



SEEK WISDOM, ELEVATE YOUR INTELLECT AND SERVE HUMANITY!



**ADDIS ABABA UNIVERSITY
COLLEGE OF LAW AND GOVERNANCE STUDIES
SCHOOL OF LAW LL, M PROGRAM**

Master of Laws (LL, M) in Human Right Law

**Adequacy of the Prevailing Regulatory Framework Relating to Hate Speech
on Social Media in Ethiopia**

Rahwa Weldeghebriel

Supervisor: Mesenbet Assefa (PhD)

**A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree
of Master of Law (LLM) in Human Rights**

December 2020

**Adequacy of the Prevailing Regulatory Framework Relating to Hate Speech
on Social Media in Ethiopia.**

Rahwa Weldeghebriel

Supervisor: MesenbetAssefa (PhD)

**A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree
of Master of Laws (LLM) in Human Rights**

December 2020

Approval Sheet by Board of Examiners

Rahwa Weldeghebriel’s thesis, entitled as “**Adequacy of the Prevailing Regulatory Framework Relating to Hate Speech on Social Media in Ethiopia.**” is approved by the undersigned members to the examining board.

Board of Examiners

Signature

Advisor

Examiner

Examiner

Acknowledgment

First and foremost, praise and gratitude to God. I am forever grateful for your immeasurable blessings in my life, only you know what is right for me, and I have faith in you.

I would like to express my deep and sincere gratitude to my advisor Dr. Mesenbet Assefa for his support and Encouragement. And I would also like to thank Dr. Biruk Haile for introducing me to this interesting topic as well as for his generosity. I would like to extend my special thanks to Yohannes Eneyew for his extraordinary cooperation and assistance by providing me valuable feedback and advice.

My friends Soreti Berhanu, Eskedar Taddesse, Seblwegel Tamiru, and Mesalea Abraham thanks for giving me beautiful memories that I am going to cherish for a lifetime. My life has been blessed for knowing you.

To my beloved parents and especially my Grandma, I can not thank you enough for your endless love and prayer throughout my life. Thank you for always believing in me, even when I didn't really believe in myself.

Most importantly, I would like to thank my caring and loving husband, Tewodros Kebede, and my beautiful boys, Abel Tewodros, Kidus Tewodros, and Barkon Tewodros, for their endless love and understanding. Thank you, I will be grateful forever for your love.

List of Abbreviations

ACHPR	The African Charter on Human and Peoples' Rights
ACHR	American Convention on Human Rights
CDA	The Communications Decency Act
COE	The Council of Europe
CPPCG	The Convention on the Prevention and Punishment of Crime of Genocide
ECHR	European Convention on Human Rights and Fundamental Freedoms
ECtHR	The European Court of Human Rights
FDRE	Federal Democratic Republic of Ethiopia
HsDPSP	The Hate Speech and Disinformation Prevention and Suppression Proclamation
ICERD	International Convention on the Elimination of All Forms of Racial Discrimination
ICCPR	International Covenant on Civil and Political Rights
ICTR	International Criminal Tribunal for Rwanda
RPA	The Rabat Plan of Action
SMSPs	Social Media Service Providers
NetzDG	Network Enforcement Act
UDHR	Universal Declaration of Human Rights
UN	United Nations

Abstract

Over the years, the abuse of social media, including the spread of different forms of hate speech, has fomented ethnic conflicts and violence's in Ethiopia. The Ethiopian government has passed a comprehensive law called Hate Speech and Disinformation Prevention and Suppression Proclamation (HsDPSP) to combat hate speech both online and offline. This thesis seeks to assess the adequacy of the Proclamation to regulate hate speech circulated through social media in Ethiopia. This study has examined the legislation in light of international and regional human rights instruments, as well as pertinent domestic laws. The study reveals that while the government's measure to regulate hate speech is progressive, the hate speech legislation fails to strike a proper balance between freedom of expression and hate speech, both in content and procedure which may pose a risk of excessive regulation short of international standards.

Key Terms: Freedom of Expression, Hate Speech, Social Media, Hate Speech and Disinformation Prevention and Suppression Proclamation

Table of Contents

Page

Acknowledgment	iii
List of Abbreviations.....	vi
Abstract.....	vi
Table of Contents.....	vi
1. INTRODUCTION.....	- 1 -
1.1 Background of the Study	- 1 -
1.2. Statement of the Problem	- 5 -
1.3 Research Objectives	- 6 -
1.5 Scope and Limitation.....	- 7 -
1.6 Significance of the Study	- 8 -
1.7 Research Methodology.....	- 8 -
1.8 Literature Review	- 8 -
1.9 Organization of the Thesis	- 10 -
CHAPTER TWO	- 11 -
SOME CONCEPTUAL FRAMEWORKS OF HATE SPEECH	- 11 -
2.2. Historical Overview of Hate Speech.....	- 11 -
2.3 Regulating Hate Speech.....	- 13 -
2.3.1 Arguments for Hate speech Regulation.....	- 13 -
2.3.2 Arguments Against Hate Speech Regulation	- 14 -
2.4. Legal Frameworks of Hate Speech	- 15 -
2.4.1 International Human Rights Legal Framework.....	- 15 -
2.4. 2 The Regional Human Rights Frameworks	- 18 -
A. The European Convention on Human Rights (ECHR)	- 19 -
B. The African Charter on Human Rights and Peoples' Rights (ACHRP).....	- 23 -
C. The American Convention on Human Rights (ACHR).....	- 24 -
2.5. Freedom of Expression, the Right to Equality Versus Hate Speech.....	- 25 -

