14-643843> [last accessed 2 December 2019]; 'Digana turns divisive', Daily FT 2018, at <http://www.ft.lk/top-story/Digana-turns-divisive-/26-650661> [last accessed 2 December 2018]; Digana (2018): 'Digana turns divisive', Daily FT 2018, at <http://www.ft.lk/top-story/Digana-turns-divisive-/26-650661> [last accessed 2 December 2018]; and communal clashes: After Sri Lanka's Easter Bombings: Reducing Risks of Future Violence, International Crisis Group, 27 September 2019, at <https://www.crisisgroup.org/asia/south-asia/sri-lanka/302-after-sri-lankas-easter-bombings-reducing-risks-future-violence>; [last accessed 28 April 2021]; also see Gehan Gunatilleke, *The Chronic and the Acute: Post-War Religious Violence in Sri Lanka*, at <http://ices.lk/wp-content/uploads/2016/01/The-Chronic-and-the-Acute-Post-War-Religious-Violence-in-Sri-Lanka.pdf> [last accessed 28 April 2021].

2. 'Free Media Movement calls for collective social consciousness for fight COVID-19 pandemic in Sri Lanka', *EconomyNext*, 10 April 2020, at <https://economynext.com/free-media-movement-calls-for-collective-social-consciousness-to-fight-covid-19-pandemic-in-sri-lanka-64951/> [last accessed 22 April 2020]. Note: In April 2020, the Human Rights Commission of Sri Lanka wrote to the Acting Inspector General of Police regarding the need for limiting the freedom of

As such, social media is often subject to criticisms as the main source and driver of hate speech in Sri Lanka.

The amount of hate speech and disinformation on social media has become a global challenge for states, private bodies, regulatory bodies, policymakers, media and civil society, especially as such online disinformation and hate speech can result in offline harm. Increasingly, this challenge is addressed by looking to more and stricter regulation of discourse on social media with a focus on controlling hate speech and disinformation. However, regulation must be carefully considered and applied to ensure that freedom of speech is upheld, so that democratic discourse and the expression of different, critical and opposing ideas are not curbed.

Countering hate speech and disinformation on social media faces many challenges. First, content constituting hate speech and disinformation is often replicated. The replication and mutation of the original content impose challenges to tracking or deleting several versions of the original, even if the original is taken down.³ The second relates to inadequacies in existing law. International

expression, only within the framework of the law. The letter highlights that many arrests have been made for the spread of misinformation and for criticizing public officials without a legal basis. For full letter see 'HRCSL concerned by police arresting people for criticizing officials' ('HRCSL Letter April 2020'), *Ada Derana*, 29 April 2020, at <http://www.adaderana.lk/news/63153/hrsl-concerned-by-police-arresting-people-for-criticizing-officials> [last accessed 29 April 2020].

3. Shilpa Samaratunge & Sanjana Hattotuwa, *Liking Violence: A study of hate speech on Facebook in Sri Lanka* (Centre for Policy Alternatives 2014), p. 3.

human rights law sets out several standards on the freedom of expression, including the types of speech that can and should be restricted. Therefore, social media content that does not fall within the ambit of hate speech and disinformation, should not be restricted in line with an individual's right to the freedom of speech and expression. However, there is uncertainty in determining which type of content would constitute hate speech or disinformation to be prohibited. Further complications could arise if social media platforms define prohibited content in different ways.⁴ Third, language poses a challenge when detecting content that should be prohibited. For instance, when certain Sinhala and Tamil content is translated into English it may not constitute hate speech as per the definitions outlined in social media guidelines, when cultural/contextual references are lost. Fourth, there are challenges in determining if and how legal models should be designed to hold different actors to account. Some of these actors include internet intermediaries, individuals, the state, and the private sector. For instance, expanding the state's authority by strengthening the legal framework related to combatting hate speech and disinformation may have serious implications on the freedom of speech. One such implication is that the misuse of laws can lead to the persecution of speech, which is merely dissent.

This study explores the existing legal and non-legal frameworks to regulate the spread of disinformation and hate speech on social media. The study presents potential implications of the spread of disinformation and hate speech on social media in the specific context of ethno-religious violence in Sri Lanka. While the study itself does not make recommendations for regulation, the study provides observations of eight key options that may be considered by stakeholders. Such stakeholders include the government, technology/service providing companies and civil society, who may explore these options to potentially mitigate the spread of disinformation and hate speech on social media.

Research Method

Democracy Reporting International (DRI) commissioned Verité Research (Verité) to conduct research and provide

an analysis on the existing legal and non-legal regulatory frameworks for social media in Sri Lanka relating to hate speech and disinformation.

This study was conducted solely as a desk-based research analysis and was compiled from a review of publicly available secondary sources of information. The conclusion has been drawn based on the research analysis. Thereafter, the observations discussed by Verité focused on the possible avenues available in relation to online content moderation in Sri Lanka. The contents of the report are updated as at July 2020. Although the study was mainly a desk-based research, the research team also incorporated feedback following multiple outreach sessions held at the end of 2020.

The study and proposals were further reviewed using a listening and feedback process to discuss the findings and specifically the possible options for combatting hate speech and disinformation on social media at these outreach sessions. Five outreach sessions were held between September 2020 to December 2020, which comprised of representatives from civil society, mainstream and social media, legal professionals and other groups.⁵ The in-person outreach sessions which focused on raising awareness of the relevant frameworks, were conducted in English and Sinhala languages in Colombo and Kandy in September 2020. However, with the outbreak of the COVID-19 pandemic, the remaining sessions in December 2020 took place virtually in Sinhala and Tamil languages, and included participants from Batticaloa, Jaffna, Mannar, Trincomalee, Galle and Matara. The substantive sections of the report have been updated as of July 2020 while the observations were updated as of December 2020 after incorporating feedback from the outreach sessions. The final draft of the report was presented at the end of February 2021. Based on feedback received at this event, the report has been revised in March 2021 to incorporate the relevant feedback.

This study discusses the 1) legal regulatory frameworks (formal) and 2) alternative non-legal frameworks for

4. This is explored in part 2 of this brief.

5. The outreach sessions included a total number of 93 participants. A key observation made at the outreach sessions were that many participants were unaware of the current legal framework on regulating hate speech and disinformation in Sri Lanka.

regulating social media in Sri Lanka. It specifically discusses these frameworks in relation to online hate speech and disinformation. Under the formal regulatory framework, first, this study outlines the international legal framework for the freedom of expression and the prohibition of incitement to hatred. Thereafter, it discusses the relevant domestic legislative and regulatory framework that could be applied with regard to social media in Sri Lanka. With regard to formal frameworks, this study concludes that, while current laws in Sri Lanka cover hate speech and disinformation on social media, their enforcement and implementation in both physical and virtual spaces reflect the risk of abuse by law enforcement through selective ‘over inclusive’ application - for instance, by ignoring egregious violations of the law while at the same time targeting speech that does not reasonably fall within the application of the law. The law itself contains ambiguous terminology and sets a weak interpretive framework, which allows for the over-inclusive’ application.

The study also contemplates alternative non-legal frameworks of regulation that can complement or substitute for the more formal regulatory mechanisms. The discussion on alternative methods assesses three frameworks for the non-legal regulation of social media: (i) voluntary self-regulation by social media companies, (ii) independent international regulation, and (iii) counter-messaging by civil society and private institution initiatives. This study finds that a combination of the three alternative non-legal frameworks could be a way to effectively address hate speech on social media in Sri Lanka. Accordingly, the study provides summary conclusions and observations of eight key options to address hate speech and disinformation in Sri Lanka through formal and alternative frameworks.

1. FORMAL REGULATORY FRAMEWORK

Brief introduction to hate speech and disinformation

There is no universally accepted definition of hate speech.⁶ A common understanding of ‘hate speech’ might simply be ‘the advocacy of hatred’. Not all speech that is harmful or offensive should be restricted through criminal sanctions. However, article 20(2) of the International Covenant on Civil and Political Rights (ICCPR) stipulates that a particular type of harmful speech, i.e., speech that advocates national, racial or religious hatred, and which incites discrimination, hostility or violence, must be prohibited by law. Many of the specific terms found in article 20(2) of the ICCPR are not defined in international law. However, non-binding “soft law” instruments such as the Camden Principles on Freedom of Expression and Equality offer guidance on interpreting the specific terms contained in article 20(2):⁷

- ‘Hatred’, ‘Hostility’ – intense and irrational emotions of opprobrium, enmity, and detestation towards the target group.
- ‘Advocacy’ – requiring an *intention* to promote hatred publicly towards the target group.
- ‘Incitement’ – statements (about national, racial or religious groups) that create an *imminent risk* of discrimination, hostility, or violence (against persons belonging to those groups).

The Rabat Plan of Action proposes a six-part test for identifying hate speech that must be prohibited.⁸ This threshold test involves the evaluation of six key factors: (i) the social and political context in which the expression is made; (ii) the speaker, for example his or her status and influence; (iii) the intent of the expression; (iv) the content or form of the expression; (v) the extent or

6. Dr. Beata Martin-Rozumiłowicz and Rasło Kužel, *Social Media, Disinformation and Electoral Integrity: IFES Working Paper* (International Foundation for Electoral Systems 2019), p. 10.

7. Principle 12, Camden Principles on Freedom of Expression and Equality (ARTICLE 19, 2009) (‘Camden Principles’).

8. Human Rights Council, *Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred* (‘Rabat Plan of Action’) 11 January 2013, A/HRC/22/17/Add.4, para.29.

reach of the expression; and (vi) the likelihood that the expression would cause harm.

Disinformation is broadly defined as false information that is knowingly or purposefully shared or disseminated with the intention of misleading or causing harm.⁹ If disinformation is part of the advocacy of national, racial or religious hatred and constitutes incitement of discrimination, hostility or violence, it would fall within the scope of article 20(2) of the ICCPR, and must be prohibited.

1.1 THE INTERNATIONAL LEGAL FRAMEWORK ON THE FREEDOM OF EXPRESSION AND THE PROHIBITION OF THE ADVOCACY OF HATRED

The freedom of expression is guaranteed under international human rights law (IHRL).¹⁰ Article 19 of the ICCPR recognises the freedom of opinion and expression.¹¹ The freedom of opinion and expression is also articulated in the Universal Declaration of Human Rights (UDHR). Although not a legally binding instrument, the UDHR is foundational to IHRL, and has influenced the framing of subsequent treaties on human rights including the ICCPR.¹² Sri Lanka ratified the ICCPR in 1980. By virtue of this ratification, Sri Lanka is legally obligated to respect, protect and promote the rights in the ICCPR.

Article 19(2) of the ICCPR guarantees the right to the freedom of expression, which includes the freedom to seek, receive and impart information and ideas of all kinds, regardless of the mode of communication. The ICCPR also states that restrictions on the freedom of expression can be imposed if they are provided for by law, and are necessary (a) to respect the rights or

reputations of others, (b) for the protection of national security or of public order (ordre public), or of public health or morals.¹³ This means that, first, restrictions on the freedom of expression must be imposed by laws that are precise and accessible to the public.¹⁴ Second, any restriction must be narrowly defined and must respond to a pressing social need. Furthermore, the restriction should be proportionate, meaning that the benefit of restricting the expression must outweigh the associated harm caused to the individual whose expression is being restricted.¹⁵ IHRL acknowledges that these standards and principles equally apply to expressions made online.¹⁶

As mentioned above, article 20(2) of the ICCPR states that any advocacy of national, racial, or religious hatred that incites discrimination, hostility or violence shall be prohibited by law.¹⁷ The International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) also refers to the freedom of expression and racist superiority. It obliges states parties to prohibit the ‘dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin, and also the provision of any assistance to racist activities, including the financing thereof’.

1.2 THE DOMESTIC LEGAL FRAMEWORK

Constitutional provisions on the freedom of speech and expression

The Constitution of Sri Lanka guarantees every citizen the freedom of speech and expression, including publication.¹⁸ This right can be restricted, as prescribed

9. Claire Wardle & Hosein Derakhshan, *Information Disorder: Toward an interdisciplinary framework for research and policy making* (Council of Europe Report 2017), p. 5; Christina Nembr & William Gangware, *Weapons of Mass Distraction: Foreign State-Sponsored Disinformation in the Digital Age* (Park Advisors March 2019), p. 4.

10. Article 19, UN General Assembly, Universal Declaration on Human Rights (UDHR) (10 December 1948); Article 19, UN General Assembly, International Covenant on Civil and Political Rights (ICCPR) (16 December 1966) 999 UNTS 171 (entered into force 23 March 1967); Camden Principles.

11. Article 19, ICCPR.

12. Jack Donnelly, *Universal Human Rights: In Theory and Practice* (3rd edn, 2013), p. 26.

13. Article 19(3), ICCPR.

14. United Nations Human Rights Committee, General Comment No. 34 – Article 19 (Freedom of Opinion and Expression) (‘HRC General Comment 34’), 12 September 2011, CCPR/C/GC/34, para. 25.

15. Rabat Plan of Action, p. 9.

16. HRC General Comment 34, para. 43.

17. Article 20(2), ICCPR. Note: Article 4(a) of the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) (21 December 1965) 660 UNTS 195 (entered into force 4 January 1969) also refers to racist expression. It obliges states parties to ‘declare an offence punishable by law all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin, and also the provision of any assistance to racist activities, including the financing thereof’.

18. Article 14(1)(a), The Constitution of the Democratic Socialist

by law in the interests of racial and religious harmony, or in relation to parliamentary privilege, contempt of court, defamation or incitement to an offence.¹⁹ The freedom of expression can also be restricted in the interests of national security, public order and the protection of public health or morality, or for the purpose of securing due recognition and respect for the rights and freedoms of others, or for meeting the just requirements of the general welfare of a democratic society.²⁰

Statutory laws relevant in combatting hate speech and disinformation

Several laws could be applied in combatting hate speech and disinformation in Sri Lanka.²¹

These include:

1. International Covenant on Civil and Political Rights Act, No. 56 of 2007
2. Penal Code, No. 2 of 1883
3. Police Ordinance, No. 16 of 1865
4. Public Security Ordinance, No. 25 of 1947
5. Prevention of Terrorism (Temporary Provisions) Act (PTA), No. 48 of 1979
6. Sri Lanka Telecommunications Act (SLTA), No. 25 of 1991
7. Computer Crimes Act, No. 24 of 2007

The following section identifies each law and discusses the substantive legal provisions, related regulations, applicable case law (if any), and recent examples of the implementation of these laws. It also identifies the

Republic of Sri Lanka 1978.

19. Article 15(2), The Constitution of the Democratic Socialist Republic of Sri Lanka 1978.

20. Article 15(7), The Constitution of the Democratic Socialist Republic of Sri Lanka 1978.

21. Although defamation laws are used in combatting hate speech in other jurisdictions, judicial interpretation and application of defamation laws in Sri Lanka has been limited to personal injury, and limited publicly accessible information is available on whether it has been applied in the larger context of hate speech and disinformation that incite violence. Presently defamation is solely tried in courts as a private dispute under Roman Dutch common law principles that recognise civil responsibilities. As this section looks at relevant statutory laws, defamation is not considered in this report. See Sulakshana Senanayake, 'Freedom Of Speech And Defamation In Sri Lanka: Where To Draw The Line', Roar Media, 21 March 2017, at <https://roar.media/english/life/reports/freedom-of-speech-and-defamation-in-sri-lanka-where-to-draw-the-line> [last accessed 13 October 2020].

institutions involved in implementation and enforcement and analyses the relevance of these institutions to combatting hate speech and disinformation. This section will also briefly discuss other regulatory mechanisms that are available and fall within the 'formal' regulatory framework.

1.2.1 International Covenant on Civil and Political Rights (ICCPR) Act, No. 56 of 2007

On 16 November 2007, Sri Lanka enacted the ICCPR Act to give effect to certain articles of the ICCPR relating to human rights that had not been given domestic recognition by law.²²

Section 3 of the Act gives effect to article 20 of the ICCPR, and criminalises the propagation of war or the advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence. Furthermore, any person who attempts or threatens to commit or aids or abets the commission of this offence is also guilty of an offence. The punishment for the offence is rigorous imprisonment for a maximum of ten years.

The legal provision is problematic as it groups different types of incitement into one offence. In fact, the Rabat Plan of Action states that only the most severe forms of incitement must be criminalised.²³ Penal offences should be precisely defined to avoid ambiguity and ensure proportionality in the imposition of penalties.²⁴ An overbroad offence may be misapplied, thereby criminalising speech that does not necessarily warrant penal sanctions. For instance, under these provisions of the Act, the state may treat merely controversial or dissenting expression as 'hate speech' warranting sanctions.

In relation to implementation and enforcement, the regular criminal procedure would apply.²⁵ The offence is 'cognisable and non-bailable', and is punishable on

22. For example, section 4 gives effect to elements of the right to a fair trial and section 5 provides for the rights of a child.

23. Rabat Plan of Action, para. 34.

24. Verité Research, *Ambiguity and Abuse: The Law on Hate Speech in Sri Lanka Briefing Note* (October 2019), p. 3.

25. The law that governs criminal procedure in Sri Lanka is the Criminal Procedure Act, No. 15 of 1979.

conviction by the High Court.²⁶ Bail can only be granted by the High Court under exceptional circumstances. Of importance is Section 3(5) of the Act, which provides that a trial in the High Court for an offence committed under section 3 must be taken up *before any other business of court* and must be held on a *day to day basis*. Postponement is only permitted owing to unavoidable circumstances. This provision highlights the priority that needs to be given to offences committed under section 3 of the Act.

Recently, the Human Rights Commission of Sri Lanka (HRCSL), in a letter to the Acting Inspector General of Police (IGP), noted the general reluctance of law enforcement authorities to enforce the ICCPR Act and take decisive action to mitigate ethno-religious violence.²⁷ The HRCSL further proposed the adoption of certain guidelines based predominantly on the Rabat Plan of Action and requests law enforcement authorities to follow the six-part test in its enforcement and implementation of the Act.