2.6. A Proposed Typology of Hate Speech.....	- 27 -
A. Hate speech that Must be Prohibited (Top Level of Hate Speech);	- 27 -
B. Hate Speech That May be Prohibited (Intermediate Level of Hate Speech)	- 28 -
C. Lawful Hate Speech (Bottom Level of Hate Speech):.....	- 28 -
Figure 1. The ‘Hate Speech Pyramid’	- 29 -
2.7.Hate speech vs Hate crime	- 30 -
2.8 Hate speech on Social Media	- 30 -
CHAPTER THREE	- 32 -
REGULATING HATE SPEECH UNDER NATIONAL LAWS.....	- 32 -
3.1. Introduction.....	- 32 -
3.2 The National Legal Frameworks.....	- 32 -
3.2.1. Legal protection of the Right to Freedom of Expression on Social Media	- 32 -
3.2.2 The Legal Protection on the Right to Equality	- 33 -
3.2.3. Ethiopia's Legal Framework for Tackling Social Media Speech	- 34 -
3.3 Lessons from other jurisdictions Relating to Social Media Hate Speech Regulation.....	- 36 -
3.3.1 Germany:.....	- 36 -
Enforcement on Social Networks '(NetzDG) 2017.....	- 36 -
3.3.2 The United State of America	- 38 -
The Communication Decency Act (CDA) 1996.....	- 38 -
CHAPTER FOUR.....	- 40 -
Assessing The Adequacy of the HsDPSP to Regulate Hate Speech on Social Media	- 40 -
4.1 Introduction	- 40 -
4.2 The duty of the Ethiopian government to tackle hate speech on social media.....	- 40 -
4.3 The Hate Speech and Disinformation Prevention and Suppression Proclamation (HsDPSP)	- 42 -
4.3.1 Compatibility of the Proclamation with International human rights law Standards on Freedom of Expression	- 43 -
4.3.2 The scope of protected groups.....	- 48 -

4.3 .4 Legal Enforcement Towards Social Media Service Providers	- 50 -
4.3. Procedural safeguards.....	- 53 -
Conclusion and Recommendations.....	- 55 -
5.1 Conclusion.....	- 55 -
5.2 Recommendations	- 57 -
BIBLIOGRAPHY	- 58 -

CHAPTER ONE

1. INTRODUCTION

1.1 Background of the study

In today's world, it is undeniable that social media plays a vital role in influencing our culture, our economy, and our overall outlook in the world.¹ Globally, the landscape of communications was changed further when social media platforms began to be used well in the 1990s.² Social media use mobile and web-based technology to build highly engaging platforms in which individuals and groups share, co-create, discuss, and change user-generated content³. Social media platforms include Facebook, YouTube, Instagram, Twitter Telegram LinkedIn, Viber, Snapchat, Tiktok, and other similar platforms. Social media platforms simply became intermediaries between two groups of people private users who share the content and the audience, who receives them. The means of communication became accessible to a large number of people.

Social media is a nascent medium that brings people to share information, participate, and unite for a cause, seek advice, and provide guidance⁴ as well as a lot of business entities use social media to generate and connect to business consumers. Social media eliminated contact barriers and built a decentralized communication system and left the door open for everyone to have a voice and to take part and participate freely including people in repressive countries.

However, this digital technology has come both with a blessing and a curse, as this low-cost and high-speed dissemination mechanism has also created new platforms for criminal and extremist activity. The Social media age has contributed to the spread of hate speech and the spread of

¹ Jacob Amedie, 'The Impact of Social Media on Society' (2015) Pop Culture Intersections. 2.

² Christopher McFadden, 'The Chronological History of Social Media' (2010) IE <<https://interestingengineering.com/a-chronological-history-of-social-media> accessed > Accessed 21 Oct 2020

³ Jan H. Kietzmann, and et al, 'Social Media? Get serious! Understanding the Functional Building Blocks of Social Media ' (2011) Business Horizons 54(3) 241

⁴ Jyoti Suraj Harchekar, 'Empact of Social Media on Society ' (2017) IJERT Vol. 6 381

cyber hatred activities, with an increase in the number of websites providing racist content and hate speech recently recorded.⁵

Social media platforms also offer violent actors the opportunity to publicize their acts. It provides an opportunity for radical groups to find like-minded individuals, to create collective identity and cohesion for a certain ideological viewpoint, to further connect and collaborate. The anonymity, immediacy, and global nature of the online platform have also made it an ideal tool for extremists and hawks to foster hatred.⁶ Therefore, the potential to influence the globe through social media platforms has enhanced, and as it has been argued; today's computer keyboards could be even more dangerous than tanks and machine guns.⁷

Jeremy Waldron describes two harmful forms of hate speech messages that expose different groups to vulnerability. The first message is aimed at victims to dehumanizing or degrading them and making them feel marginalized in society⁸ this can be exemplified by the COVID-19 pandemic related hate speech, after the outbreak of the pandemic, individuals perceived as ethnically Chinese or Asian, or belonging to certain ethnic and religious minorities, migrants, and foreigners have been blamed and vilified for spreading the virus and due to the rumors and misinformation marginalized populations have also been recorded⁹. The second one is to reinforce a sense of an in-group that is (purportedly) under threat.

On the other hand, the right to freedom of expression is a fundamental human right recognized in all major human rights systems as well as in domestic legislation. This fundamental right includes the freedom to seek, receive, or share information or ideas of any kind, regardless of borders, through any media of a person's choice.¹⁰ At the same time, it is not an absolute right, it

⁵ F .Cassim , 'Regulating Hate Speech And Freedom Of Expression On The Internet: Promoting Tolerance and diversity ' (2015) SACJ 3, 303

⁶ James Banks , 'Regulating Hate Speech Online ' (2010) I R LCT 233

⁷ Lashel Shaw, 'Hate Speech in Cyberspace: Bitterness without Boundaries' (2012), Notre Dame Journal of Law, Ethics & Public Policy, 282

⁸ Jeremy Waldron ,'The Harm In Hate speech' (2012) London: Harvard University Press 2-3.

⁹ UN, 'United Nations Guidance Note on Addressing and countering Covid-19 Related Hate speeches',2020

¹⁰ International convention on civil and political rights (ICCPR) 1966 and Universal Declaration of Human Rights (UDHR), 1948 Art 19, African charter on human and peoples right (ACHPR), 1982 Art 9, American Convention

can be curtailed to protect national security, public order, public health, or morals; and, respecting the rights or reputations of others.¹¹ Thus some laws contain provisions that provide for balancing freedom of expression against those other interests in the particular context of hate speech. For instance, Article 20(2) of the International Covenant on Civil, and Political Rights (ICCPR) prohibits the incitement of discrimination, hatred, or violence and requires to be punishable by law, and Article 4 of the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) further requires legislation which not only prohibits such incitement but also makes 'an offense punishable by law.'¹²

Under international law, there is no universally accepted definition of hate speech; and to build a globally unified framework that addresses hate speech; the United Nations Strategy and the Hate Speech Guidance Action Plan, a non-binding document, defined the term hate speech as

"Any form of communication in speech, writing or behavior that attacks or uses a pejorative discriminatory language with reference to a person or group based on which they are, that is based on their religion, ethnicity, nationality, race, color, descent, gender or other identity factors."¹³

Besides, the Rabat Plan of Action was developed by international experts to better direct all stakeholders, including the national legislature and the judiciary, in the enforcement of the international obligation of prohibition incitement to hate under Article 20(2) of ICCPR. The instrument provides the six-part threshold test, bearing in mind (1) the social and political context, (2) the position of the speaker, (3) the intention to incite the audience against the target

on Human Rights (AmCHR), 1969 Art 13, European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR),¹⁹⁵ Article 10 and fConstitution of the Federal Democratic Republic of Ethiopia (FDRE), 1995 Art 29

¹¹ ICCPR Art. 19(3)

¹² Ibid Art 20 (2) and International Convention on the Elimination of All Forms of Racial Discrimination (ICERD),1965 Art 4

¹³ UN,United Nations Strategy and Plan of Action on Hate Speech: Detailed Guidance2020

group, (4) the content and form of the expression, (5) the degree of its propagation and (6) the probability of harm, including imminence.¹⁴

Social media platforms need to moderate hate speech while promoting freedom of expression. As the Social media platform has been used both to exercise the fundamental right to freedom of speech and to spread hatred at the same time many countries have enacted social media regulatory legislation to strike a balance between freedom of expression and group rights, as well as concepts of dignity, liberty, and equality. For example, countries like Canada, Germany, and France have further enacted special laws to halt hate speech online beyond general hate speech laws.¹⁵ There are also several other ways to minimize hateful content or its negative impacts by encouraging people to avoid uploading such content in the first place, which is a preventive method rather than a reactive one.

According to the latest Digital 2020 survey, there are 21,14 million Internet users in Ethiopia with an Internet penetration of 19 %. Among these 6,20 million are social media users, the number of users of social media increased by 237 thousand as of April 2019.¹⁶ Even though the country's Internet penetration is even less than half the African average, hate speech and misinformation have become a major and growing problem in Social media platforms. Ethiopia, as a State party to the ICCPR and ICERD which require States to enact legislation banning hate speech, has an obligation to regulate offline and online hate speech by laws that are consistent with international standards. The research is motivated by the need to assess the adequacy of the existing frameworks for the regulation of hate speech in social media with respect to the country's international obligations, international standards of freedom of speech, and the principle of non-discrimination.