There are two observations relating to the implementation of the ICCPR Act: (i) the absence of convictions or inadequate prosecutions under the ICCPR Act; and (ii) the overinclusive application of the ICCPR Act which are illustrated below via selected case stories.

(i) Absence of convictions or inadequate prosecutions under the ICCPR Act

Since its enactment, the ICCPR Act has not been enforced to secure a single conviction of perpetrators responsible for inciting anti-minority violence.²⁸ (See Case Stories A, B, C, D and E for examples).

26. As per section 2 of the Code of Criminal Procedure Act, a 'cognizable' offence is an offence for which a peace officer can arrest a person without a warrant. As per section 2, 'peace officer' includes a police officer, and a Divisional Assistant Government Agent and a Grama Seva Niladhari who have been appointed by a Government Agent in writing to perform police duties.

27. Human Rights Commission of Sri Lanka, 'Letter to the Acting Inspector General of Police', 23 May 2019, at <https://srilankabrief.org/wp-content/uploads/2019/09/HRCSL-letter-to-Acting-Inspector-General-of-Police.pdf> [last accessed 30 March 2020].

28. See *Confronting Accountability for Hate Speech in Sri Lanka* (Centre for Policy Alternatives (CPA) 2018) at p. 20 where the authors state that there have been no reported judgments or trials concluded under the section.

Case Story A: Aluthgama, Dharga Town, Valipanna and Beruwala Violence in 2014

Since 2012, with the formation of the far-right Buddhist group, the Bodu Bala Sena (BBS), sporadic violence against the Muslim community began to increase.²⁹ In June 2014, in Aluthgama and the surrounding areas, anti-Muslim violence and riots broke out and continued for over two days. Even prior to the violence, anti-Muslim rhetoric was being spread via social media. The background to the violence was a clash between a Buddhist monk and three Muslim youth who allegedly assaulted the monk, creating a tense situation in the area. In light of the already prevailing anti-Muslim sentiment, the BBS organised a rally in the area. At the BBS rally, General Secretary of the BBS, Ven. Galagoda Aththe Gnanasara Thera delivered an inflammatory speech targeting Muslims.³⁰ Following the rally, violence erupted, and mobs attacked mosques and Muslim-owned properties in the area.³¹

Initially, the violent activities that followed the BBS rally were groups of Muslims and Sinhalese throwing stones at each other. However, following the declaration of a curfew by the police that same day, mob attacks began in areas surrounding Aluthgama, affecting both Muslim-owned businesses and homes, and some Sinhala-owned houses.³² Attacks continued throughout the night to the next day when more attacks took place.³³ The violence reportedly resulted

29. Bodu Bala Sena (BBS or 'Buddhist Power Force) mandate as per the group's official website is the, 'protection of Buddhism for future generations'. See, Website of Bodu Bala Sena, 'Goals and Objectives of Bodu Bala Sena', at <http://www.bodubalasena.org/> [last accessed 26 June 2020]. The group has been identified as a 'Buddhist Nationalist' group and have been reported to allegedly propagate extremism and racial hatred towards other communities. See, 'Sri Lanka on alert as Buddhist monks meet over presidential vote', *Aljazeera*, 07 July 2019, at <https://www.aljazeera.com/news/2019/07/sri-lankan-alert-buddhist-monks-meet-presidential-vote-190707105833379.html> [last accessed 26 June 2020].

30. 'Fear, shock among Sri Lankan Muslims in aftermath of Buddhist mob violence', *CNN World*, 24 June 2014, at <https://edition.cnn.com/2014/06/19/world/asia/sri-lanka-muslim-aluthgama/> [last accessed 11 July 2020].

31. *ibid.*

32. Farzana Haniffa, Harini Amarasuriya and Vishakha Wijenayake, *Where Have All the Neighbours Gone? Aluthgama Riots and its Aftermath: A Fact-Finding Mission to Aluthgama, Dharga Town, Valipanna and Beruwala* (Law & Society Trust 2014) at <https://sri.lanka.org/wp-content/uploads/2017/11/aluthgama-report-final-english.pdf> [last accessed 10 June 2020], p. 4.

33. *ibid.*

in the death of three people, and the destruction of approximately 152 homes and businesses.³⁴

Reportedly no persons have been convicted under the ICCPR Act for the instigation of violence that took place in June 2014.³⁵

Case Story B: Anti-Muslim Violence in Gintota in 2017

The violence in Gintota in November 2017 was reportedly due to an incident that took place the night before the violence. A roadside accident escalated into a minor altercation which then escalated to violence. Following the roadside accident, it was rumoured that a local Muslim politician led a mob to attack Sinhalese residences in Gintota.³⁶ Following this rumour, several Muslim homes, properties, and businesses were attacked. It was reported that four persons were injured during these attacks.³⁷ The government deployed the regular police, the Special Task Force (of the police), and the Tri-Forces to the area.³⁸ Curfew was imposed in the area, and was subsequently extended to the neighbouring areas of Kurunduwatta, Maha Hapugala, Welipitmodara, Ukwatta and Piyadigama.³⁹

News reports suggest that the ICCPR Act was used to initially arrest 22 persons allegedly involved in the violence.⁴⁰ However, there is no information on the actual prosecution of any perpetrators under the Act.

34. Ministry of Foreign Affairs, *Right of Reply Sri Lanka 26th Session of the Human Rights Council Agenda Item 4*, at https://www.mfa.gov.lk/images/stories/pdfs/ROR_on_Aluthgama_incident.pdf (last accessed on 19 August 2020).

35. Gehan Gunatilleke, *The Chronic and the Entrenched: Ethno-religious Violence in Sri Lanka* (ICES 2018). For an account of the incident see, *Where have all the neighbours gone?* (Law & Society Trust 2014), *op. cit.*, p. 1-6.

36. *The Chronic and the Entrenched: Ethno-religious Violence in Sri Lanka* (ICES 2018), *op. cit.*

37. 'Gintota unrest: 19 suspects remanded, curfew again today', *ColomboPage*, 18 November 2017, at http://www.colombopage.com/archive_17B/Nov18_1511026781CH.php [last accessed 26 June 2019].

38. 'Gintota violence: PM warns troublemakers', *Daily News*, 20 November 2017, at <https://www.dailynews.lk/2017/11/20/local/134986/gintota-violence-pm-warns-troublemakers> [last accessed 26 June 2020].

39. *ibid.*

40. 'Suspects charged under Int'l Covenant on Civil and Political Rights Act', *The Island*, 3 December 2017, at http://www.island.lk/index.php?page_cat=article-details&page=article-details&code_title=176019 [last accessed 28 March 2020].

Case Story C: Attacks in Ampara in 2018

In February 2018, inter-communal tensions arose in Ampara after a Sinhalese customer claimed that a Muslim-owned restaurant had mixed 'sterilisation pills' in his food.⁴¹ Video footage of the alleged incident was widely shared on Facebook. The video falsely portrayed the owner of the restaurant supposedly admitting to the act by nodding his head when confronted by the customer. The video was then widely spread on social media and was essentially used to confirm the fear of an alleged Muslim plot to 'sterilise' the Sinhala-Buddhist community.⁴² The restaurant owner later claimed that he does not understand Sinhala and was nodding out of fear.⁴³ Following the video of the alleged confession, the restaurant in question was attacked. The news spread and a mosque and several vehicles were also attacked the following day.⁴⁴

Subsequently, the report of the Government Analyst revealed that the alleged 'sterilisation pill' was actually clumps of flour.⁴⁵ Reportedly, no arrests were made under the ICCPR Act in relation to the incident.

Case Story D: Attacks in Digana and Teldeniya in 2018

Anti-Muslim attacks erupted in Digana and Teldeniya during the first week of March 2018. The violence was reportedly triggered as a reprisal for an assault of a Sinhalese man by four Muslims after a traffic incident. The assault allegedly took place a week prior to the violence.⁴⁶ Mobs attacked Muslim-owned businesses, and places of worship. The government deployed

41. 'Ampara tense following attack on shop and mosque', *Sunday Leader*, 28 February 2018, at <http://www.thesundayleader.lk/2018/02/27/ampara-tense-following-attack-on-shop-and-mosque/> [last accessed 15 June 2020].

42. 'Tension in Ampara after fake 'sterilization pills' controversy', *Sunday Observer*, 4 March 2018, at <http://www.sundayobserver.lk/2018/03/04/news/tension-ampara-after-fake-%E2%80%98sterilization-pills%E2%80%99-controversy> [last accessed 10 July 2020].

43. *ibid.*

44. *ibid.*

45. 'Sterilization Pills' Found To Be Clumps Of Flour: Government Analyst Rules Out Rumours Of Attempted Ethnic Cleansing', *Asian Mirror*, 8 March 2018, at <https://asianmirror.lk/news/item/27037-sterilization-pills-found-to-be-clumps-of-flour-government-analyst-rules-out-rumours-of-attempted-ethnic-cleansing> [last accessed 10 July 2020].

46. 'Digana turns divisive!', *Daily FT*, 6 March 2018, at <http://www.ft.lk/top-story/Digana-turns-divisive-/26-650661> [last accessed 15 June 2020].

approximately 200 Special Task Force (STF) personnel and declared a curfew in response to the attacks.

According to news reports, one of the main suspects involved in instigating the violence, Mahason Balakaya leader Amith Weerasinghe, and nine others were arrested for the violence.⁴⁷ It was also reported that several others were arrested in Teldeniya in connection with the violence.⁴⁸ It is, however, unclear what law they were arrested under. Reportedly, no arrests were made under the ICCPR Act.⁴⁹

Law enforcement officials have been described as being 'complicit in attacks by virtue of omission'.⁵⁰ For instance, STF has been accused of failing to intervene in some cases of violence and looting.

Case Story E: Violence subsequent to the 2019 Easter Sunday Attacks in Chilaw, Gampaha and Kurunegala Districts

Following the Easter Sunday Attacks in April 2019, which were attributed to National Thowheeth Jamaath, an Islamist militant group, several incidents of violence took place targeting the Muslim community in the country. The first spate of violence took place in Chilaw when several mosques and Muslim-owned stores were pelted with stones. Subsequently, mobs attacked several villages in the Kurunegala district in the North-Western Province, destroying houses, property, and mosques. A Police curfew, which was initially imposed in the area, was later extended to the whole country.⁵¹

47. 'The Digana-Kandy Racial Riots: What You Need To Know', *Roar Media*, 07 March 2018, at <https://roar.media/english/life/in-the-know/the-digana-kandy-racial-riots-what-you-need-to-know> [last accessed 15 June 2020].

48. 'Digana turns divisive!', *op. cit.*

49. See 'Abuse of ICCPR Act has 'chilling effect' on fundamental freedoms', *Sunday Observer*, 16 June 2019, at <http://www.sundayobserver.lk/2019/06/19/news-features/abuse-iccpr-act-has-%E2%80%98chilling-effect%E2%80%99-fundamental-freedoms> [last accessed 21 April 2020]; 'Kandy communal violence: Main suspect arrested', *Daily News*, 9 March 2018, at <http://www.dailynews.lk/2018/03/09/local/145064/kandy-communal-violence-main-suspect-arrested> [last accessed 21 April 2020].

50. *Where have all the neighbours gone?* (Law and Society Trust 2014), *op. cit.*, p. 70. See Gehan Gunatilleke, *Ethno-religious Conflict Resolution Mechanisms in Sri Lanka* (The Asia Foundation, 2018), p. 28-29.

51. See 'Police impose island-side curfew till Tuesday morning amid fears of escalation in violence', *Sri Lanka Brief*, 13 May 2019, at <https://srilankabrief.org/2019/05/mobs-attack-mosques-muslim-owned-shops-and-homes-in-sri-lankas-kurunegala-district/> [last accessed 17 April 2020].

Over 5,500 policemen were deployed to the North-Western Province to contain the violence.⁵² Thereafter, violence erupted in the Gampaha district as well.⁵³ According to news reports, damage caused as a result of the violence included property, businesses, religious places, and vehicles.⁵⁴

Mahason Balakaya Leader Amith Weerasinghe was arrested in connection with the violence on the grounds of 'breaching the peace' and was reportedly investigated under the provisions of the Public Security Ordinance and Emergency Regulations.⁵⁵ The director of the Nawa Sinhale National Organisation, Dan Priyasad, was also arrested in connection with the violence.⁵⁶

Moreover, it was reported that suspects were arrested under the ICCPR Act subsequent to the violence in the Minuwangoda and Gampaha areas.⁵⁷ However, there is a lack of information on whether these perpetrators were actually prosecuted. Inaction by law enforcement officials was visible during the incidents of violence that took place in the aftermath of the Easter Sunday attacks.⁵⁸

(ii) Overinclusive application of the ICCPR Act

The ICCPR Act has been applied in an overinclusive manner in the recent past (see Case Stories F, G and H for examples).

52. 'Police remand 78 involved in violence', *Daily FT*, 16 May 2019, at <http://www.ft.lk/news/Police-remand-78-involved-in-violence/56-678230> [last accessed 15 June 2020].

53. See 'Police impose island-side curfew till Tuesday morning amid fears of escalation in violence', *op. cit.* See also, 'Dozens arrested after anti-Muslim violence in Sri Lanka', *The Hindu*, 14 May 2019, at <https://www.thehindu.com/news/international/one-killed-in-anti-muslim-riots-in-sri-lanka/article27123889.ece> [last accessed 26 June 2020].

54. 'Anti-Muslim attacks: Process of compensation causes trickle of unease', *Daily Mirror*, 30 May 2019, at <http://www.dailymirror.lk/news-features/Anti-Muslim-attacks:-Process-of-compensation-causes-trickle-of-unease/131-168220> [last accessed 26 June 2020].

55. 'Leader of the Mahason Balakaya Amith Weerasinghe further remanded', *News 1st*, 28 May 2019, at <https://www.newsfirst.lk/2019/05/28/leader-of-the-mahason-balakaya-amith-weerasinghe-further-remanded/> [last accessed 26 June 2020].

56. 'Police remand 78 involved in violence', *op. cit.*

57. '13 arrested for riots in Minuwangoda to be produced in court today', *ColomboPage*, 14 May 2019, at http://www.colombopage.com/archive_19A/May14_1557817003CH.php [last accessed 17 April 2020].

58. 'Methodist Church Bishop Perera says police yet to act against church attacks', *Sunday Times*, 20 April 2019, at <http://www.sundaytimes.lk/article/1084560/methodist-church-bishop-perera-says-police-yet-to-act-against-church-attacks> [last accessed 13 June 2020].

Case Story F: Arrest of the author and poet Shakthika Sathkumara, April 2019

Author and poet Shakthika Sathkumara wrote a short story 'Ardha', which he posted on Facebook. The fictional short story alluded to sexual abuse involving a member of the Buddhist clergy.⁵⁹ Sathkumara was arrested following a complaint filed by a group of monks from the Buddhist Information Centre. This group considered the short story as 'derogatory and defamatory to Buddhism' and as having 'insulted the life of the Buddha and the Maha Sangha'.⁶⁰

Sathkumara was arrested under section 3 of the ICCPR Act and section 291 of the Penal Code, and held in remand for 130 days.⁶¹ Social media activism in support of Sathkumara involved campaigns such as '#FreeShakthika', and public appeals that were sent to officials requesting they discontinue the investigations.⁶²

Case Story G: Arrest of Abdul Raheem Masaheena, May 2019

Subsequent to the Easter Sunday Attacks, the Muslim community faced backlash, and increased scrutiny. In this context, a Muslim woman was arrested under section 3 of the ICCPR Act for wearing a dress decorated with the logo of a ship's helm on it. The police arrested her because the logo resembled a

sacred Buddhist symbol ('Dharmachakraya').⁶³ Social media commentary involved the demand of justice on behalf of the individual, who was wrongly detained.⁶⁴

Case Story H: Arrest of Ramzy Razeek, April 2020

Ramzy Razeek, a retired government officer in Kandy, was reportedly arrested under the ICCPR Act and Computer Crimes Act on 9 April 2020, for posting certain opinions on his private Facebook account.⁶⁵ Razeek had reportedly emphasised the importance of an 'ideological jihad' in the context of hate propagated against Muslims through a Facebook post. Following this post, Razeek had received several death threats. In response to these threats, on 9 April, Razeek had filed an online complaint with the police. However, this complaint led to the police arresting him under the above mentioned laws.⁶⁶ Many local and international advocacy groups, including Amnesty International, have called for his release.⁶⁷

Upon closer scrutiny of the ICCPR Act, there is nothing to suggest that section 3 cannot be used to regulate expressions, including disinformation, that advocates national, religious, or racial hatred constituting

59. For description of the incident see, 'Abuse of ICCPR Act has 'chilling effect' on fundamental freedoms', *op. cit.*

60. 'Arrest of writer Sathkumara sparks debate on freedom of expression', *Daily Mirror*, 12 April 2019, at <http://www.dailymirror.lk/news-features/Arrest-of-writer-Sathkumara-sparks-debate-on-freedom-of-expression/131-165392> [last accessed 17 April 2020].

61. 'Sri Lankan authorities delay on whether to prosecute award-winning writer Shakthika Sathkumara' *World Socialist Web Site*, 16 December 2019, at <https://www.wsws.org/en/articles/2019/12/16/sath-d16.html> [last accessed 17 April 2020].

62. 'Shakthika has been bailed. But he is not yet free', *Sri Lanka Campaign for Peace & Justice*, 05 August 2019, at <https://www.srilankacampaign.org/shakthika-has-been-bailed-but-he-is-not-yet-free/> [last accessed 26 June 2020]; Indi Samarajiva, 'The Law Is Not Fair', *Medium*, 19 June 2019, at <https://medium.com/@indica/the-law-is-not-fair-78210cc34fc4> [last accessed 26 June 2020]; 'Sri Lanka: Drop investigation into award-winning writer Shakthika Sathkumara', *Pen International Site*, 8 January 2020, at <https://pen-international.org/news/sri-lanka-renewed-calls-to-drop-investigation-into-award-winning-writer-shakthika-sathkumara> [last accessed 26 June 2020]; Facebook Page of Hashtag Generation, 'Public Petition Demanding Immediate Release of Author Shakthika Kumara', 06 April 2019, at https://docs.google.com/forms/d/e/1FAIpQLSegVZAhcqKjHaOdOZHamaOl-JgkXAjNN8gi6e7pigM-5ulnQg/viewform?fbclid=IwAR1NklitNNbF-cfjhAa7O7cOezfQjA4Y9NYGrKD3RZjYT7m_n09X05Zqq4Q&usp=embed_facebook [last accessed 26 June 2020].