¹⁴ OHCHR /The Rabat Plan of Action (RPA), 'Freedom of expression vs incitement to hatred' <https://www.ohchr.org/en/issues/freedom_of_opinion/articles19-20/> accessed 10 Oct 2020

¹⁵ Alexander Thesis, 'Hate in Cyberspace: Regulating Hate Speech on the Internet' (2001), 38 SAN DIEGO L. REV. 817,

¹⁶ DATAREPORTAL, 'Digital 2020 Ethiopia' <https://datareportal.com/reports/digital-2020-ethiopia> accessed 4 Mar 2020

1.2. Statement of the problem

Ethiopia is a multi-ethnic nation with diversified culture, language, religion, and ethnicity; such vibrant culture of tolerance and respect between peoples must be built up. But ethnic conflicts and tensions have recently become major challenges across the country. As a result of these conflicts and inter-communal violence, around 2.9 million people have been displaced from their habitual residence and 68,000 new displacements have occurred in the first half of 2020.¹⁷

In previous years, speeches by government officials, activists, and others in Ethiopia have spread quickly through social media and helped trigger or fuel violent conflict in the country. There is a consensus that the rise in irresponsible use of social media, such as the use of a platform to spread hate speech and disinformation, has inflamed ethnic violence in various parts of the country. Ethiopian Prime Minister Dr. Abiy Ahmed also admits in his speech accepting the Nobel Peace Prize that social media is now being used to sow "hate and division" and undermine the delicate political transition in Ethiopia.¹⁸

Thus, the government has been asked by national and international human rights experts to take adequate measures that can effectively address the issue of hatred. For example, David Kaye, the U.N.'s special rapporteur on the right to freedom of opinion and expression regarding the country's tensions between the various ethnic groups urges the Government to develop a diverse culture of tolerance and respect and to consider the lessons of Resolution 16/18 (2011) of the Human Rights Council, which calls for tolerance and meaningful action to address hatred.¹⁹

¹⁷ IDMC, 'Global Report on Internal Displacement' <https://www.internal-displacement.org/countries/ethiopia> accessed 1 Oct 2020

¹⁸ Abdi Latif Dahir, Nobel Peace Laureate Says Social Media Sows Hate in Ethiopia, *The New York Times*, 10 Dec 2019 < <https://www.nytimes.com/2019/12/10/world/africa/nobel-peace-abiy-ahmed.html> > accessed 5 Oct 2020

¹⁹ David Kaye, Visit to Ethiopia, 2-9 December 2019 End of mission statement. Ohchr.org <<https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25402&LangID=E>> accessed 8 Oct 2020

In Ethiopia, both in online and offline conversations, there are so many forms of hate speeches.²⁰ There are several political, ethnic, and derogatory terminologies in social media conversations. For example: ‘*Sefari*’ or ‘*Mettie*’ refer to ‘settlers’ and ‘newcomers’ in other regions. The use of one or more of these inflammatory expressions would give rise to anger, and at times exclusion, internal displacement, and discrimination.

Concerning the regulation of online hate speech, it is encouraging that the government has recently adopted a special law called Hate Speech and Disinformation Prevention and Suppression Proclamation (HsDPSP) No 1185/2020. In fact, under Article 20(2) of ICCPR and Article 4 of ICERD, the Government is obliged to prohibit by law advocacy of national, racial, and religious hatred that constitutes incitement to discrimination, hostility, and violence. However, that does not mean that the law on hate speech could be adopted without limitation, because the instruments referred to above, which allowed the state to prohibit hate speech by law, required any hate speech law to conform with international standards on freedom of expression and also adhere to non-discrimination provisions.²¹ And this study assesses the substantive and procedural gaps within the Ethiopian hate speech proclamation that cripple the regulation of hate speech on social media in the light of the State's international obligations.

1.3 Research Objectives

General objective

The general objective of the research is to review and assess the adequacy of Ethiopia's prevailing regulatory legal framework (Hate Speech and Disinformation Prevention and Suppression Proclamation) to tackle social media hate speeches.

The specific objectives of the study are

- To assess the compliance of the content of the prevailing regulatory framework with expected international standards on the right to freedom of expression;

²⁰ Elias Meseret, Hate speech and disinformation concerns escalate in Ethiopia, *Devex*, 06 May 2020 <https://www.devex.com/news/hate-speech-and-disinformation-concerns-escalate-in-ethiopia-97095?fbclid=IwAR1ZUqpjSMQD6z_cFh0kCjXcGF9AWAfL3CPR7S6mBUDgbGEJbD62SaUqbnY> accessed October 21, 2020.