63. For description of the incident see, 'Abuse of ICCPR Act has 'chilling effect' on fundamental freedoms', *op. cit.*

64. See Facebook post, 'Abdul Raheem Masaheena, the woman who got arrested for wearing a dress had a ship wheel printed on it, filed a fundamental rights petition against Hasalaka Police. #lka #SriLanka #Human Rights', 18 June 2019, at <https://www.facebook.com/azeem.salam.5/posts/10217114348723715> [last accessed 26 June 2020]; See also, Twitter post, 'Abdul Raheem Masaheena, the women who got arrested for wearing a dress had a ship wheel printed on it, filed a fundamental rights petition against Hasalaka Police yesterday. #lka #SriLanka #HumanRights', at <https://twitter.com/Welikumbura/status/1139826153315885057?s=20> [last accessed 26 June 2020].

65. Damith Chandimal and Ruki Fernando, 'Sri Lanka: The Truth About the Arrest and Detention of Ramzy Razeek', *Sri Lanka Brief*, 14 April 2020, at <https://srilankabrief.org/2020/04/sri-lanka-the-truth-about-the-arrest-and-detention-of-ramzy-razeek/> [last accessed 17 July 2020].

66. *ibid.*

67. Amnesty International, 'SRI LANKA: HEALTH CONCERNS FOR DETAINED BLOGGER: RAMZY RAZEEL', 20 May 2020 at <https://www.amnesty.org/en/documents/asa37/2357/2020/en/> [last accessed 17 July 2020]; Website of Change.org, 'Free Ramzy Razeek and others imprisoned for free expression in Sri Lanka', at https://www.change.org/p/inspector-general-of-police-sri-lanka-free-ramzy-razeek-and-others-imprisoned-for-free-expression-in-sri-lanka?recruiter=1044921961&utm_source=share_petition&utm_medium=twitter&utm_campaign=psf_combo_share_abi&utm_term=share_petition&recruited_by_id=490871e0-5728-11ea-ac5d-3fa66276688d [last accessed 17 July 2020]; Website of Sri Lanka Campaign for Peace and Justice, 'Ramzy Razeek release #FreeRamzyRazeek', at <https://www.srilankacampaign.org/threatened-with-death-then-arrested-still-detained-help-freeramzyrazeek/ramzy-razeek-infographic-min/> [last accessed 17 July 2020].

incitement to discrimination, violence and hostility on social media. The issue, however, lies in how section 3 is framed. On the one hand, its scope is quite broad, and many key terms, such as ‘hostility’ remain undefined. Such breadth makes section 3 vulnerable to abuse. On the other hand, it conflates several types of speech offences, such as, incitement to violence and incitement to discrimination, and potentially sets the same penalty for all of the offences (i.e. a maximum of ten years rigorous imprisonment). Such conflation is problematic, as it mixes very harmful forms of hate speech (i.e. incitement to violence) with comparatively less harmful forms of hate speech (e.g. calls for boycotts of Muslim businesses). While both types of offences ought to be dealt with, imposing a mandatory penalty of incarceration for lesser offences is disproportionate.

1.2.2 Penal Code, No. 2 of 1883

Chapter XV of the Penal Code deals with offences relating to religion. Two specific provisions seek to restrict speech that ‘wounds’ the religious feelings of others, provided that malicious and deliberate intention is proved.

Section 291A provides that uttering words deliberately intended to wound religious feelings is an offence punishable with imprisonment for a term up to one year, or a fine, or both.⁶⁸ Section 291B provides that *deliberate and malicious* acts intended to outrage the religious feelings of any class, by insulting its religion or religious beliefs, is a punishable offence with imprisonment for a term up to two years, or a fine, or both.⁶⁹ It is argued that both these provisions can be interpreted to cover expressions that may advocate religious hatred constituting incitement to discrimination, hostility, or violence.⁷⁰ However, there is a lack of jurisprudence

on sections 291A and 291B of the Penal Code.⁷¹ The Indian Penal Code has similar provisions to the penal offences in Sri Lanka. Jurisprudence in India elicits that the provisions must be interpreted strictly. For instance, in *Narayan Das and Another vs. State*, the Indian High Court made it clear that a ‘deliberate intention’ had to be proved, as opposed to mere ‘knowledge’.⁷² This approach could be followed in Sri Lanka when interpreting its penal provisions to prevent undue restrictions to free speech.⁷³

With respect to section 291B specifically, the Supreme Court referred to this section in the *Naomi Michelle Cokeman* fundamental rights case which is explained under Case Story I.⁷⁴

*Case Story I: The Naomi Michelle Cokeman Case, 2017*⁷⁵

The Petitioner, Naomi Cokeman, was a British citizen who visited Sri Lanka in 2014. She had a tattoo of Lord Buddha on her arm ‘as an apt tribute to and as a personal expression of her devotion to Buddhist teachings’.⁷⁶ Upon her arrival to Sri Lanka, she was arbitrarily arrested and detained, subjected to degrading treatment, and subsequently deported. The Police reportedly arrested her for violating section 291B of the Penal Code. She claimed that her fundamental rights under articles 11, 12(1) and 13(1) of the Sri Lankan Constitution had been violated.

The Supreme Court observed that there was no reasonable basis for the arrest. Referring to the report by the police, the police themselves had acknowledged

⁶⁸ [title=166041](#) [last accessed 17 April 2020].

⁷¹ *Confronting Accountability for Hate Speech in Sri Lanka* (CPA 2018), *op. cit.*, p. 30-35.

⁷² See *Confronting Accountability for Hate Speech in Sri Lanka* (CPA 2018), *op. cit.*, at p. 32-34 which cites *Narayan Das and Anr vs. State* (1952) ARI Ori 149. It also cites *Chakra Behera and Ors Balakrushna Mohaptra* AIR 1963 Ori 23 which referred to the threshold for ‘deliberate intention’. The Court in *Shalibhadra Shah & Ors. Vs. Sawmi Krishna Bharati And Anr* 1980 1981 CriLJ 113 stated that the section did not apply to ‘published articles’, and only referred to words said ‘orally’.

⁷³ *ibid.*

⁷⁴ *Naomi Michelle Cokeman v. The Attorney General and Others*, SC (FR) Application 136/2014 (Supreme Court of Sri Lanka) judgment delivered on 15 November 2017, at http://www.supremecourt.lk/images/documents/sc_fr_136_2014.pdf [last accessed 18 April 2020].

⁷⁵ The following information is taken from the judgment.

⁷⁶ *ibid.*, p. 3.

⁶⁸ Section 291A provides ‘Whoever, with the deliberate intention of wounding the religious feelings of any person, utters any word or makes any sound in hearing of that person, or makes any gesture in the sight of that person, or places any object in the sight of that person, shall be punished with imprisonment of either description for a term which may extend to one year, or with fine, or both.’

⁶⁹ Section 291B provides ‘Whoever, with deliberate and malicious intention of outraging the religious feelings of any class of persons, by words, either spoken or written, or by visible representations, insults or attempts to insult the religion or the religious beliefs of that class, shall be punished with imprisonment of either description for a term which may extend to two years, or with fine, or with both.’

⁷⁰ ‘BASL issues severe indictment on IGP, AG, govt over attacks on minorities’, *The Island*, 1 June 2017, at http://www.island.lk/index.php?page_cat=article-details&page=article-details&code_

that the petitioner in the case had no ‘intention’ to outrage religious feelings. The court observed that no case could be made based on section 291B of the Penal Code, as there was no acceptable evidence placed before the court of the ‘possibility of public outcry’. The Court held that the petitioner’s rights under articles 11, 12(1) and 13(1) of the Constitution had been violated, and ordered the State and Respondents in the case to pay compensation to the Petitioner.

While the case was not decided on section 291B of the Penal Code, the Court’s reference to it implies the importance of an ‘intention’ to outrage religious feelings, and thereby provides some clarity on the provision.

Section 120 of the Penal Code on ‘exciting or attempting to excite disaffection’ is of relevance as well. This section provides:

Whoever by words, either spoken or intended to be read or by signs, or by visible representation or otherwise, excites or attempts to excite feelings of disaffection to the State, or excite hatred or contempt of the administration of justice, or excites or attempts to excite the People of Sri Lanka to procure, otherwise than by lawful means, the alteration of any matter by law established or attempts to raise discontent or disaffection, or to promote feelings of ill-will and hostility between different classes of such people, shall be punished with simple imprisonment for a term which may extend to two years.

A few instances in which this section has been used to make arrests have been documented.⁷⁷ The Supreme Court has held that the ‘essence’ of section 120 ‘is whether the words in question incite the people to commit acts of violence and disorder and not whether the words are defamatory or not’.⁷⁸ With respect to

77. See Human Rights Watch, ‘Sri Lanka: Halt Harassment of Media: Websites Blocked, Journalists Face Charges’, 3 July 2017, at <https://www.hrw.org/news/2012/07/03/sri-lanka-halt-harassment-media> [last accessed 6 April 2020] for an instance when the Sri Lanka Mirror was charged for exciting or attempting to excite the feelings of disaffection to the President or the government. See also, ‘Four arrested for circulating harmful SMSs’, *Daily Mirror*, 6 August 2014, at <http://www.dailymirror.lk/50702/tech> [last accessed 17 April 2020] for an instance when four people were arrested for allegedly circulating an SMS with the intent of creating disharmony among communities.

78. *Sisira Kumara Wahalathanthri and Another v. Jayantha Wickra-*

implementation and enforcement, the regular criminal procedure would apply. Available news reports suggest that section 120 has been the most commonly used to stifle expression by journalists and media that evoked anti-government sentiments – similar to the common law offence of sedition.⁷⁹ By contrast, there is a lack of information pertaining to the use of section 291A.

Section 291B was reportedly used in relation to the recent arrest of Shakthika Sathkumara⁸⁰ (see Case Story F under heading 1.2.1 on the ICCPR Act for further details.) However, there is a lack of information on how widely the section has been used to make arrests, or charge perpetrators under the provisions.

Of the provisions referred to above, sections 291B and 120 can be applied in the context of social media. There is no element in the provision itself that seeks to limit the scope of these sections. The problem, however, arises with the lack of clarity on the precise scope and interpretation of specific terms like ‘outrage religious feelings’, ‘discontent or disaffection’ and ‘promote feelings of ill-will’ which feature in these provisions. Such ambiguity can lead to the misapplication and abuse of the law, resulting in a violation of the freedom of expression (see Case Story J and figure 01 for examples).

Case Story J: Spread of false information during COVID-19

An individual was reportedly arrested for using his Facebook account to criticise the appointment of Basil Rajapaksa to the head of the Presidential Task Force on COVID-19.⁸¹ Similarly, an individual was arrested for using Facebook to criticise a Divisional Secretariat for ‘injustices’ that had occurred during the coronavirus

maratne and Others, SC(FR) Application 768/2009 (Supreme Court of Sri Lanka) judgment delivered on 5 November 2015, at www.supremecourt.lk/images/documents/sc_fr_768_2009.pdf [last accessed 17 April 2020], p. 19.

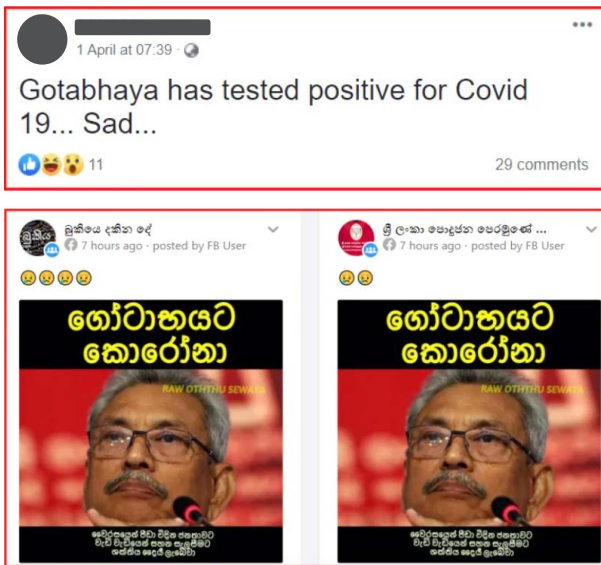
79. See ‘Penal Code is being used fraudulently: JVP’, *Daily Mirror*, 22 February 2012, at http://www.dailymirror.lk/breaking_news/penal-code-is-being-used-fraudulently-jvp/108-17020 [last accessed 29 April 2020]. See also ARTICLE 19, *An Agenda for Change: the right to freedom of expression in Sri Lanka*, 27 October 1994, at <https://www.ref-world.org/publisher,ART19,,LKA,475418a40,0.html> [last accessed 29 April 2020], at p. 26 which refers to a case in 1992 when section 120 was used to ‘silence peaceful criticism’.

80. Referred to in the previous section.

81. ‘Sri Lanka government intensifies crackdown on social media’, *World Socialist Web Site*, 9 April 2020, at <https://www.wsws.org/en/articles/2020/04/09/medi-a09.html> [last accessed 21 April 2020].

eradication and quarantine programme.⁸² Another individual was arrested for reportedly claiming that the president was infected with COVID-19 on Facebook.⁸³ The respective news report on these arrests states that the individuals were arrested under the Computer Crimes Act and the Penal Code. However, it is unclear which provision(s) of the Penal Code were used to make these arrests.

Figure 1: Posts extracted from social media that falsely claimed President Rajapaksa had contracted COVID-19.⁸⁴ The first post was made by an individual, following which it was shared among several groups.



1.2.3 Police Ordinance, No. 16 of 1865

The Police Ordinance vests the police with several powers to control situations where there is a threat to public peace and public order.⁸⁵ Concerning the propagation

82. *ibid.*

83. *ibid.*

84. Website of Fact Crescendo, ‘ජනාධිපති ගෝඨාභය රාජපක්ෂ මහතාට COVID-19 වැළඳී ඇති බවට පළවූ සාවිද්‍යා ප්‍රකාශයක්!’ (Fake news posted on President Gotabaya Rajapaksa being infected by COVID-19), 08 April 2020, at <https://srilanka.factcrescendo.com/2020/04/08/fact-check-sl-president-is-not-infected-with-covid19/> [last accessed 26 June 2020].

85. For example, see Section 79(2): ‘Any person who in any public place or at any public meeting uses threatening, abusive or insulting words or behaviour which is intended to provoke a breach of the peace or whereby a breach of the peace is likely to be occasioned, shall be guilty of an offence under this section.’; Section 56: ‘Every Police Officer shall for all purposes in this Ordinance contained be considered to be always on duty, and shall have the powers of a Police Officer in every part of Sri Lanka, It shall be his duty—(a) to use his best endeavours and ability to prevent all crimes, offences, and public nuisances; (b) to preserve the peace; (c) to apprehend disorderly and suspicious characters; (d) to detect and bring offenders to justice; (e) to collect and com-

of speech or disinformation that may advocate national, religious or racial hatred, section 98 of the Ordinance is of relevance. This section provides that any person who spreads false reports with the view of alarming the inhabitants of any place within Sri Lanka and creating ‘a panic’ will be guilty of an offence.⁸⁶ Furthermore, the provision is broad enough to be interpreted to include false reports on social media. However, the publicly available information about the use of this section to make any arrests or charge perpetrators is limited.

During the outbreak of the COVID-19 pandemic, a letter by the HRCSL to the Acting IGP mentions that section 98 of the Police Ordinance was cited in Police ‘B’ reports as the basis for certain arrests made in relation to the spread of misinformation regarding the pandemic.⁸⁷ This communication by the HRCSL demonstrates that the section has been relied on to make arrests.

1.2.4 Public Security Ordinance, No. 25 of 1947

The Public Security Ordinance (PSO) provides for the enactment of emergency regulations or the adoption of other measures in the interests of public security and the preservation of public order and for the maintenance of supplies and services essential to the life of the community.⁸⁸ Under the PSO, the president has the power to make Emergency Regulations (ERs) if it appears ‘necessary or expedient in the interests of public security and the preservation of public order and the suppression of mutiny, riot or civil commotion, or for the maintenance of supplies and services essential to the life of the community’.⁸⁹ ERs can be issued for several purposes including the detention of people, entry and

municate intelligence affecting the public peace; and (f) promptly to obey and execute all orders and warrants lawfully issued and directed to him by any competent authority.’

86. The punishment is ‘a fine not exceeding two hundred rupees, or to imprisonment, with or without hard labour, for any period not exceeding twelve months; and if he shall be convicted a second time, or shall persist in the offence after warning to desist, he shall be liable to corporal punishment not exceeding twenty lashes.’ Note: A sentence of whipping as a punishment was repealed through section 3 of Corporal Punishment (Repeal) Act, No. 23 of 2005.

87. HRCSL Letter April 2020, *op. cit.*

88. Preamble, Public Security Ordinance, No. 25 of 1947 (PSO).

89. Section 5, PSO.

search, and the acquisition of private property.⁹⁰ An operational ER supersedes any other law, except the Constitution, and cannot be called into question by any court.⁹¹

During past states of emergency, several ERs have been issued.⁹² Most recent examples of ERs have included prohibition on spreading false information which are illustrated through two case stories.

Emergency Regulations under the PSO

Case Story D1: Digana Violence in 2018

Subsequent to the religious violence in Digana in March 2018, former President Maithripala Sirisena issued new ERs.⁹³ According to regulation 15, the communication or spreading of any rumour or false statements or any information or image or message, which was likely to cause public alarm or public disorder, racial violence or incite the committing of an offence, was made an offence.⁹⁴ The provision explicitly refers to modes of communication, which include ‘digital and or social media’.

Case Story E1: 2019 Easter Sunday Attacks

Subsequent to the Easter Sunday Attacks in April 2019, the former President Sirisena declared a state of emergency and issued new ERs. The new ERs were similar in content to the ERs issued in 2018,

with additional provisions to deal with terrorism.⁹⁵ The provision on communication or spreading of rumours and false statements was not as detailed as the 2018 regulation.⁹⁶ It did not specifically mention social media. Instead, it stated ‘by any other means whatsoever’, which is broad enough to cover all forms of communication. Further, regulation 33 made it an offence to ‘print or publish a document recording or giving information or commenting about, or any pictorial representation, photograph or cinematograph film of any matter likely, directly or indirectly to create communal tension’.⁹⁷

Subsequent news reports suggested that any propaganda or misleading information could be penalised under the 2019 ERs.⁹⁸ It was also reported that a special investigation unit had been established to identify and enforce the law against individuals who spread extremist content on social media that incited communal hatred.⁹⁹ Publishing or even sharing such content was reportedly considered an offence for which a person could be sentenced to three to seven years of imprisonment under the ERs.¹⁰⁰ This regulation is overbroad and could have been used to disproportionately regulate content on social media.

Case stories D1 and E1 illustrate that, in the last two years, ERs in response to violence and terrorist attacks have included broad language that can cover content on social media. While this broadness allows for social

90. Section 5, PSO.

91. Sections 7 & 8, PSO.

92. For example: Emergency (Prohibition on Publication and Transmission of Sensitive Military Information) Regulation No. 1 of 1998 published in Gazette Extraordinary No. 1030/28 of 5th June 1998 as amended on 6th June 1999 prohibited the publication of “any publication pertaining to official conduct, morale, the performance of the Head or any member of the Armed Forces or the Police Force or of any person authorised by the Commander - in - Chief of the Armed Forces for the purpose of rendering assistance in the preservation of national security.” In *Sunila Abeyskera v Ariya Rubesinghe and Other* [2000] 1 Sri LR 314, the Court held that the restrictions imposed by the ERs were not disproportionate to the legitimate aim of the regulations (i.e. interest of national security).

93. Emergency Regulations Extraordinary Gazette No 2061/21, issued on 6 March 2018, at http://www.documents.gov.lk/files/egz/2018/3/2061-21_E.pdf [last accessed 17 April 2020].

94. Regulation 15, Emergency Regulations 2018: ‘No person shall, by word of mouth or by any other means whatsoever including through digital and or social media communicate or spread any rumour or false statement or any information or image or message which is likely to cause public alarm or public disorder, racial violence or incite the committing of any offence.’

95. Extraordinary Gazette No. 2120/5, issued on 22 April 2019, at http://www.documents.gov.lk/files/egz/2019/4/2120-05_E.pdf [last accessed 17 April 2020].

96. Regulation 32: ‘No person shall, by word of mouth or by any other means whatsoever, communicate, disseminate or spread any rumour or false statement which is likely to cause public alarm or public disorder.’

97. Regulation 33(e).

98. For example see ‘Desist from fake propaganda activities, warns Defence Ministry’, *Daily News*, 29 April 2019 at <http://www.dailynews.lk/2019/04/29/local/184092/desist-fake-propaganda-activities-warns-defence-ministry> [last accessed 18 April 2020] which reports that Military Spokesman Brigadier Sumith Atapattu stated that “publication of false and misleading stories, images and information and conduct of media provocations by various interested sections/individuals/organizations that could incite/lead to communal passions and disunity among ethnic groups is liable for prosecution under Emergency Regulations, now in force.”

99. Police to establish special unit to penalise social media users who spread content inciting communal hatred’, *Sunday Times*, 13 May 2019, at <http://www.sundaytimes.lk/article/1087840/police-to-establish-special-unit-to-penalise-social-media-users-who-spread-content-inciting-communal-hatred> [last accessed 18 April 2020].

100. *ibid.*

media to be regulated through the legal framework, it can be problematic given Sri Lanka's history of misusing emergency rule and regulations.¹⁰¹

1.2.5 Sri Lanka Telecommunications Act (SLTA), No. 25 of 1991

The Sri Lanka Telecommunications Act, No. 25 of 1991 provides for the establishment of the Telecommunications Regulatory Commission of Sri Lanka (TRCSL), sets out the powers of the Minister in respect of telecommunication, prescribes certain offences and penalties, and generally recognises the power of the government to regulate the use of telecommunication. The Act provides that the TRCSL has the power to take 'regulatory measures as may be prescribed to comply with any general or special directions that may be given to it from time to time by the Government of Sri Lanka in the interest of national security, public order and the defence of the country'.¹⁰²

Section 58 provides that every person who by himself, or through an agent or servant 'tenders for transmission at any telecommunication office any message of an indecent, obscene, seditious, scurrilous, threatening or grossly offensive character' is guilty of an offence.¹⁰³ The provision can be interpreted to cover expression

and disinformation that advocates national, religious or racial hatred constituting incitement to, hostility or violence in particular.

The Act also provides for the power of the government to prohibit or restrict the transmission and reception of telecommunication messages due to the occurrence of 'any public emergency or in the interest of public safety and tranquillity'.¹⁰⁴ This provision can be used to regulate social media in the context of expression and disinformation that advocates national, religious, or racial hatred constituting incitement to discrimination, hostility or violence.

In the context of COVID-19, in March 2020, the TRCSL sent out a message advising the public to refrain from 'misusing all forms of telecommunication services to circulate or share false or fabricated information on COVID-19 to create unnecessary panic among the general public' as it is a 'punishable offence'.¹⁰⁵ However, the provisions upon which the TRCSL formulated this punishable offence remains unclear.

In terms of the implementation and enforcement of the Act, as well as the powers exercised under the Act, there have been instances where the TRCSL has regulated or blocked content for multiple reasons.¹⁰⁶ (See Case Stories K and case stories D2 and E2 for examples).

Case Story K: Blocking of websites 2015-2018

In 2017, the LankaNews website was blocked under the direction of the TRCSL. The block triggered a right to information request by citizenship journalism website Groundviews, who made a request under the Right to Information Act for the TRCSL to disclose information

101. For an overview of emergency rule in Sri Lanka, see Deepika Udagama, 'An Eager Embrace: Emergency Rule and Authoritarianism in Republican Sri Lanka' in Asanga Welikala (ed), *Reforming Sri Lankan Presidentialism: Provenance, Problems and Prospects* (CPA 2015), at <http://www.mediareform.lk/wp-content/uploads/2020/01/Deepika-Udagama-2015-An-Eager-Embrace-Emergency-Rule-and-Authoritarianism-in-Republic-Sri-Lanka.pdf> [last accessed 29 April 2020], p. 286-332. See also International Commission of Jurists, *Sri Lanka: Briefing Paper: Emergency Laws and International Standards* (March 2009), at <https://www.icj.org/wp-content/uploads/2012/05/SriLanka-emergencylaws-advocacy-2009.pdf> [last accessed 21 April 2020].

An example of when ERs have been misapplied, was in the case against Sarath Fonseka, the former Army General. The High Court convicted him under regulation 28 of the 2005 ERs on communicating or spreading any rumour or false statement which was likely to cause public alarm or public disorder. The conviction was based on a statement he allegedly made during an interview with a journalist regarding the allegation that LLTE leaders who were surrendering with white flags were shot on order of the Defence Secretary during the war. For information on this case, see Jayantha de Almeida Guneratne, Kishali Pinto-Jayawardena (ed) & Gehan Gunatilleke, *The Judicial Mind in Sri Lanka; Responding to the Protection of Minority Rights* (Law & Society Trust 2014), at https://www.academia.edu/12987171/The_Judicial_Mind_in_Sri_Lanka_Responding_to_the_Protection_of_Minority_Rights [last accessed 29 April 2020], p. 254-258.

102. Section 5(f), Sri Lanka Telecommunications Act (SLTA), No. 25 of 1991 (SLTA).

103. The penalty for the offence is a fine not exceeding five thousand rupees or imprisonment up to six months, or to a fine and imprisonment.

104. Section 69, SLTA. Note: The Minister must do this by an Order published in the Gazette. Further the Minister can either prohibit or order supervision or control of telecommunication. In the latter instance, the Minister can either prescribe conditions for the supervision or control, or make provision as deemed necessary for the interception or censoring of all or specified messages via telecommunication.

105. 'Refrain from circulating false info on Covid-19: TRCSL', *Daily Mirror*, 13 March 2020, at http://www.dailymirror.lk/breaking_news/Refrain-from-circulating-false-info-on-Covid-19-TRCSL/108-184897 [last accessed 21 April 2020].

106. Center for Policy Alternatives, 'The Internet as a medium for free expression: A Sri Lankan legal perspective', J C Weliamuna, (2013), at <https://www.cpalanka.org/the-internet-as-a-medium-for-free-expression-a-sri-lankan-legal-perspective/> [last accessed 16 March 2021] for case law and legal proceedings on instances of websites that were blocked by TRCSL.

with respect to blocking of several websites, including Lankaenews.¹⁰⁷ In response, the TRCSL revealed that 13 websites (some sites publishing political news and pornographic material) had been blocked since 2015. It also provided information on the process for blocking these websites.¹⁰⁸ However, initially, it did not specify reasons for blocking Lankaenews. Overall, the TRCSL's response revealed that four websites had allegedly been blocked for 'publishing incorrect information and damaging the President's reputation', while two other websites had allegedly been blocked for 'publishing false information'.¹⁰⁹

Case Story D2 and E2: Social media restrictions subsequent to violence in Digana in 2018 and the 2019 Easter Sunday Attacks

A key mechanism through which the former government sought to regulate speech which (including disinformation) advocates national, racial, or religious hatred that incites discrimination, hostility or violence was restricting access to social media. Following the anti-Muslim violence in Digana in 2018, access to certain social media platforms was temporarily restricted to curb the spread of inflammatory information and prevent communal violence. The temporary restriction was enforced by the TRCSL in compliance with a request made by the Ministry of Defence.¹¹⁰ The government claimed that the temporary restriction was enforced as Facebook was being used by 'extremists' to organise and recruit

supporters. It also claimed that several videos of the attacks, hate speech and inflammatory videos had spread at a rapid pace.¹¹¹ The temporary restriction continued even after the violence in the Kandy District had been contained. Similarly, after the Easter Sunday terror attacks in April 2019 the government restricted access to social media to 'prevent the spread of misinformation'.¹¹²

Case story K and case stories D2 and E2 indicate that the previous government used social media restrictions as an almost 'default' position to supposedly curb violence. It appears that the TRCSL assumes broad regulatory powers that extend to social media, and disproportionately uses such powers during crises.

Challenges: Politicisation of the TRCSL

Presently, the trend of moving cyber-security from technical experts to the political sphere is dominant across the globe.¹¹³ A common factor in this trend is the militarisation of cyber-security.¹¹⁴

In Sri Lanka, one of the key issues regarding the TRCSL has been the politicisation of this body. There have been instances when the President's Secretary has also served as the Chairperson of the TRCSL.¹¹⁵ The dual roles occupied by the President's Secretary raise questions over the independence of the TRCSL. In March 2019, the former President Sirisena brought the TRCSL under his control (under the purview of the Ministry of Defence).¹¹⁶ Similarly, subsequent to electing the new President

107. 'Blocked: RTI requests reveal process behind blocking of websites in Sri Lanka', *Groundviews*, 8 December 2017, at <https://groundviews.org/2017/12/08/blocked-rti-requests-reveal-process-behind-blocking-of-websites-in-sri-lanka/> [last accessed 6 April 2020].

108. *ibid.* The process reveals that the Ministry of Parliamentary Reforms and Mass Media issued the final order to the TRC, but the initial order came directly from the Presidential Secretariat, for a few of the websites. In turn, the TC sent a letter to the CEOs of the main internet service providers.

109. *ibid.* Note: Subsequently, the TRC revealed information about the blocking of Lankaenews in 2017. See 'RTI Reveals Lanka E News Blocked On Order from President's Office', *Groundviews*, 11 April 2018, at <https://groundviews.org/2018/04/11/lanka-e-news-blocked-on-order-from-presidents-office-rti-reveals/> [last accessed 6 April 2020].

110. Facebook and other Social Media Networks are temporarily blocked', *News 1st*, 7 March 2018, at <https://www.newsfirst.lk/2018/03/07/measures-taken-monitor-social-media-websites-trc/> [last accessed 6 April 2020]; 'Sri Lanka blocks social media networks to stop sectarian violence', *Reuters*, 7 March 2018, at <https://www.reuters.com/article/sri-lanka-clashes-internet/sri-lanka-blocks-social-media-networks-to-stop-sectarian-violence-idUSL4N1QP39X> [last accessed 6 April 2020].

111. 'TRC to lift social media ban on Friday', *Daily FT*, 14 March 2018, at <http://www.ft.lk/top-story/TRC-to-lift-social-media-ban-on-Friday/26-651218> [last accessed 6 April 2020].

112. 'Sri Lanka blocks social media after deadly Easter explosions', *CCNET*, 21 April 2019, at <https://www.cnet.com/news/sri-lanka-shuts-down-social-media-following-deadly-easter-explosions/> [last accessed 6 April 2020]; 'Sri Lanka attacks: The ban on social media', *BBC News*, 23 April 2019, at <https://www.bbc.com/news/technology-48022530> [last accessed 17 April 2020].

113. Myriam Dunn Cavelty, 'The Militarisation of Cyber Security as a Source of Global Tension' (2012), *STRATEGIC TRENDS ANALYSIS*, Zurich, Möckli, Daniel, Wenger, Andreas, eds., Center for Security Studies, 2012, at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2007043 [last accessed 26 June 2020].

114. *ibid.*

115. 'RTI Reveals Lanka E News Blocked On Order from President's Office', *Groundviews*, 11 April 2018, at <https://groundviews.org/2018/04/11/lanka-e-news-blocked-on-order-from-presidents-office-rti-reveals/> [last accessed 6 April 2020].

116. 'Sri Lanka President takes control of telecommunication authority', *ColomboPage*, 8 March 2019, at http://www.colombopage.com/archive_19A/Mar09_1552151677CH.php [last accessed 6 April 2020].

*Gotabaya Rajapaksa, the Secretary to the Ministry of Defence was appointed as the Chairperson of the TRCSL in December 2019.*¹¹⁷

The TRCSL's mandate and powers extends to the regulation of social media. However, the politicisation of the TRCSL, coupled with its overbroad mandate and powers, could lead to it disproportionately regulating social media.

1.2.6 Prevention of Terrorism (Temporary Provisions) Act (PTA), No. 48 of 1979

The PTA was enacted under the government of President J. R. Jayewardene in 1979. The government 'rushed' the bill through Parliament as an 'urgent bill', which resulted in a lack of opportunity for public comment or opposition.¹¹⁸ The PTA was a special law intended to make temporary provisions to deal with terrorism threats for three years, which was subsequently made permanent.¹¹⁹ A draft bill titled the Counter Terrorism Act was proposed in 2018, in order to repeal and replace the PTA. The bill was under consideration by the Sectorial Oversight Committee on International Relations. On 3 January 2020, the Cabinet approved a decision to withdraw the draft bill.¹²⁰

Two sections of the Act are directly relevant for the purpose of this brief. Section 2(1)(h) makes it an offence for any person to cause or intend to cause the 'commission of acts of violence or religious, racial or communal disharmony or feelings of ill-will or hostility between different communities or racial or religious group', by words, signs, visible representations or otherwise. The Act also penalises persons who abet, conspire, or incite the commission of offences identified

in section 2.¹²¹ This section appears to prohibit certain forms of hate speech, and can be extended to social media platforms as it is not restricted to particular modes of communication.

Furthermore, section 14 provides that, subject to certain conditions,¹²² no person should print or publish in any newspaper any matter relating to 'incitement to violence, or which is likely to cause religious, racial or communal disharmony or feelings of ill-will or hostility between different communities or racial or religious groups' without the approval of a competent authority.¹²³ Any person who contravenes this provision is guilty of an offence.¹²⁴

In terms of implementation and enforcement, the Act provides broad powers for law enforcement officials in relation to the powers of entry, search, seizure, and arrest. The Act also provides for detention and restriction orders. Over the years, most of these provisions have been disproportionately used against several persons including journalists, human rights defenders, and political activists.¹²⁵ For example, the case of *Tissanaiyagam* in 2008 was an instance where section 2(1)(h) was relied on. Tissanaiyagam was an editor of a magazine. He was arrested by the Terrorism Investigation Division for publishing articles that allegedly incited the commission of acts of violence or racial or communal disharmony. In 2009, he was sentenced to 20 years rigorous imprisonment. The case has been criticised for the use of a 'tampered' confession as evidence in the proceedings, and for the failure to prove that the

117. 'Defence Secretary takes office as TRC Chairman', *Daily FT*, 7 December 2019, at <http://www.ft.lk/front-page/Defence-Secretary-takes-office-as-TRC-Chairman/44-691169> [last accessed 6 April 2020].

118. *Reforming Sri Lankan Presidentialism: Provenance, Problems and Prospects* (CPA 2015), *op. cit.* p. 296- 297.

119. Section 29 ('Duration of the Act') of the original Act provided that the Act would operate for 3 years. This section was repealed by the Prevention of Terrorism (Temporary Provision) Amendment Act, No. 10 of 1982.

120. 'Sri Lanka government decides to withdraw Counter Terrorism Bill', *ColomboPage*, 03 January 2020, at http://www.colombopage.com/archive_20A/Jan03_1578067614CH.php [last accessed 26 January 2020].