²¹ ICCPR, Human Rights Committee, General comment No. 34 Article 19: *Freedoms of opinion and expression*, (2011) para 26

- To examine the compatibility of the hate speech law's scope of protection with the international principle of non-discrimination;
- To assess the legal enforcement and feasibility of the obligations imposed on social media service provider services ;
- To identify the procedural gaps of the existing hate speech regulatory law that cripple combating social media hate speech;

1.4 Research Questions

Principal research question :

How adequate is Ethiopia's prevailing legal framework to regulate social media hate speeches in compliance with international standards?

In order to answer this principal research question, the study will address the following specific questions

- Whether the scope of protection under the prevailing hate speech law is consistent with the international principle of non-discrimination?
- Whether the content of the Hate Speech and Disinformation Prevention and Suppression Proclamation is compatible with international standards on freedom of expression?
- Are the obligations imposed on social media service providers legally enforceable and feasible?
- What are the procedural limitations in the prevailing hate speech regulatory framework to counter social media hate speech?

1.5 Scope and Limitation

The focus of the study is limited to examining the adequacy of Ethiopia's hate speech regulatory frameworks to combat hate speech on social media. Legal analysis will be carried out in light of the applicable national laws and ratified regional and international human rights documents and as well as soft laws and guidelines. Owing to the restricted reach of the thesis, it is not possible to address various approaches to the regulation of hate speech; the focus of this analysis is on legal regulation.

1.6 Significance of the Study

Hate speech has become a pressing issue in Ethiopia worth proper regulation. This thesis is being undertaken to answer the question of how legal frameworks are fit to address social media hate speech. Thus, by identifying the gaps in the legal framework, this research will provide insights for legislative and policymakers to address social media hate speech in line with Ethiopia's international obligation arising from the ratification of the international human rights instruments. It will also enrich the knowledge of the law on the regulation of social media hate speech and also serve as a stepping stone for more academic research on the subject.

1.7 Research Methodology

To address the research questions, the study will use the doctrinal method of legal research approach. Initially, the study examines all relevant domestic legislation relating to the regulation of hate speech, which is the constitution, criminal code, computer crime proclamation No.958/2016 and Telecom Fraud Offence Proclamation No.761/2012 then, with a particular emphasis on the regulation of social media hate speech under the special law, i.e Hate Speech and Disinformation Prevention and Suppression Proclamation. The research will review these laws in light of applicable international and regional human rights binding instruments. It will also explore other jurisdictions' experiences. Also, books and articles from secondary sources in both print and electronic formats will be reviewed. To obtain the necessary data and information, the website sources will also be consulted.

1.8 Literature Review

Hate speech, both online and offline, has become a pressing issue because it has exacerbated social and ethnic divisions between different groups. Thus the subject receives prevalent international attention from both legal and non-legal researchers. Although legal research on the regulation of hate speech in social media is scanty in domestic literature, there are various studies on freedom of expression as well as the regulation of hate speech from various points of view.

For instance, Gedion [2010]²² argues the gap in Ethiopia's jurisprudence concerning freedom of expression has adverse effects for the enjoyment of the right, for the dignity of the bearers of the right and the establishment of democratic order and culture. Moreover, Mesenbet [2016]²³ argues freedom of expression offers the most important way by which individuals can actively participate entirely in the political life of the community, but the draconian laws and regulations, like the 2009 of Anti-Terrorism and Charities and Associations Proclamation, and the 2008 Mass Media and Access to Information Proclamation, have basically demolished political pluralism and multi-party democracy, as well as drowned independent media outlets.

Yared Legesse [2012]²⁴ argues though equality demands that certain marginalized ethnic groups be protected from hate speech, there is considerable difficulty in regulating ethnic hate speech when ethnicity is a primary medium for political organization and mobilization. The paper correlates with the challenge of regulating hate speech on the ground.

Baraki Belay [2019]²⁵ argues that the scattered provisions found in various legislation before the new specific hate speech law can be used to prosecute online hate speech. The study is limited to the prosecution of hate speech offenses under the scattered provisions of general law, such as the Criminal Code and the Law on Computer Crime, but the new specific hate speech law is outside the scope of the study.

Yohannes Eneyew [2020]²⁶ has assessed the limiting measures towards freedom of expression on the internet including internet shutdowns, hate speech and disinformation regulation, draconian

²² Gedion Timothewos, 'Freedom of Expression in Ethiopia: The Jurisprudential Dearth' (2010) *Mizan Law Review* Vol. 4 No.2, 201

²³ Mesenbet A Tadege 'Freedom of Expression and the media landscape in Ethiopia: contemporary challenges' (2016) *Henonline* 5 U. Balt. J. Media L. & Ethics 69

²⁴ Yared Legesse Mengistu, 'Shielding Marginalized Groups from Verbal Assaults Without Abusing Hate Speech Laws' in Michael Herz and Peter Molnar (eds) *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge University 2012) 352

²⁵ Baraki Belay, 'Prosecution of criminal behaviors committed Over the social media in Ethiopia: Acts of Hate speech' in *Focus* (Gonder University Bachelor of Law Thesis 2019)

²⁶ Y Eneyew Ayalew 'Assessing the limitations to freedom of expression on the internet in Ethiopia against the African Charter on Human and Peoples' Rights' (2020) 20 *African Human Rights Law Journal* 315-345

national laws, and internet censorship. But the assessment of measures, including the law on hate speech, was mainly focused on the context of the Banjul Charter.