121. Section 3, Prevention of Terrorism (Temporary Provisions) Act, No. 48 of 1979 (PTA).

122. Upon an order being made in the Gazette by the Minister.

123. Section 14(2)(ii), PTA.

124. Section 14(3), PTA. The punishment is imprisonment up to 5 years. The Court may also order other penalties.

125. 'Sri Lanka's draft Counter Terrorism Act: a license for continued state oppression, intimidation and torture', *Sri Lanka Campaign for Peace & Justice*, 4 May 2017, at <https://www.srilankacampaign.org/sri-lankas-draft-counter-terrorism-act-license-continued-state-oppression-intimidation-torture/> [last accessed 18 April 2020]. See also Human Rights Watch, 'Locked Up Without Evidence: Abuses under Sri Lanka's Prevention of Terrorism Act', 29 January 2018, at <https://www.hrw.org/report/2018/01/29/locked-without-evidence/abuses-under-sri-lankas-prevention-terrorism-act> [last accessed 18 April 2020] which refers to several instances when the provisions of the PTA were misapplied.

respective articles could incite ethnic disharmony.¹²⁶ This case is one such example of the misapplication of the PTA.

Section 2(1)(h) can be applied to the regulation of social media, as it does not limit the modes of communication.

1.2.7 Computer Crimes Act, No. 24 of 2007

The Computer Crimes Act was enacted to identify computer crimes and to provide the procedure for the investigation and prevention of such crimes, and for other connected matters.¹²⁷

Section 6 provides that it is an offence for a person to intentionally cause a computer to perform any function, knowing or having reason to believe that such a function will result in danger or imminent danger to (a) national security, (b) the national economy, or (c) public order.¹²⁸ Accordingly, if any person uses a computer to potentially create, or even share content on social media that advocates national, religious or racial hatred constituting incitement to hostility or violence, such person will be guilty of an offence under this section. In light of the fact that the grounds deal with national security and public order, the section does not necessarily cover incitement to discrimination.

Furthermore, the Act makes it an offence for a person who (without lawful authority) 'produces, sells, procures for use, imports, exports, distributes or otherwise makes available (a) any device, including a computer or computer program; b) a computer password, access code or similar information by which the whole or any part of a computer is capable of being accessed', with the intent of it being used by anyone for the purpose of committing an offence.¹²⁹ This provision could be of relevance if a person makes such a system available for

the dissemination of content on social media that should otherwise be prohibited.

Any attempt to commit the above offences,¹³⁰ abetting in these offences,¹³¹ and conspiring to commit these offences under the Act,¹³² are also made offences by the Act.

In terms of implementation and enforcement, all offences are cognisable and are investigated under the provisions of the Code of Criminal Procedure Act, No. 15 of 1979.¹³³ The Act vests powers in the police to make arrest¹³⁴ and also to carry out searches and seizures with a warrant.¹³⁵ It also stipulates the powers the police have in the course of investigations.¹³⁶ It further provides for the appointment of a panel of 'experts', i.e., public officers who have the required qualifications and experience in electronic engineering or software technology, to assist any police officer in the investigation of offences.¹³⁷

Case Story J1: Arrests during COVID-19

In March 2020, the Act was used to arrest persons for spreading 'misinformation via social media' in relation to COVID-19.¹³⁸ A media website reported that one suspect was arrested under section 6 of the Computer Crimes Act for 'creating false propaganda' about the virus and sending it to almost 5,000 people on Facebook. The arrest was reportedly justified on the

126. See *The Judicial Mind in Sri Lanka; Responding to the Protection of Minority Rights*, (Law & Society Trust 2014) *op. cit.*, p. 243-244; Niran Anketell & Gehan Gunatilleke, 'Emergency Law in the Context of Terrorism – Sri Lanka', at http://www.southasianrights.org/wp-content/uploads/2009/10/SL-Terrorism-and-Emergency-Laws_Draft1.pdf [last accessed 18 April 2020] p. 23-24.

127. Preamble, Computer Crimes Act, No. 24 of 2007 (CCA).

128. The punishment is imprisonment for a term up to five years.

129. The punishment is a fine between one hundred thousand rupees and three hundred thousand rupees or to imprisonment between six months and three years, or to both.

130. Section 11, CCA.

131. Section 12, CCA.

132. Section 13, CCA.

133. Sections 15 and 16, CCA.

134. Section 21, CCA.

135. Section 18, CCA.

136. Section 19, CCA – requiring the preservation of information; Section 22- record and afford access to seized data; Section 24 – maintaining confidentiality of data.

137. Section 17, CCA. Note - the The Act further describes who an expert would be, procedure for the appointment, and powers of the expert (eg: to enter upon the premises along with a police officer, orally examine any person etc). It also states that the appointment takes place by order published in the Gazette by the Minister of Science and Technology, in consultation with the Minister of Justice.

138. 'Man suspected of spreading false information on coronavirus granted bail', *ColomboPage*, 16 March 2020, at http://www.colombopage.com/archive_20A/Mar16_1584376889CH.php [last accessed 18 April 2020]. Note - subsequent news reports also stated that more arrests will be made under the Computer Crimes Act for the spread of false information about the coronavirus. See 'Investigations underway to arrest 40 suspects for spreading fake news on social media', *News 1st*, 17 March 2020, at <https://www.newsfirst.lk/2020/03/17/investigations-underway-to-arrest-40-suspects-for-spreading-fake-news-on-social-media/> [last accessed 18 April 2020].

basis that the content could harm national security and public peace.¹³⁹ Other persons were also arrested for criticising public officials on Facebook, and spreading rumours.¹⁴⁰ For instance, it was reported that an administrator of a private school was remanded for allegedly uploading misinformation regarding the virus.¹⁴¹

As of July 2020, there are no reported cases of the Act being specifically used in respect of hate speech on social media. However, Case Story J1 on arrests during COVID-19 suggests that the Act may potentially be applied to cover certain forms of hate speech on social media.

1.3 OTHER REGULATORY MECHANISMS: INSTITUTIONS

1.3.1 SLCERT – Sri Lanka Computer Emergency Readiness Team / Coordination Centre¹⁴²

SLCERT was established in 2006.¹⁴³ It is a state institution and currently functions under the purview of the Ministry of Defence. SLCERT provides advice on the latest threats and vulnerabilities affecting the country's computer systems and networks. It also provides expertise in responding to and recovering from cyber-attacks.

For instance, according to reports, as of January 2020 the Ministry of Defence was slated to introduce a new Cyber Security Act under the National Cyber Security Strategy. The proposed Act was drafted by SLCERT and would cover the gambit of issues in cyber security from credit card fraud, hacking and intellectual property crimes, revenge porn, hacking websites, cyber terrorism, hacking of financial data in addition to removal of 'ethically and religiously sensitive posts that spread hate speech' according to Defence Secretary Kamal

Gunaratne.¹⁴⁴ The proposed Act was handed over to the Legal Draftsman Department as of January 2020 and there is no publicly available information on the status of this as of March 2021.

With respect to incidents on social media, specifically, SLCERT *only* provides *technical assistance* to resolve social media incidents. According to its website, SLCERT does *not* provide support to trace or take legal action against perpetrators. It also does not remove content on social media platforms or block websites. However, SLCERT can provide support for removing fake accounts, hacked accounts, and reporting content that violates the privacy policy/community standards of social media platforms and other websites. These conditions point to the limited role that SLCERT can play to combat hate speech online. Therefore, SLCERT can only provide support if a social media account is hacked, or if a fake profile is created and used to generate content constituting hate speech.

1.3.2 Cyber Crimes Division of the Criminal Investigation Department

The Criminal Investigation Department (CID) of the Sri Lanka Police carries out special investigations into specific crimes. In 2015, it was reported that the CID had introduced 'Hi-tech Crime Investigation Units' across the country to fight cybercrimes and increase the expediency with which complaints are handled.¹⁴⁵ It was reported that the units would focus on the sexual harassment of women on social media, threats to children, and financial crimes.¹⁴⁶ There is no follow up data available on the performance of these units. The 2017 Performance Report of the Sri Lanka Police reports that the CID apprehended a number of individuals for financial scams online under the category of cybercrimes.¹⁴⁷

139. 'Man suspected of spreading false information on coronavirus granted bail', *op. cit.*

140. 'Sri Lanka government intensifies crackdown on social media', *op. cit.*

141. 'Educational administrator remanded for uploading fake news', *Daily News*, 25 March 2020, at <http://www.dailynews.lk/2020/03/26/law-order/215175/educational-administrator-remanded-uploading-fake-news> [last accessed 26 June 2016].

142. The status of SLCERT within the report is as at July 2020.

143. Note: The following information is taken from the SLCERT website, at <https://www.cert.gov.lk/aboutUs.php> [last accessed 18 April 2020].

144. 'Govt. to bring new laws to combat online hate speech', *DailyFT*, 24 January 2020 at <http://www.ft.lk/front-page/Govt-to-bring-new-laws-to-combat-online-hate-speech/44-694230> [last accessed 16 March 2021].

145. 'CID to fight cyber-crime with Hi-Tech Crime Investigation Units', *Sunday Observer*, 1 November 2015, at <http://www.sundaytimes.lk/151101/news/cid-to-fight-cyber-crime-with-hi-tech-crime-investigation-units-169982.html> [last accessed 21 April 2020].

146. *ibid.*

147. Sri Lanka Police, *Performance Report (2017)*, at <https://www.parliament.lk/uploads/documents/paperspresented/performance-report-srilanka-police-2017.pdf> [last accessed 21 April 2020] p. 10. Note - under 'Human Resource Development', the report refers to overseas training of personnel in relation to cybercrimes.

Based on these reports, the mandate of the Cyber Crimes Division of the CID is unclear, and it does not appear to

specifically deal with content on social media that could lead to violence.

1.4 OVERVIEW OF THE LEGAL FRAMEWORK AND REGULATORY BODIES

Table 1: Regulatory bodies - Enforcement authority and role of relevant minister¹⁴⁸

Legislation/ Regulatory body	Relevant law enforcement authority	Role of the minister	Other bodies / personnel involved
ICCPR Act	Sri Lanka Police	Section 9 deals with the power of the Minister to make regulations Note: It is difficult to locate the subject ministry of the ICCPR Act	
Penal Code	Sri Lanka Police		
Police Ordinance	Sri Lanka Police	The Police Ordinance gives several powers to the Minister in relation to the administration of the Police Sri Lanka police currently functions under the purview of the Minister of Defence	
Public Security Ordinance and Emergency Regulations	1) Sri Lanka Police 2) Armed forces: As per section 12(2) of the PSO, when the armed forces are called out to maintain public order, they too will have powers of search and arrest which is conferred on police by provisions of the PSO or any other written law, except powers in Chapter XI of the Code of Criminal Procedure Act		The President exercises power under the PSO Under section 6 of the PSO, ERs may delegate powers to authorities or persons to make orders and rules for the purposes for which the ERs are made The President also has the power to call out the armed forces under section 12 when there is a circumstance endangering public security or such circumstance is imminent. The armed forces are also housed under the Minister of Defence
PTA	Sri Lanka Police	Minister of Defence The PTA refers to the powers of the 'Minister' to make detention orders (section 9); and orders of restriction of movement in certain cases (Section 11)	
SLTA, Telecommunications Regulatory Commission of Sri Lanka (TRCSL)	Sri Lanka Police	Minister of Defence	

148. The status of regulatory bodies within the report is updated as at July 2020.

Legislation/ Regulatory body	Relevant law enforcement authority	Role of the minister	Other bodies / personnel involved
CCA	Sri Lanka Police	<p>Several ministers are in charge of various functions under the CCA</p> <p>For certain crimes, the minister in charge is dependent on the nature of the crime</p> <p>E.g.</p> <p>In relation to the prosecution of offences under (a) or (c) in section 6, a certificate under the hand of the Secretary to the ministry of the Minister of Defence is necessary</p> <p>In relation to offence under section 6(b), a certificate under the hand of the Secretary to the Ministry of the Minister in charge of Finance is necessary.</p> <p>Additionally, appointment of a panel of experts (public officers) having qualification and experience in electronic engineering or software technology to assist police in the investigation of an offence in the CCA</p> <p>Section 17(1) provides that the Minister in charge of Science and Technology may in consultation with the Minister in charge of Justice, appoint this panel by order published in the gazette</p>	
SLCERT		Ministry of Defence	
Cyber Crimes Division of the Criminal Investigation Department	Sri Lanka Police	Ministry of Defence	

2. ALTERNATIVE NON-LEGAL FRAMEWORKS

Outside the state's legal regulatory framework (formal), there are alternative means of combatting hate speech and disinformation on social media. Non-legal and social responses to hate speech can substitute or complement formal regulatory frameworks.¹⁴⁹

This section discusses the following alternative mechanisms:

1. Voluntary self-regulation by social media companies
2. Independent international regulation
3. Counter-messaging by civil society and private institution initiatives

An effective strategy to counter online hate speech and disinformation may need to incorporate elements of all three of these mechanisms. Comparative global and local experience suggests that relying solely on one mechanism is not as effective as adopting multiple parallel mechanisms.

2.1 VOLUNTARY SELF-REGULATION BY SOCIAL MEDIA COMPANIES

Among other regulatory mechanisms, voluntary self-regulation is one mechanism adopted by social media organisations.¹⁵⁰ This mechanism is not motivated by legal sanctions, but rather by the need to allow for the 'development and credibility' of the media industry.¹⁵¹

In terms of voluntary self-regulation online, social media platforms routinely remove content from their platforms. This removal is based on the terms of service

or community standards that all users of a platform are bound to accept. Content is removed on the platform's own initiative or based on the complaints of users in discovering that the content is in violation of the platform's community guidelines.¹⁵²

Community standards: Definitions for hate speech and enforcement

Widely used social media platforms have community standards that define the types of content that constitute hate speech. The following section discusses the different definitions and approaches used by selected social media platforms through their community standards. It specifically discusses the definitions, standards and enforcement mechanisms used by Facebook, YouTube, Twitter, and Instagram.


149. UN Human Rights Council Special Rapporteur on Minority Issues (HRC 2015), cited in *Countering Online Hate Speech*, UNESCO Series on Internet Freedom (United Nations Educational, Scientific, and Cultural Organisation 'UNESCO' 2015), at <https://unesdoc.unesco.org/ark:/48223/pf0000233231> [last accessed 27 March 2020].

150. See Article 19, *Self-regulation and 'hate speech' on social media platforms* (2018), at https://www.article19.org/wp-content/uploads/2018/03/Self-regulation-and-%E2%80%99hate-speech%E2%80%99-on-social-media-platforms_March2018.pdf [last accessed 27 March 2020] p.9. The report briefly discussed the three types of media regulatory mechanisms - statutory regulation, co-regulation and self-regulation.

151. *ibid.*

152. *ibid.*, p.15.

Table 2: Community standards¹⁵³: Definitions for hate speech and enforcement

Social Media Platform	Relevant Community Guideline/Policy/Standard	Enforcement
<p>Facebook</p> 	<ul style="list-style-type: none"> • Under ‘objectionable content’, the community guidelines define hate speech as ‘a direct attack against people’ based on the ‘protected characteristics’ of race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disease or disability.¹⁵⁴ • ‘Attack’ is defined as ‘violent or dehumanising speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing, and calls for exclusion or segregation.’¹⁵⁵ • The guidelines are categorised into three tiers based on the severity of content that targets a person or a group of people in relation to the protected characteristics.¹⁵⁶ • Characteristics such as age and occupation are provided protection when referenced along with a protected characteristic. Additionally, Facebook protects refugees, migrants, immigrants, and asylum seekers from the most severe attacks, though they allow commentary and criticism of immigration policies.¹⁵⁷ • When the intent is clear, Facebook may allow people to share someone else’s hate speech content to raise awareness or discuss whether the speech is appropriate to use, to use slurs self-referentially in an effort to reclaim the term or for other similar reasons.¹⁵⁸ 	<ul style="list-style-type: none"> • Facebook may remove content or cover content with a ‘warning screen’.¹⁵⁹ • The platform continues to develop detection technology in order to flag hate speech.¹⁶⁰ • The platform mainly determines whether content needs to be removed or not by either ‘automatically’ removing hate speech or by proactive detection of ‘potentially violating content’.¹⁶¹

153. The status of community standards within the report is updated as at July 2020.

154. Website of Facebook, Community Standards, at https://www.facebook.com/communitystandards/recentupdates/hate_speech/ [last accessed 11 March 2021].

155. *ibid.*

156. *ibid.*



157. *ibid.*

158. *ibid.*

159. Facebook Transparency, *Community Standards Enforcement Report (November 2019)*, at <https://transparency.facebook.com/community-standards-enforcement#hate-speech> [last accessed 28 March 2020].

160. *ibid.*

161. *ibid.*

Social Media Platform	Relevant Community Guideline/Policy/Standard	Enforcement
<p>YouTube</p> 	<ul style="list-style-type: none"> Under 'hateful content', the community guidelines specifically state that hate speech is not allowed. Promoting violence or hatred against individuals or groups based on certain attributes will be removed. These attributes are – age, caste, disability, ethnicity, gender identity and expression, nationality, race, immigration status, religion, sex/ gender, sexual orientation, victims of a major violent event and their kin and veteran status.¹⁶² 	<ul style="list-style-type: none"> Guidelines range from a preliminary warning to restricting the content that can be uploaded. Access to other features on YouTube may be restricted or the user's channel may be removed.¹⁶³ YouTube will be testing new features such as: (i) filter on YouTube Studio, which will automatically hold potentially inappropriate and hurtful comments for review; and (ii) warn users before posting potentially offensive comments.
<p>Twitter</p> 	<ul style="list-style-type: none"> Under the 'hateful conduct policy', promoting violence against or directly attacking or threatening other people or dehumanising people based on race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease is prohibited.^{164 165} The policy provides examples of the types of content that would violate their standards.¹⁶⁶ 	<ul style="list-style-type: none"> Twitter has a range of enforcement options for content that violates its policies. These options include tweet-level enforcement, direct-message-level enforcement, account-level enforcement and action against non-violating content.¹⁶⁷
<p>Instagram</p> 	<ul style="list-style-type: none"> Instagram follows Facebook's rules on defining hate speech. These rules are further strengthened to ban more implicit forms of hate speech, like content depicting Blackface and common antisemitic tropes.¹⁶⁸ 	<ul style="list-style-type: none"> Instagram may delete content, disable accounts or may work with law enforcement if there is a risk of physical harm or threat to public safety.¹⁶⁹

162. Website of YouTube, Hate Speech Policy, at <https://support.google.com/youtube/answer/2801939?hl=en> [last accessed 28 March 2020]. The site offers examples of what would amount to 'hate speech'.

163. Website of YouTube, Community Guidelines strike basics, at <https://support.google.com/youtube/answer/2802032?hl=en> [last accessed 28 March 2020].

164. Website of Twitter, Rules and Policies, at <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> [last accessed 28 March 2020].

165. Website of Twitter, 'Updating our rules against hateful conduct', at https://blog.twitter.com/en_us/topics/company/2019/hatefulconductupdate.html [last accessed 11 March 2021].

166. For example: inciting fear about a protected category.

167. For details on the different types of action, see Website of Twitter, Enforcement options, at <https://help.twitter.com/en/rules-and-policies/enforcement-options> [last accessed 28 March 2020].

168. Website of Instagram, 'An update on our work to tackle abuse on Instagram', at <https://about.instagram.com/blog/announcements/an-update-on-our-work-to-tackle-abuse-on-instagram> [last accessed 11 March 2021].

169. *ibid.*

Effectiveness of voluntary self-regulation

The use of community guidelines and standards offer a voluntary non-legal approach to regulating content on social media. It has been noted that the publication of yearly transparency reports is a step towards ensuring greater accountability with regards to implementing such voluntary self-regulation standards.¹⁷⁰

There is a general sense of acceptance by social media companies that they need to combat hate speech.¹⁷¹ For instance, in Sri Lanka, steps have been taken by Facebook to respond to mis/disinformation and hate speech. In particular, it was reported that in early 2019, Facebook created a team to specifically work on issues in Sri Lanka and other countries where content online could lead to off-line harm or violence.¹⁷² The team seeks to conduct digital literacy workshops with Sri Lankan non-profit organisations and enforce Facebook's community policies.¹⁷³

In an interview in early 2019, a Facebook spokesperson disclosed that it was developing artificial intelligence (AI) that could 'proactively flag posts that break [their] rules'.¹⁷⁴ Facebook also stated that the company had expanded its automatic translation to Sinhala to identify Sinhala content more effectively.¹⁷⁵ It has been identified, however, that AI systems may not be effective in monitoring and detecting misinformation and hate speech in Sinhala and Tamil.¹⁷⁶ Commenting

on Facebook's response to the recently released Human Rights Impact Assessment (HIRA) of the Facebook platform in Sri Lanka, researcher Sanjana Hattotuwa also highlights the obscurity on the number of content moderators working in Sinhala and Tamil to review content in Sri Lanka.¹⁷⁷ Furthermore, Hattotuwa expressed concerns over the platform's 'enduring unevenness' in creating awareness and accessibility to reporting tools in Sinhala and Tamil—allowing time for harmful content to go viral.¹⁷⁸ Moreover, in its response, Facebook underscored its measure to 'strengthen partnerships with local fact-checking organisations in Sri Lanka'.¹⁷⁹ However, despite its current partnerships with third party fact-checkers for Sri Lanka—AFP and Fact Crescendo—a recent report revealed that news stories confirmed to be false by these fact-checkers continue to remain on Facebook with no disclaimers.¹⁸⁰

Case Story E3: Facebook's response post-Easter Sunday Attacks in 2019

In response to the spate of misinformation that was disseminated in the aftermath of the Easter Sunday Attacks, it was reported that Facebook was taking steps to 'remove content' that violates their policies.¹⁸¹ These steps included working with other actors in the region and civil society organisations to identify misinformation that 'has the potential to contribute to imminent violence', 'identifying content that violates

170. *Countering Online Hate Speech* (UNESCO 2015), *op. cit.*

171. *Self-regulation and 'hate speech' on social media platforms*, (2018), *op. cit.*, p. 16.

172. Mazin Hussain, 'The deadly Facebook cocktail: Hate Speech and Misinformation in Sri Lanka', *ReadME*, 14 May 2019, at <https://www.readme.lk/facebook-hate-speech-misinformation-sri-lanka/> [last accessed 22 April 2020].

173. *ibid.* See also, 'Facebook focuses on Digital Literacy to improve online safety in Sri Lanka', *Lanka Business Online*, 21 January 2019, at <https://www.lankabusinessonline.com/facebook-focuses-on-digital-literacy-to-improve-online-safety-in-sri-lanka/> [last accessed 28 June 2020]; 'How Facebook fights false news', *Daily FT*, 04 June 2019, at <http://www.ft.lk/columns/How-Facebook-fights-false-news/4-679370> [last accessed 30 June 2020].

174. *ibid.*

175. *ibid.* See also 'Facebook Apologizes for Role in Sri Lankan Violence', *Bloomberg*, 13 May 2020, at <https://www.bloomberg.com/news/articles/2020-05-12/facebook-apologizes-for-role-in-sri-lankan-violence> [last accessed 15 June 2020] in which Facebook states that subsequent to the Digana riots, the company has taken steps including hiring content moderators with local language skills, implementing technology to automatically detect hate speech, and deepening relationships with local CSOs.

176. See Twitter Account of Yudhanjaya Wijeratne, Tweet dated 24 April 2019, at <https://twitter.com/yudhanjaya/sta->

[tus/1121054915692908545?s=20](https://twitter.com/yudhanjaya/sta-tus/1121054915692908545?s=20) [last accessed 15 June 2020].

177. Sanjana Hattotuwa, 'Facebook's Human Rights Impact Assessment (HIRA) on Sri Lanka: Some brief thoughts', *ICT for Peacebuilding*, 13 May 2020, at <https://ict4peace.wordpress.com/2020/05/13/facebook-human-rights-impact-assessment-hria-of-sri-lanka-some-brief-thoughts/> [last accessed 30 June 2020].

178. Meera Srinivasan, 'Online hate and its offline costs', *The Hindu*, 16 May 2020, at <https://www.thehindu.com/news/international/online-hate-and-its-offline-costs/article31603139.ece> [last accessed 30 June 2020].

179. Website of Facebook, Newsroom, Miranda Sissons, 'An update on Facebook's Human Rights work in Asia and around the world', 12 May 2020, at <https://about.fb.com/news/2020/05/human-rights-work-in-asia/> [last accessed 30 June 2020].

180. Hashtag Generation, *Findings from the social media monitoring exercise during the 2019 Sri Lankan Presidential Election* (2019), at <https://hashtaggeneration.org/wp-content/uploads/2020/01/social-media-monitoring-report.pdf> [last accessed 30 June 2020] p. 19; See Twitter account of Sanjana Hattotuwa, Tweet dated 06 March 2020, at https://twitter.com/sanjanah/status/1235656870724915200?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Cwterm%5E1235656870724915200%7Ctwgr%5E&ref_url=https%3A%2F%2Fict4peace.wordpress.com%2F2020%2F05%2F13%2Ffacebook-human-rights-impact-assessment-hria-of-sri-lanka-some-brief-thoughts%2F [last accessed 30 June 2020].

181. 'The deadly Facebook cocktail: Hate Speech and Misinformation in Sri Lanka', *op. cit.*

[Facebook's] policies' and 'ensuring language support for content review'.¹⁸²

The company also began removing content that was found to be 'praising or supporting the attacks'.¹⁸³ The Community Operations team began to support law enforcement in Sri Lanka and worked with CSOs in order to identify misinformation that could lead to imminent violence or physical harm, particularly in relation to content in Sinhala and Tamil.¹⁸⁴

During this period, Hattotuwa documented an unprecedented rate at which pages and posts inciting hate were reported and removed on Facebook.¹⁸⁵ However, reports monitoring Facebook during this period also revealed that over ten posts, which incited violence toward Muslim minorities in Sri Lanka remained visible on the platform at the time of the attacks and for days afterward.¹⁸⁶ Moreover, Facebook also made changes to its Facebook Messenger and WhatsApp apps to limit the number of message forwards in an attempt to reduce virality of content that has the potential to exacerbate violence.¹⁸⁷

More recently, in the context of COVID-19, several social media and online platforms pledged to take steps to tackle misinformation relating to the pandemic.¹⁸⁸ A joint statement by Facebook, Google, LinkedIn, Microsoft, Reddit, Twitter and YouTube indicated that the platforms were working together in their COVID-19 response efforts.¹⁸⁹

182. *ibid.*

183. *ibid.*

184. *ibid.*

185. See Twitter account of Sanjana Hattotuwa, Tweet dated 29 April 2019, at <https://twitter.com/sanjanah/status/1122632338863558657>; Twitter account of Sanjana Hattotuwa, Tweet dated 30 April 2019, at <https://twitter.com/sanjanah/status/1122988532480405505> [last accessed 30 June 2020].

186. Paul M. Barrett, *Who moderates the social media giants? A Call to end outsourcing* (New York University Stern Center for Business and Human Rights 2020), at https://issuu.com/nyusterncenterforbusinessandhumanri/docs/nyu_content_moderation_report_final_version/21?fr=sZWfMzE0OTcyNDk,p [last accessed 30 June 2020] p. 21.

187. Website of Facebook, Newsroom, Samidh Chakrabarti, 'Understanding social media and conflict', 20 June 2019, at <https://about.fb.com/news/2019/06/social-media-and-conflict/> [last accessed 30 June 2020].

188. 'Why is it so hard to stop COVID-19 misinformation spreading on social media', *The Conversation*, 14 April 2020, at <https://theconversation.com/why-is-it-so-hard-to-stop-covid-19-misinformation-spreading-on-social-media-134396> [last accessed 22 April 2020].

189. See Twitter Account of Facebook Newsroom, Tweet

Case Story J2: COVID-19 in 2020

In Sri Lanka, Facebook pledged that they would work with national health organisations to limit the spread of misinformation and harmful content on the COVID-19 pandemic.¹⁹⁰ In Sri Lanka, Facebook directs users to link to health advice and information from the WHO, UNICEF and Sri Lanka's Health Promotion Bureau.¹⁹¹ They also claimed they were taking measures to remove content containing false claims or conspiracy theories that could cause physical harm to people.¹⁹² Facebook's Sri Lanka COVID-19 Fact Sheet revealed that Facebook, in March 2020, displayed warnings on over 40 million posts related to COVID-19 and 'removed hundreds of thousands' of pieces of misinformation that could lead to imminent physical harm.¹⁹³

It should be noted that a recent report by Hashtag Generation depicts marginal improvement in relation to language detection during the parliamentary election period from 15 June to 5 August 2020.¹⁹⁴ However, the report provides that, of forty incidents related to hate speech, only fifty percent of content was removed by Facebook.

Voluntary self-regulation may have its merits, as social media platforms may be incentivised to be proactive when the alternative is regulation by the state. However, revelations of Russian interference in the lead up to the Brexit referendum and the 2016 United States

dated 17 March 2020, at <https://twitter.com/fbnewsroom/status/1239703497479614466?s=20> (last accessed 15 June 2020). See also Website of Facebook, Newsroom, Kang-Xing Jin, 'Keeping People Safe and Informed about the Coronavirus' at <https://about.fb.com/news/2020/06/coronavirus/#joint-statement> [last accessed 15 June 2020].

190. For detailed information on steps Facebook is taking, See 'Facebook takes action against misinformation on COVID-19 and keeps people safe and informed in Sri Lanka', *Daily Mirror*, 19 March 2020, at <http://www.dailymirror.lk/business/Facebook-takes-action-against-misinformation-on-COVID-19-and-keeps-people-safe-and-informed-in-Sri-L/215-185321> [last accessed 22 April 2020].

191. Website of Facebook, 'Sri Lanka COVID-19 Fact Sheet', 16 April 2020, at <https://economynext.com/wp-content/uploads/2020/04/Sri-Lanka-COVID-19-Responsibility-and-Integrity-Media-Briefing-Fact-Sheet.pdf> [last accessed 30 June 2020].

192. *ibid.*

193. *ibid.*

194. Hashtag Generation, *Sri Lanka: Social Media and Electoral Integrity; Findings from the 2020 parliamentary Election* (03 September 2020), at https://drive.google.com/file/d/1qQKubeK1HtAtom-3J5v2BkoQ5hToH9ma_/view?fbclid=IwAR3tEPm6Ke3Oz6RaQ9cu-4WgcSHK-s0r45wx9ocA9H64eYjPWY0diYPLZqKc [last accessed 13 October 2020].

presidential election, suggest that reliance on voluntary self-regulation alone may not be adequately effective in dealing with widespread disinformation.¹⁹⁵ Moreover, social media platforms face the challenge of lacking the capacity to takedown hate speech in local languages.

Apart from the lack of capacity, there is often a lack of clarity and transparency in the process used by social media platforms to enforce guidelines and standards. For instance, there have been increased calls for transparency surrounding Facebook's advertising policy. Social media companies that are geared towards making profits, such as Facebook, which primarily generate revenue through advertisements, have been observed to be disincentivised to take down material that may contain disinformation. Several investigative reports on the Brexit referendum and the 2016 US presidential election suggest that Brexit would not have taken place and Donald Trump would not have been elected president if not for the influence of 'fake news' allowed by Facebook's advertising engine.¹⁹⁶ In effect, Facebook has often failed to restrict content that can undermine rights.¹⁹⁷ Therefore, it makes sense to opt for multiple alternative means (i.e. means outside formal state interventions) through which online hate speech and disinformation can be combatted.

2.2 INDEPENDENT INTERNATIONAL REGULATION

Many stakeholders recommend independent external mechanisms for social media as an effective option.¹⁹⁸ They suggest the creation of a council at a national or

international level, or a combination of both.¹⁹⁹ A report by ARTICLE 19 for example recommends that such a council can set elaborate ethical standards specific to the online distribution of content, and cover other areas such as community guidelines and content regulation practices of social media companies.²⁰⁰ It has been suggested that such a council can engage with the public through appropriate consultative processes, while relying on principles of transparency, and peer and public pressure.²⁰¹

The only guide relating to self-regulation of social media in Sri Lanka, which is available at the time of publishing this report, is the Social Media Declaration (2019): Code of Conduct for Responsible Social Media Use. This Code sets out important standards that several organisations aspire to achieve.²⁰² It also stipulates the responsibilities of social media platforms, the government, media organisations and civil society. In effect, it recognises the importance of multi-stakeholder responsibility in the regulation of social media.

Appealing to internationally accepted norms, principles and ethics can be a powerful means through which social media can be regulated. Towards that end, setting up an international independent council to guide and advise on the standards and hear serious complaints from individual national contexts can be a useful initiative. An international initiative is much more likely to succeed over national initiatives due to the global scale of social media platforms and the global legitimacy that is required for the adoption and modification of standards and their proper implementation. In order to enable such an international council to be inclusive in its deliberations, it is sensible to simultaneously invest in national-level citizen-driven initiatives that can draw attention to the unchecked spread of problematic content, and engage with the envisaged international council on the necessary adoption of and improvements to standards from a local context and perspective.

195. *Self-regulation and 'hate speech' on social media platforms*, (2018), *op. cit.*, p. 15-16; See also *Countering Online Hate Speech*, (UNESCO 2015), *op. cit.*, p. 30-32.

196. Hunt Allcott and Matthew Gentzkow, 'Social Media and Fake News in the 2016 election' (2017), Volume 31, Number 2, *Journal of Economic Perspectives* 211-236, at <https://web.stanford.edu/~gentzkow/research/fakenews.pdf> [last accessed 30 June 2020] p. 212; Timothy Garton Ash, Robert Gorwa and Danae Metaxa, *GLASNOST! Nine ways Facebook can make itself a better forum for free speech and democracy* (2019), at https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-01/Garton_Ash_et_al_Facebook_report_FINAL_0.pdf [last accessed 30 June 2020] p. 14. Note: Such findings are only preliminary and there have been no conclusive findings presented in a court of law.

197. 'Facebook's Human Rights Impact Assessment (HIRA) on Sri Lanka: Some brief thoughts', *op. cit.*

198. *Self-regulation and 'hate speech' on social media platforms*, (2018), *op. cit.* p. 20.

199. *ibid.*

200. *ibid.*

201. *ibid.*, at p. 21-22 for factors that must be considered in the creation of the council.

202. The signatories include several think-tanks, policy organizations and online news and blog spaces.

2.3 COUNTER-MESSAGING BY CIVIL SOCIETY AND PRIVATE INSTITUTION INITIATIVES

Counter-speech means raising awareness, improving education and building the capacity to speak out against hate speech.²⁰³ Counter-speech is often preferred to the suppression or restriction of free speech.²⁰⁴ Counter-messaging or speech can take place at two levels: institutional and individual.

Institutional level:

- Counter-messaging initiatives by civil society, media actors and activists

Civil society plays a pivotal role in protecting and promoting the freedom of expression, and countering or responding to hate speech.²⁰⁵ Civil society initiatives are among the most innovative and effective means of countering hate speech and responding to incidents of violence.²⁰⁶

From a comparative perspective, certain instances depict how civil society actors mobilise to condemn hate speech online and call for alternatives to censorship. For example, the Panzagar (flower speech) campaign was launched in April 2014 as part of an attempt to quell a growing tide of online hate speech in Myanmar.²⁰⁷ Other examples include the 'No Hate Speech Movement' project by the Council of Europe. The project provides a platform for sharing initiatives and best practices by civil society campaigns to prevent and respond to hate speech.²⁰⁸

Alternative models in other parts of the world also include codes of conduct adopted jointly by companies and public institutions. One such example is the EU Code of Conduct on countering illegal hate speech online

203. ee generally, *Countering Online Hate Speech*, (UNESCO 2015), *op. cit.*

204. *ibid.* p. 5.

205. Article 19, *Responding to 'hate speech' with positive measures: A case study from six EU countries* (2018), p. 28.

206. *ibid.*

207. *Countering Online Hate Speech*, (UNESCO 2015), *op. cit.*, p. 38-39.