This thesis assesses the adequacy of the existing legal framework to regulate hate speeches on social media in compliance with international standards which is outside the domain of the above literatures.

1.9 Organization of the Thesis

This study is divided into five separate sections; the first section will be an introductory part providing the background, statement of the problem, research objective, scope and limitation of the research, and the research methodology. The second section will cover a literature review with a focus on exploring the legal framework for the limitation of freedom of expression, as well as the regulation on hate speech. The third section will discuss national laws on the regulation of hate speech and other selected foreign experiences linked to the regulation of online hate speech. The fourth section will analyze the gap in the prevailing recent hate speech regulator framework and areas of improvement to properly regulate social media hate speeches. Finally, the fifth section will summarize the main findings of this study.

CHAPTER TWO

SOME CONCEPTUAL FRAMEWORKS OF HATE SPEECH

2.1 Introduction

This chapter aims to discuss certain conceptual frameworks of freedom of expression as well as hate speech. It includes an overview of the relevant frameworks under international and regional human rights treaties.

2.2. Historical Overview of Hate Speech

No one has been born hating or fearing others, but political leaders have used various types of rhetoric to systematically turn groups of people against each other throughout human history, vilifying and degrading others.²⁷ Nazi regime is a prime example of how Jews were portrayed as an unwelcome and alien race, through propaganda which inciting hatred. And the demand to curtail hate speech is inherently related to the roots of contemporary international human rights discourse, to the post-Holocaust era, and crimes committed on a large scale during the Second World War²⁸, which lasted from 1939 to 1945.²⁹

After the Second World War, a series of military tribunals were created, the Nuremberg Trials, in which prominent members of the political, military, judicial, and economic leadership of Nazi Germany were prosecuted.³⁰ In the case against Julius Streicher, the publisher of '*Der Stürmer*', a journal actively instigating the persecution and expulsion of the Jews, the International Military Tribunal (Nuremberg) ruled that Streicher guilty by stating that his incitement to murder and extermination at a time when Jews in the East were murdered under the most appalling

²⁷ Suzan Bensch, Catherine Buerger, Tonei Glavnic, 'Dangerous Speech: a Practical Guide', 2018

²⁸ Mari J. Matsuda , ' Public Response to Racist Speech: Considering the Victim's Story' (1989) 87 Mich.L.R ,2320

²⁹ Jennifer Rosenberg, Timeline of World War II From 1939 to 1945, *Thought Co*<<https://www.thoughtco.com/world-war-ii-timeline-1779991>> Accessed 23 Oct 2020

³⁰ Ibid

circumstances clearly constitutes persecution on political and ethnic grounds in accordance with war crimes as specified in the Charter and constitutes a crime against humanity.³¹

Similarly, the International Criminal Tribunal for Rwanda (ICTR) follows the same concept as Nuremberg in the Rwandan Genocide of the 1990s, which is believed to have killed one million civilians.³² In the case of Nahimana, a senior official of the "*Radio Television Libre des Mille Collines*" (RTLM), a radio station inciting the population against Tutsis, the ICTR found Nahimana guilty of overt and public incitement to commit genocide as well as persecution and extermination as crimes against humanity.³³

With regard to the legal framework relating to hate speech while the UDHR, the guarantor of human rights, ratified by the United Nations General Assembly on 10 December 1948, offers equal protection against any incitement to such discrimination³⁴ but the topic of hate speech was the subject of tough discussions between states during the drafting process. The member states have different views on the topic. There were three positions; generally speaking; the Soviet Union and its allies, who had a strong interest and concern in dealing with hate speech and war propaganda; the United States; that privileged free speech and other western democracies and developing countries who tried to hold the middle ground.³⁵ Besides, the contrasting world ideas arise again on the inclusion of hate speech provisions on the UN treaties;³⁶ Article 4 of ICERD and Article 20 of ICCPR which require state parties to prohibit speech conceived as advocacy to hate that constitutes incitement to discrimination, hostility, or violence.³⁷ Despite these

³¹ The Prosecutor v Streicher [1946] International Military Tribunal (Nuremberg), 1946[502] <<https://www.legal-tools.org/doc/45f18e/pdf>> accessed 23 Oct 2020

³² 'Say no to hate speech and xenophobia', urges Guterres, marking 26 years since genocide in Rwanda, UN News (April 2020) Human Rights <<https://news.un.org/en/story/2020/04/1061242>> Accessed 24 Oct 2020

³³ The Prosecutor v Ferdinand Nahimana, Jean-Bosco Barayagwiza Hassan Ngeze [2003] Case No ICTR-99-52, [1073] <http://www.worldcourts.com/ictr/eng/decisions/2003.12.03_Prosecutor_v_Nahimana.pdf> accessed 24 Oct 2020

³⁴ UDHR Art 7

³⁵ Johannes Morsink, 'The Universal Declaration of Human Rights: Origins, Drafting, And Intent' 66 (1999).