208. For more examples in Europe, see *Responding to 'hate speech' with positive measures: A case study from six EU countries*, (2018), *op. cit.*, p. 28-30.

developed by the European Commission and key digital technology companies.²⁰⁹

In Sri Lanka, there are examples of civil society engagement on social media and monitoring of social media platforms for potentially misleading and dangerous content. For instance, during the presidential elections in 2019, Hashtag Generation, a youth-led CSO,²¹⁰ monitored social media for dangerous speech, false information, and violations of electoral legislation.²¹¹ Similarly, CSO co-operation has been evident in capacity-building for better social media monitoring. One example of an initiative to strengthen social media monitoring was a workshop conducted by Democracy Reporting International (DRI) in December 2019. The workshop was conducted for the Centre for Monitoring Election Violence (CMEV) and People's Action for Free and Fair Elections (PAFFREL).²¹² The workshop identified key issues that need to be monitored, which included reconciliation and peace, and religious and ethnic issues.²¹³ Such examples highlight the opportunities for civil society to engage in counter-messaging at an institutional level.

While civil society has an important role in countering hate speech and disinformation, multiple obstacles serve to undermine its effectiveness. Primarily, social media analysis conducted by Hattotuwa observed that a majority of CSO Facebook pages gain marginal traction for their content in comparison to political gossip and meme pages.²¹⁴ This lack of traction was particularly

209. European Commission, *The Code of Conduct on countering illegal hate speech online*, (31 May 2016).

210. Hashtag Generation is a 'A youth-led movement advocating for full and effective participation of young people in policy making, implementation and evaluation at local, national, regional & international levels.' See https://www.facebook.com/pg/hashtaggenerations/about/?ref=page_internal [last accessed 11 June 2020].

211. Support afforded by Hashtag Generation cited in European Union Election Observation Mission, *Final Report: Sri Lanka Presidential election, 16 November 2019* (2020), at http://www.epgencms.europa.eu/cmsdata/upload/570dea26-4277-40e6-abb8-f50c59c8d403/Sri-Lanka_presidential_election_16_November_2019_EU_EOM_report.pdf [last accessed 11 June 2020] p.43.

212. Website of Democracy Reporting International, 'Sri Lanka: DRI works with local partners to strengthen social media monitoring', 16 December 2019, at <https://democracy-reporting.org/sri-lanka-dri-works-with-local-partners-to-strengthen-social-media-monitoring/> [last accessed 11 June 2020].

213. *ibid.*

214. See Twitter account of Sanjana Hattotuwa, Tweet dated 14 October 2019, at <https://twitter.com/sanjanah/status/1183661078258962434> [last accessed 30 June 2020].

evident during last year's presidential election. Additionally, Sri Lanka's low rates of digital and media literacy and a digital divide between urban and rural areas impede CSO counter-messaging initiatives.²¹⁵

- Fact-checking and ethical engagement by private or social institutions

Private or social institutions can play a vital role in ensuring the accuracy of information online. Disinformation that could constitute hate speech can be verified by online fact-checking platforms. For instance, FactCheck.lk, Watchdog Sri Lanka, and FactCrescendo aim to combat online disinformation in Sri Lanka.²¹⁶ Their mandates deal with determining the veracity of information. However, these mandates could be extended to identify hate speech online. Furthermore, 'Hate Speech Monitor Sri Lanka', comprising civil society volunteers, was created after the Easter Sunday Attacks to address hate speech and misinformation. In particular, this initiative encourages civil society engagement and public discourse on hate speech on social media. However, a report commissioned to monitor social media during the 2019 Sri Lankan presidential election observed that AFP and Fact Crescendo—which worked with Facebook during the election to debunk false news—did not have the reach (among Sri Lankan social media users) that false narratives had.²¹⁷ Furthermore, the report observed that despite catering to all local languages, the fact-checking platforms failed to achieve widespread reach.²¹⁸

Private institutions can counter hate speech and disinformation by appealing not only to factual issues in the content, but also to ethical issues. For instance, Ethics

Eye, a social media platform run by Verité Research, seeks to hold mainstream media organisations to account for publishing content that violates standards of ethical reporting. The platform also covers content shared on social media by mainstream media actors, and has built a strong base of support in society to encourage media houses to stop publishing content that violates these ethical standards. A similar model could be applied to content shared on social media platforms. Building social support for the recognition and rejection of content that violates not just factual, but also ethical standards can be an important first step towards modifying standards and moderation policies of social media companies.

The table below illustrates a breakdown of these platforms.






215. *Final Report: Sri Lanka Presidential election, 16 November 2019* (2020), *op. cit.*, p. 35.

216. Mandates: FactCheck.lk – 'FactCheck is a non-profit platform that monitors a select sample of Sri Lankan print media (Sinhala, Tamil, and English) to identify and fact check statements attributed to high-level decision makers in public office on subjects of public interest.' (<http://factcheck.lk/>); Watchdog Sri Lanka – A multidisciplinary team of individuals who deal with misinformation on social media channels. (<https://watchdog.paladinanalytics.com/about.html>); Fact Crescendo—An independent digital journalism initiative that engages in fact-checking of both social media and mainstream media. (<https://srilanka.fact-crescendo.com/about/>).

217. *Findings from the social media monitoring exercise during the 2019 Sri Lankan Presidential Election* (2019), *op. cit.*, p. 19; *Final Report: Sri Lanka Presidential election, 16 November 2019* (2020), *op. cit.*, p. 35.

218. *ibid.*

Table 3: Fact-checking platforms - Mandate and followers

Name of Platform	Mandate	Number of followers (as at 30 June)
<p>Factcheck.lk</p> 	<p>A non-profit platform that monitors a select sample of Sri Lankan print media (Sinhala, Tamil, and English) to identify and fact check statements attributed to high-level decision makers in public office on subjects of public interest²¹⁹</p>	<p>Facebook 3,170 likes</p> <p>Twitter 1,287 followers</p>
<p>Watchdog Sri Lanka</p> 	<p>A multidisciplinary team of individuals who deal with misinformation on social media channels²²⁰</p> <p>The initiative primarily works through an App</p>	<p>Facebook 6,495 likes</p> <p>Twitter 3,431 followers</p>
<p>FactCrescendo</p> 	<p>An independent digital journalism initiative that engages in fact-checking of both social media and mainstream media²²¹</p> <p>The team reviews and rates the accuracy of stories on Facebook and other social media platforms in Sri Lanka²²²</p>	<p>Facebook 14,708 likes</p> <p>Twitter 1,249 followers</p>
<p>Hate Speech Monitor Sri Lanka</p> 	<p>A civil society initiative, created in the aftermath of the Easter Sunday Attacks to address hate and misinformation</p> <p>The parameters used to determine hate speech are:</p> <ul style="list-style-type: none"> • Messages that call for discrimination or violence against any group or community • Fake news and misinformation • Business and trade boycotts • Disrespect for religious/cultural practices and beliefs in order to incite²²³ 	<p>Facebook 4,708 likes</p>
<p>Ethics Eye</p> 	<p>A research-based public education platform that serves to foster ethical journalism in Sri Lanka. The platform engages both the public and journalists on current ethical concerns that arise from daily reportage with a high reach in Sinhala²²⁴</p>	<p>Facebook 25,819 likes</p> <p>Twitter 3,185 followers</p>

219. Website of FactCheck, at <http://factcheck.lk/> [last accessed 11 June 2020].

220. Website of WatchDog, at <https://watchdog.paladinanalytics.com/about.html> [last accessed 30 June 2020] Note: WatchDog's social media platforms appear to be inactive as of the revision of the report in March 2021.

221. Website of FactCrescendo, at <https://srilanka.factcrescendo.com/about/> [last accessed 30 June 2020].

222. Facebook page of FactCrescendo, at <https://www.facebook.com/factcrescendosl/> [last accessed 30 June 2020].

223. Facebook Page of Hate Speech Monitor, at <https://www.facebook.com/hatespeech.lk/> [last accessed 30 June 2020].

224. Facebook Page of Ethics Eye, at <https://www.facebook.com/ethicseye/> [last accessed 30 June 2020].

Individual level:

This level concerns several interrelated mechanisms for countering hate speech on social media. Primarily, citizens can be educated on the elements of hate speech and the types of content that should be prohibited. Information literacy can be improved, thereby enabling individuals to counteract hate speech on their own volition. In its 2018 report, ARTICLE 19 makes suggestions on the measures that states could take to promote counter-speech, and how members of the public can respond to hate speech. They suggest that investing in digital literacy skills, will result in the realisation of the benefits of engagement online, and the importance of such engagement.²²⁵

Presently, CSOs and other institutions have developed initiatives to enhance social media literacy, particularly among youth. For instance, HackaDev,²²⁶ an initiative that commenced in 2015, seeks to meaningfully engage youth in Sri Lanka to identify sustainable and innovative solutions for identified development challenges.²²⁷ The initiative has potential to engage youth in social media literacy through the development of innovative solutions for this engagement.

Other examples include initiatives by Search for Common Ground (SFCG), Sri Lanka, an organisation that primarily works on initiatives for sustainable peace through dialogue, media and community level engagement.²²⁸ For instance, in 2019, SCFG mobilised youth to counter increased ethno-religious violence and hate speech on social media.²²⁹ This was done through their 'Work together, Win together' project, of which the second phase focused on social-media campaigns

to counter hate speech and promote reconciliation.²³⁰ The project sought to empower youth to 'produce and disseminate positive stories of peaceful co-existence and reconciliation'.²³¹ 'Cyber Guardian', another initiative by SFCG, focused on combatting the spread of fake news and hate speech on social media. The objective was fulfilled by empowering youth to use their social media platforms to counter fake news and hate speech, as well as capacity building on using social media responsibly.²³²

In addition, there are other initiatives which have worked on empowering youth on digital literacy. These include (but are not limited to) Digital Story Telling²³³ - an educational initiative that works on citizen journalism and digital literacy,²³⁴ Hashtag Generation which worked on countering hate speech during COVID-19,²³⁵ the website Bakamoono.lk²³⁶ that engages on digital activism,²³⁷ and iVoice - a digital content platform that promotes evidence based community journalism on social issues concerning women, the elderly, the youth and the community.²³⁸ Further, United Nations Development Programme (UNDP) launched an initiative titled 'Engage to Disengage' - an awareness building and training programme on 'Critical Media and Information Literacy' for young people in Sri Lanka.²³⁹ Other civil society groups have also focused on engaging and empowering individuals on the responsible consumption of content.

225. *Responding to 'hate speech' with positive measures: A case study from six EU countries* (2018), *op. cit.*, p. 23; See also *Countering Online Hate Speech* (UNESCO 2015), *op. cit.*, at p. 45-52 which refers to examples of initiatives that seek to improve digital literacy.

226. HackaDev is a joint initiative of the Ministry of Digital Infrastructure and Information Technology, United Nations Development Programme, Information and Communication Technology Agency and National Youth Services Council. See Website of HackaDev, at <https://hackadev.lk/#!/about> [last accessed 15 June 2020].

227. Website of HackaDev, at <https://hackadev.lk/#!/about> [last accessed 15 June 2020]. See also Website of UNDP, 'HACKADEV', at <https://www.lk.undp.org/content/srilanka/en/home/youth-and-innovation/projects/HackaDev.html> [last accessed 15 June 2020].

228. Website of Search for Common Ground, at <https://www.sfcg.org/what-we-do/> [last accessed 15 June 2020].

229. Search for Common Ground, *Search Sri Lanka 2019*, at <https://www.sfcg.org/wp-content/uploads/2020/01/Search-Sri-Lanka-2019-Annual-Report.pdf> [last accessed 15 June 2020] p. 3.

230. *ibid.*

231. *ibid.*, p. 11.

232. *ibid.*, p. 13.

233. Facebook Page for Digital Storytelling, at <https://www.facebook.com/DSTjourney/> [Last accessed on 11 March 2021].

234. National Youth Dialogue, *Benislos Thushan*, at <https://hackadev.lk/nyd2021/personnel/benislos-thushan/> [Last accessed on 11 March 2021].

235. Internews, *Young Sri Lankans Use Video to Counter COVID-19 Hate Speech*, at <https://internews.org/updates/young-sri-lankans-use-video-counter-covid-19-hate-speech> [Last accessed on 11 March 2021].

236. Website of Bakamoono.lk, at <https://freedomhouse.org/country/sri-lanka/freedom-net/2018> [Last accessed on 11 March 2021].

237. Freedom House, *Key Developments, June 1, 2017- 31 May, 2018* at <https://freedomhouse.org/country/sri-lanka/freedom-net/2018> [Last accessed on 11 March 2021].

238. Website of invoice, at <https://srilanka.unfpa.org/en/news/unfpa-sdjf-launch-wwwivoicelk-%E2%80%93-tri-lingual-online-platform-youth-discuss-social-issues-and> [Last accessed on 11 March 2021].

239. UNDP, *Engage to Disengage: Importance of Media and Information Literacy amidst a pandemic*, at <https://www.lk.undp.org/content/srilanka/en/home/presscenter/articles/2020/Engage-to-Disengage-Importance-of-Media-Information-Literacy-amidst-a-pandemic.html>. [Last accessed on 11 March 2021].

Beyond these local initiatives, the European Union (EU) Election Observation Mission concluded that government initiatives to increase digital and media literacy and reduce the digital divide are yet to manifest their anticipated impact.²⁴⁰ These initiatives involved providing free WiFi points and opening community based digital e-libraries. The engagement of CSOs and other stakeholder institutions to create awareness and enhance social media literacy among citizens is likely to have a positive impact in reducing the harm caused by problematic content as well as building social resistance to the proliferation of such content.²⁴¹

CONCLUSIONS AND OBSERVATIONS

Primary Observations

The purpose of this study was to outline the legal and alternative non-legal frameworks in Sri Lanka that can be used to regulate online hate speech and disinformation.

Accordingly, this study evaluated several laws with sufficiently broad provisions in the existing ‘formal’ legal framework that can be applied to regulate social media, as well as several non-legal mechanisms that can be deployed.

Although legal provisions cover social media, the formal regulatory framework carries with it the risk of abuse and overinclusive or selective application. Several substantive legal provisions in the existing laws are vague and have not been clearly defined nor interpreted by judicial bodies. In this context there is space for the misapplication or selective application of these legislative provisions under the ostensible guise of regulating social media. The use of laws to target critical and dissenting voices, and the lack of enforcement of laws against those inciting discrimination, hostility or violence against minority groups, point to the inherent

risks in relying on legal mechanisms alone to effectively deal with online hate speech and disinformation.

Despite these risks, regulatory bodies, such as the Telecommunications Regulatory Commission of Sri Lanka (TRCSL) can exercise significant control over the regulation of communication. While the law affords the TRCSL and the relevant ministerial powers to regulate social media, the politicisation of the TRCSL in effect means that these powers may be exercised in an overbroad manner.

Other formal regulatory mechanisms do not have structured systems in place to regulate social media in the context of expression that may advocate hatred constituting incitement to discrimination, hostility or violence.

In terms of the alternative framework with non-legal mechanisms, this study identified three key mechanisms, i.e. voluntary self-regulation by social media companies, independent international regulation, and counter-messaging by civil society and private institution initiatives. Voluntary self-regulation by social media companies based on their community standards and guidelines has been acknowledged as having been successful to a certain extent. However, the adequacy of voluntary self-regulation is in doubt. While there has been evidence of social media companies, particularly Facebook, responding to hate speech in Sri Lanka, questions have been raised about their transparency and ability to regulate the spread of hate speech. To this extent, voluntary self-regulation alone is not considered to be sufficient in the domain of alternative non-legal frameworks.

Another option that this study explores is counter-messaging by civil society actors and private institutions – at both the institutional and individual levels. The study cites examples of both types of initiatives, and stresses the importance of social media literacy as a tool for counter-messaging. Private and social institutions can be crucial in providing and ensuring accurate information through initiatives like fact-checking when it comes to combating online disinformation. It is suggested that both CSOs and other private institutions explore this

240. European Union Election Observation Mission, *Final Report: Sri Lanka Presidential election (16 November 2019)*, January 2020, *op. cit.*, p. 35. at https://eeas.europa.eu/sites/default/files/sl2019_final_report_24_january_2020.pdf [last accessed 25 March 2021].

241. *ibid.*

space more to empower individuals to identify different types of expression on social media.

The study observes that a combination of all three mechanisms may be needed to effectively regulate social media in Sri Lanka. An alternative non-legal regulatory framework should operate as a model that complements the formal regulatory framework.

Observations of Eight Key Options:

This section provides observations of eight key options to improve the formal and alternative non-legal frameworks that regulate online hate speech and disinformation in Sri Lanka. The eight key options for interventions are categorised according to the relevant category of implementation of each intervention: (i) legislative interventions; (ii) CSO-led interventions; and (iii) cross-cutting interventions.

i. Observations of Legislative Options

This section evaluates four legislative options in relation to the formal regulatory frameworks available to government to bring about change through the relevant laws and their application.

a. Formal Regulatory Framework

(1) Amending the ICCPR Act to improve clarity

Since the ICCPR Act is currently framed using overbroad terminology which collapses multiple offenses into one, it can lead to the government abusing the law through the misapplication of the law in terms of what constitutes incitement to discrimination or hostility. Therefore, the terminology and interpretative framework can be clarified and criminal sanctions involving incarceration under section 3(1) of the ICCPR Act can be confined to instances of incitement to violence – where there is a clear nexus between the advocacy of national, racial or religious hatred and an actual risk of violence.

(2) Amending the provisions in the Penal Code to clarify the scope of the offences

There is a lack of clarity on what several phrases in the provisions of the Penal Code mean. Moreover, there is

insignificant jurisprudence that clarifies the ambit and scope of the provisions. On the one hand, these sections can be misapplied. On the other hand, the lack of clarity with regard to their scope could result in their disuse. The provisions set out in table 4 could be amended to provide clarity on the scope of offenses.

Table 4: Phrases that lack clarity in the Penal Code

Penal Code Section	Phrase
<ul style="list-style-type: none"> Section 291A: Uttering words with deliberate intent to wound religious feelings. 	<ul style="list-style-type: none"> 'Wounding the religious feelings.'
<ul style="list-style-type: none"> Section 291B: Deliberate and malicious acts intended to outrage the religious feelings of any class, by insulting its religion or religious beliefs. 	<ul style="list-style-type: none"> 'Outraging the religious feelings.'
<ul style="list-style-type: none"> Section 120: Exciting or attempting to excite disaffection. 	<ul style="list-style-type: none"> 'Discontent or disaffection.' 'Promote feelings of ill-will.'

(3) Extension of SLCERT's mandate

SLCERT has some oversight functions in relation to cyberspace, but it currently has a limited role. It does not have a structured system in place to regulate social media in the context of expression that may advocate hatred constituting incitement to discrimination, hostility or violence.

Given their expertise in handling issues on cyberspace, the mandate of SLCERT could be extended to include a structured and transparent system to evaluate social media content in relation to hate speech and disinformation. Such a mandate should provide a transparent and formal acceptable basis of evaluation within the local context, and legitimacy, thereby, for requests for removal of such harmful content.

However, any such increase in the scope of engagement would require SLCERT to be subject to a multi-stakeholder accountability mechanism to ensure adequate checks and balances that evaluation of content does not infringe on an individual's freedom of expression/speech. Any content moderation initiatives

would require sensitisation to not just concerns over freedom of speech but also gender and ethno-religious concerns. In the absence of adequate checks and balances any increase in the scope of government action is at risk of being abused as outlined earlier in the analysis. Furthermore, such an expansion in its mandate requires capacity building of relevant staff as a prerequisite. Capacity building programs should include knowledge and training on 'digital rights', i.e. for instance the intersection between the freedom of expression and the use of social media platforms.²⁴²

(4) Intermediary liability imposed by the state

Another option that is being explored by states is the imposition of liability on internet intermediaries to regulate content on their respective platforms.

One such effort is Germany's Network Enforcement Act ('NetzDG'), which took a novel approach to social media liability by enacting stringent intermediary liability legislature.²⁴³ Germany was the first country in the European continent to hold social media platforms responsible for combatting online speech deemed illegal under domestic law.²⁴⁴ The law requires online platforms to set up effective complaint mechanisms, and to delete or block explicitly illegal content. In Germany, failure to delete or block such content may result in the imposition of fines up to €50 million.²⁴⁵ Germany's legislation has been criticised for placing an unreasonable onus on intermediaries to monitor content. It has also been criticised for incentivising intermediaries to err on the side of caution to avoid heavy penalties, and even censor legitimate free speech. Human Rights Watch and other international bodies also opposed the legislation, as it

242. Kay Mathiesen, 'Human Rights for the Digital Age' (2014) 29 (1) *Journal of Mass Media Ethics* 2-18, at <https://doi.org/10.1080/08900523.2014.863124> [last accessed 25 June 2021].

243. Patrick Zurth, 'The German NetzDG as Role Model or Cautionary Tale? Implications for the Debate on Social Media Liability' (2021) 31 (4) *Fordham Intell. Prop. Media & Ent. L.J.* 1084 – 1153.

244. David Morar and Bruna Martins dos Santos, 'The push for content moderation legislation around the world' (21 September 2020), at <https://www.brookings.edu/blog/techtank/2020/09/21/the-push-for-content-moderation-legislation-around-the-world/> [last accessed 25 June 2021].

245. William Echikson and Olivia Knodt, *Germany's NetzDG: A key test for combatting online hate*, Research Report No. 2018/09, (Counter Extremism Project 2018), at http://wp.ceps.eu/wp-content/uploads/2018/11/RR%20No2018-09_Germany's%20NetzDG.pdf [last accessed 27 March 2020].

would set a dangerous precedent for states around the world to restrict online speech.²⁴⁶

The Indian model for intermediary liability began with intermediary protection laws following the enactment of the Information Technology Act 2000. The Act provides intermediaries with immunity from liability in relation to unlawful third-party content so long as the intermediary does not have actual knowledge.²⁴⁷ Under the Act, intermediaries are required to remove content deemed to be unlawful by order of the court or a state directive. Due diligence requirements as to whether actual knowledge is present under the Act is prescribed under the Information Technology (Intermediary Guidelines) Rules 2011 (Intermediary rules).²⁴⁸

In 2018, India announced that it would be amending its rules concerning intermediary liability. Draft amendments were published by the Ministry of Electronics and IT in 2018, which were aimed at modifying the 2011 rules.²⁴⁹ These amendments sought to place a higher onus on intermediaries in relation to user-generated content. The proposed amendments included requiring the traceability of information including encrypted messaging platforms and proactively identify and remove unlawful content.²⁵⁰ The proposed amendments have been criticised for being overly broad and in effect reduce online security.²⁵¹ However, as of July 2020, the 2020 regulations are yet to be finalised.

246. Human Rights Watch, 'Germany: Flawed Social Media Law', 14 February 2018, at <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law> [last accessed 19 April 2020].

247. Vikram Jeet Singh and Prashant Mara, 'India: Liable vs. Accountable: How Criminal Use Of Online Platforms And Social Media Poses Challenges To Intermediary Protection In India', *mondaq*, 5 May 2020, at <https://www.mondaq.com/india/social-media/928106/liable-vs-accountable-how-criminal-use-of-online-platforms-and-social-media-poses-challenges-to-intermediary-protection-in-india> [last accessed 13 October 2020].

248. *ibid.*

249. Website of Ministry of Electronics & Information Technology – Government of India, 'The Information Technology [Intermediaries Guidelines (Amendment) Rules] 2018', at https://www.meity.gov.in/writereaddata/files/Draft_Intermediary_Amendment_24122018.pdf [last accessed 13 October 2020].

250. Arjun Jayakumar, 'The changing face of intermediary liability in India', *The Leaflet*, 18 November 2019, at <https://www.theleaflet.in/the-changing-face-of-intermediary-liability-in-india/> [last accessed 13 October 2020].

251. Christine Runnegar, 'Making Intermediaries Liable for Encrypted Content Breaks Trust and Security', *InternetTrust*, 4 June 2020, at <https://www.internetsociety.org/blog/2020/06/making-intermediaries-liable-for-encrypted-content-breaks-trust-and-security/> [last accessed 13 October 2020].

The Malaysian Anti-Fake News Act 2018 also makes online intermediaries liable for disinformation. Such laws may have a chilling effect on the freedom of expression, as stringent criminal penalties can incentivise intermediaries to overcautiously remove even lawful content in anticipation that the content violates the law.²⁵² Social media companies may not be able to make informed decisions about the legitimacy of content.²⁵³ It should be noted that the mentioned Act was repealed in December 2019, presently, however, calls are being made to reinstate the Act.²⁵⁴

Intermediary liability is an option that is available to Sri Lanka as well. However, the comparative experience illustrates the negative consequences of this type of regulation on freedom of expression and the potential for misuse by the state.

ii. Observations of options for interventions from Civil Society Organisation

This section assesses two key options to improve the alternative non-legal framework for the regulation of expression that may advocate hatred constituting incitement to discrimination, hostility or violence on social media. These options can be spearheaded by civil society organisations.

a. Alternative Non-Legal Framework

(5) Advocating for and emulating an independent international council for regulation

Independent external regulatory mechanisms for social media have been suggested as an effective option.²⁵⁵ The creation of a council at an international level could allow such a body to: (i) engage with concerns arising

from across nations through appropriate consultative processes; and (ii) independently develop and legitimise standards through which social media can be regulated. In the absence of a fully independent international mechanism, CSOs in Sri Lanka can work with other CSOs in the Asian region and can begin to emulate such a mechanism through the creation and emulation of such a council in Asia.

The existing regional networks such as the Asia Democracy Network (ADN) and the Asia Democracy Research Network (ADRN) can be a starting point for such an initiative.²⁵⁶

iii. Observations of two options that require multiple stakeholders

The following section provides an assessment of two cross-cutting options within the formal and alternative regulatory framework which require joint intervention from multiple stakeholders.

a. Formal Regulatory Framework

(6) Creating more awareness about the ICCPR Act

More resources need to be invested towards training and raising awareness among law enforcement officials on the ambit and scope of the ICCPR Act and its limitations. This option on creating more awareness of the ICCPR Act requires cohesive action from law enforcement, institutions, CSOs and Media. Some of these institutions, given the changes stemming from the 20th Amendment to the Constitution, may face challenges of politicisation. Therefore, a multi-stakeholder approach to raising awareness on the ICCPR Act could be the most appropriate way forward.²⁵⁷

For instance, institutions such as the Human Rights Commission of Sri Lanka (HRCSL) and the National Police Commission (NPC), which have a mandate in this regard, could take the necessary initial action for such initiatives.

252. Website of ARTICLE 19, 'Malaysia: Anti-Fake News Act', 24 April 2018, at <https://www.article19.org/resources/malaysia-anti-fake-news-act/> [last accessed 22 April 2020].

253. Human Rights Council, *Mandate of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, 1/2017, 1 June 2017, at <https://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf> [last accessed 19 April 2020].

254. Yiswaree Palansamy, 'MIC urges govt to reinstate Anti-Fake News Act after influx of false news related to Covid-19', MalayMail, 31 March 2020, at <https://www.malaymail.com/news/malaysia/2020/03/31/mic-urges-govt-to-reinstate-anti-fake-news-act-after-influx-of-false-news-r/1852041> [last accessed 16 December 2020].

255. *Self-regulation and 'hate speech' on social media platforms*, (2018), *op. cit.*, p. 20.

256. Website of Asia Democracy Network, 'About Us', at <https://adnasia.org/home/about/about-us/> [last accessed 15 October 2020].

257. For further details on the changes from 19th Amendment to the 20th Amendment, see Verité Research, *Strengthening the presidency: Weakening the democracy: A brief analysis of the Twentieth Amendment to the Constitution of Sri Lanka* (22 January 2021), at <https://www.veriteresearch.org/publication/20a-brief-analysis-sri-lanka/> [last accessed 25 June 2021].

The mandate of the NPC includes providing training in order to improve the efficiency and independence of the police service.²⁵⁸ The HRCSL is vested with the power to advise and assist the government in formulating legislation and issuing administrative directives and procedures which ensure the protection of fundamental rights. Additionally, it also has the authority to make recommendations to the government on the necessary measures to ensure that national laws and administrative practices are in accordance with international human rights standards.²⁵⁹

Moreover, it is important to raise awareness among the judiciary on the Act, its application and the international best practices in order to ensure effective judicial interventions to mitigate the impact of any misapplication of the law by law-enforcement authorities. Institutions like the Judicial Service Commission, which has the mandate to make rules regarding the training of High Court Judges and judicial officers,²⁶⁰ and the Judges Training Institute, which is vested with the powers to facilitate such training, can play a significant role in taking the lead in relation to such training initiatives.²⁶¹

CSOs can advocate for the proper application of the ICCPR Act and can publicly hold law enforcement accountable in instances of misapplication by highlighting these instances to the public. CSOs can also partner with relevant institutions to provide capacity building and training to law enforcement on the ambit and scope of the ICCPR Act. The mainstream media plays a large role in how Sri Lankans receive their news and confers legitimacy on government actions.²⁶² Therefore, better understanding of and reporting among journalists on the proper application of the ICCPR Act can help erode the legitimacy and tendency for the misapplication of the Act by law enforcement authorities.

258. Article 155 G (3), The Constitution of the Democratic Socialist Republic of Sri Lanka 1978.

259. Section 10, Human Rights Commission of Sri Lanka Act, No. 21 of 1996.

260. Article 111H (2) (a), The Constitution of the Democratic Socialist Republic of Sri Lanka 1978.

261. Section 5, Sri Lanka Judges Institute Act, No. 46 of 1985.

262. Website of International Media Support, 'Consuming News in Turbulent Times', February 2021 at <https://www.mediasupport.org/publication/consuming-news-in-turbulent-times/> [last accessed 16 March 2021].

b. Alternative Regulatory Framework

(7) Investing in citizen-driven initiatives that can draw attention to and counter online hate speech and disinformation using international human rights and ethical standards²⁶³

It is valuable to have citizen-driven initiatives for combatting online hate speech and disinformation, such as fact-checking platforms and platforms such as Ethics Eye (for examples see Table 3: Fact-checking platforms - Mandate and followers). Such content can be published on a myriad of social media platforms to maximise reach. Such initiatives may also be vital to notifying relevant state institutions tasked with responding to harmful online content.

Meanwhile, investing in improving digital literacy in the country can greatly benefit the long-term operation and success of such citizen-driven initiatives. This could be done by the Ministry of Education, by effectively engaging CSOs and other relevant stakeholders, in order to create a special curriculum on digital literacy and e-safety. This would promote responsible online behaviour at an early age. For instance, Sarvodaya-Fusion in partnership with Facebook, with the support of the Ministry of Education and Information and Information Communication Technology Agency of Sri Lanka (ICTA) launched the 'We Think Digital' digital literacy program in July 2020.²⁶⁴ The program aims to improve digital skills, and online responsibility among groups comprising of youth, students, parents and teachers.

(8) Forming a national civil society coalition to build digital literacy²⁶⁵

Studies conducted on the sustainability of CSOs in Sri

263. This need to increase and advance different types of citizens driven initiatives was a sentiment expressed during the outreach sessions, specifically those held in Kandy.

264. Website of Daily FT, 'Facebook launches 'We Think Digital' with Sarvodaya-Fusion to build digital literacy skills', at <http://www.ft.lk/business/Facebook-launches-We-Think-Digital-with-Sarvodaya-Fusion-to-build-digital-literacy-skills/34-702576> [last accessed 22 October 2020].

265. The need to possibly establish a national CSO coalition was a sentiment and suggestion received at the outreach session during September 2020 – December 2020. It was stated that through such a CSO coalition network, groups (especially grassroots level organisations) that advocate against hate speech and disinformation would benefit in advancing their work and would have a safe space to voice their concerns.

Lanka reveal that although the sectorial infrastructure available to CSOs is average, there is a lack of resource centres available.²⁶⁶ CSOs that operate outside of the Colombo district for instance are affected by this resource insufficiency. The creation of a national coalition to boost digital literacy would provide for increased access to resources for CSOs outside the Colombo district. For instance, many CSOs outside of Colombo may have an increased need for access to knowledge, whereas Colombo based CSOs may need support to access networks and people to boost digital literacy in other parts of Sri Lanka.

Therefore, to improve coordination it could be beneficial for CSOs to create a strong network at a national level. This could be done through CSOs and other relevant activists forming a coalition to enhance and strengthen these citizen-driven educational initiatives for combatting hate speech and disinformation through digital literacy.

The following table provides a summary of the eight key options:

266. Sierra Amarasiri and Shamana Amjah (Verité Research), *2019 Civil society Organisation Sustainability Index: Asia Region: Sri Lanka Country Report* (August 2020), at <https://www.veriteresearch.org/publication/cso-sustainability-index-2019-sri-lanka-country-report/> [last accessed 15 October 2020].

Table 5: Summary of eight key options

Stakeholder	Frame-work	No.	Options:
Legislative Interventions	Formal Regulatory Framework	1.	<ul style="list-style-type: none"> Consider amending section 3(1) of the ICCPR Act.
		2.	<ul style="list-style-type: none"> Consider amending the provisions in the Penal Code to clarify the scope of offences provide under sections 120, 291A and 291B.
		3.	<ul style="list-style-type: none"> Consider an extension of SLCERT’s mandate through a transparent and structured system to evaluate content on social media in relation to hate speech and disinformation and request removal of harmful content. This expansion of SLCERT’s mandate should be accompanied with adequate checks and balances to ensure protection of freedom of speech/expression.
		4.	<ul style="list-style-type: none"> Assessing the possibility of introducing intermediary liability in Sri Lanka. However, such laws, if considered, must be drafted without undue restrictions on the freedom of speech given the comparative experience.
CSO-led Interventions	Alternative Non-Legal Frameworks	5.	<ul style="list-style-type: none"> Consider advocating for the formation of an independent international council to engage with the public and apply media ethics to regulate social media.
Cross-Cutting Interventions	Formal Regulatory Framework	6.	<ul style="list-style-type: none"> CSOs could advocate for the proper application of section 3(1) of the ICCPR Act by law enforcement and partner with relevant institutions for training. Institutions such as HRCSL, NPC, Judicial Service Commission, and Judges Training Institute could provide capacity building and training to law enforcement on the ambit and scope of the ICCPR Act. Media outlets should ensure media ethics are followed in relation to reporting.
			Alternative Non-Legal Frameworks
	Alternative Non-Legal Frameworks	8.	
		Alternative Non-Legal Frameworks	

ABOUT DEMOCRACY REPORTING INTERNATIONAL

Democracy Reporting International (DRI) is a non-partisan, independent, not-for-profit organisation registered in Berlin, Germany. DRI promotes political participation of citizens, accountability of state bodies and the development of democratic institutions world-wide. DRI helps find local ways of promoting the universal right of citizens to participate in the political life of their country, as enshrined in the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights.

<http://www.democracy-reporting.org>



Democracy Reporting International (DRI) has produced this study with the support of Strengthening Reconciliation Processes in Sri Lanka (SRP). SRP is co-financed by the European Union and the German Federal Foreign Office and implemented by Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH and the British Council, in partnership with the Government of Sri Lanka. The contents of this publication are the sole responsibility of Democracy Reporting International and can in no way be taken to reflect the views of the Strengthening Reconciliation Processes in Sri Lanka (SRP).



This study was commissioned by DRI in partnership with SRP, and was produced by Verité Research. Verité Research aims to be a leader in the provision of information and analysis for negotiations and policy making in Asia, while also promoting dialogue and education for social development in the region. The firm contributes actively to research and dialogue in the areas of economics, sociology, politics, law, and media, and provides services in data collection, information verification, strategy development, and decision analysis.