³⁶ Pesins Antonis 'The Regulation of "Hate speech"; the Meaning of "Incitement" Under the Case-law of European Court of Human Rights and the Jurisdictions of European Union; the United Kingdom and Greece' (2015) CEU 1

³⁷ ICCPR Art 20(2) and ICERD Art 4

controversial debates, the provisions survived on the conventions for the regulation of hate speech. But it is undeniable fact many state parties to the convention made a reservation on these provisions, which indicates their unwillingness to the realization and application of the provisions at the national level.³⁸

2.3 Regulating Hate Speech

Although the need to condemn and eliminate hate speech is indisputable³⁹, there are debates concerning its regulation.

2.3.1 Arguments for Hate speech Regulation

Scholars argue in favor of the regulation of hate speech from a different view of point. Primarily, it has been argued that hate speech should be regulated because it weakens democracy by reducing the tone of public discourse and also damages the dignity of members of the target population by stigmatizing them and denying them the right to live as responsible members of society.⁴⁰ Jeremy Waldron, the most influential advocate of the regulation of hate speech, firmly argues that hate speech laws provide the assurance that members of disadvantaged minorities should be able to rely on; thus, should be regulated as part of our commitment to human dignity and equality and respect for members of marginalized minorities.⁴¹ It is sometimes argued beyond the target community, banning hate speech reassures all members of society by stripping away the environment filled with hate that it helps to achieve a sense of shared humanity.

³⁸ United Nations Treaty Collection, International Convention on the Elimination of All Forms of Racial Discrimination, New York, 7 March 1966 <<https://treaties.un.org/doc/Publication/MTDSG/Volume%20I/Chapter%20IV/IV-2.en.pdf>>

³⁹ Ritika Patni; Kasturika Kaumudi, 'Regulation of Hate Speech', (2009) 2 NUJS L. Rev. 749

⁴⁰ Bhikhu Parekh, 'Is There a Case for Banning Hate Speech' in Michael Herz and Peter Molnar (eds) *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge University 2012) 37

⁴¹ Jeremy Waldron, 'Dignity and Defamation: The Visibility of Hate' (2010) Harv. L. Rev. 1596

2.3.2 Arguments against Hate Speech Regulation

Some free speech theorists argued that hate speech laws are likely to be manipulated by those in power, who would always be able to characterize their opponents' speech or politics as hate speech or similar and it is likely to minimize the democratic cultural self-understanding that disputes should be dealt with as political rather than violent struggles.⁴²

Moreover, it is also argued that the only way to resolve bad ideas is not by banning them, but by subjecting them to critical analysis and challenging them with better ideas; the reaction to hate speech is not less speech, but rather more.⁴³ Some further argue restrictions on hate speech can weaken the validity of anti-discrimination legislation, both in terms of widespread recognition and, more significantly, in terms of the morality of their implementation.⁴⁴ It is also argued the regulation of hate speech gives the State the right to judge the content of the speech and to determine what kind of speech is good or bad, and should or should not be permitted to do so as well as it is argued that the law itself can not alter people's perceptions and eradicate hate.⁴⁵

Irrespective of the arguments posed above, there has been an almost universal inclination towards regulating certain forms of speech. Many countries prohibit different forms of hate speech in their domestic law. For instance, Britain bans violent, offensive, and threatening speech; Denmark and Canada ban speech that is offensive and degrading; and India and Israel ban speech that incites ethnic and religious hatred or is likely to stir up animosity between groups. In the Netherlands, it is a criminal offense to make public remarks offending classes of people. Australia forbids speech that offends, insults, humiliates, or intimidates persons or groups,⁴⁶ and Ethiopia bans speech that promotes hate, discrimination, or attacks against an

⁴² C. Edwin Baker, 'Hate Speech' in Michael Herz and Peter Molnar (eds) *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge University 2012) 57

⁴³ Parekhsupra note 38

⁴⁴ James Weinstein, 'Hate Speech Bans, Democracy, and Political Legitimacy' (2017). *Constitutional Commentary*. 465

⁴⁵ Parekh supra note 38

⁴⁶ Ibid

individual or a discernible identity group based on ethnicity, religion, race, gender, or disability through broadcast, print or social media using text, image, audio or video.⁴⁷

2.4. Legal Frameworks of Hate Speech

Owing to the absence of a universally accepted definition of hate speech in international human rights law, it is hard to define it clearly and consistently. International and regional human rights instruments contain provisions focusing on various forms of expressions that would generally be considered to constitute hate speech.

2.4.1 International Human Rights Legal Framework

The word hate speech is not expressly specified in many international human rights instruments but it is implicitly invoked by some of the values of human dignity and freedom of expression. For instance, the initial international instrument UDHR does not explicitly name the term hate speech. However, this instrument implicitly carries the prohibition of hate speech, as they indirectly include clauses focusing on various forms of expression that would usually be viewed as constituting hate speech.⁴⁸

The non-binding document 1948 Universal Declaration of Human Rights (UDHR), which was drawn up as a response to the atrocities of World War II, incorporates the right to equal protection against any discrimination in violation of the declaration and against any incitement to such discrimination.⁴⁹ In the meantime, it recognizes the right to freedom of expression, which includes freedom to hold opinions without interference and to seek, obtain, and exchange information and ideas through any media, regardless of borders.⁵⁰ And taken these articles together, i.e can be understood that everyone has the right to free speech and the right to be

⁴⁷ Hate Speech and Disinformation Prevention and Suppression Proclamation 1185/2020 (HsDPSP), Federal Negarit Gazette 26th Year, Addis Ababa, 23 March 2020 Art 4 and 2(1)

⁴⁸ Iginio Gagliardone, 'Countering Online Hate Speech' (2015) UNESCO < https://www.researchgate.net/publication/284157227_Countering_Online_Hate_Speech_-_UNESCO > accessed Oct 26, 2020

⁴⁹ UDHR ART 7 and Art 19

⁵⁰ Ibid Art 19 and 20

protected against discrimination; that is, everyone has the right to be protected against hate speech in so far as such speech contains discriminatory objectives.⁵¹

In the same vein, the ICCPR, even though it does not make specific reference to the term hate speech, it imposes restrictions on freedom of expression which requires States to prohibit certain types of speech intended to sow hatred, that is, the advocacy of national, racial or religious hatred, which constitute an incitement to discrimination, hostility or violence.⁵² However, States are not necessarily obliged to criminalize those expressions.

There are also other instruments that contain provisions focusing on types of hate speech. For example, the Convention on the Prevention and Punishment of Crime of Genocide (CPPCG) forbids overt and public incitement to commit genocide.⁵³ Also, the Statute of the International Criminal Court (ICC) holds a person accountable and liable for a crime within the jurisdiction of the Court if that person directly and publicly incites others to commit a crime.⁵⁴

Unlike the above instruments the ICERD that was adopted by the UN General Assembly in 1965, is the first international treaty to deal directly with the issue of hate speech and requires states to render several types of expression punishable by law i.e States are obliged to punish those involved in the propagation of ideas based on racial superiority or hatred, incitement to racial discrimination, acts of abuse or incitement against any race or groups of people of a different color or ethnic origin.⁵⁵ It imposes more far-reaching responsibilities for States Parties, *inter alia*, to criminalize hate speech, to support racist activities, and to ban and criminalize membership in organizations that encourage and incite racial discrimination.

However, many state parties have reservations about this Article and perceive it as prohibiting measures that infringe freedom of expression and association. This manifests the unwillingness of the State Parties to be bound by such a clause. Also, certain states have already made an

⁵¹ Iginio Gagliardone Supra note 46

⁵² Ibid

⁵³ Convention on the Prevention and Punishment of Crime of Genocide (CPPCG) 1848 , Art 3(c)

⁵⁴ Rome Statute of the International Criminal Court (RSICC) 2000, Art 25

⁵⁵ CERD Art 4

interpretative declaration that the right to freedom of opinion and expression and the right to peaceful assembly and association can not be jeopardized by such implementation of hate speech legislative steps.⁵⁶ Accordingly, the Committee on the Elimination of Racial Discrimination clarifies that the prohibition on the propagation of all ideas centered on racial superiority or hatred is compatible with the right to freedom of opinion and speech enshrined in the Universal Declaration of Human Rights (Article 19) and Article 5(d) of ICERD.⁵⁷

In order to develop a common understanding of legislative patterns, judicial practices, and policies regarding the concept of incitement to national, racial, or religious hatred, the Rabat Plan of Action (RPA) was adopted in 2012 by experts, which includes guidelines for state parties on implementing their obligations under Article 20(2) of the ICCPR. It outlines a six-part test taking into account (1) the social and political context, (2) status of the speaker, (3) intent to incite the audience against a target group, (4) content and form of the speech, (5) extent of its dissemination and (6) likelihood of harm, including imminence. This "six-part test" can be used by the UN as a framework to consistently analyze the nature and assess the severity of any hate speech they can experience and to adjust their actions and interventions in response.⁵⁸

In addition, the Camden Principles on Freedom of Expression and Equality referred to in the Rabat Action Plan, have set out principles and recommendations to promote better consensus on the proper relationship between respect for freedom of expression and the promotion of equality to avoid misuse of the concept of hate speech.⁵⁹ It requires all states to enact legislation banning “any advocacy of national, ethnic, or religious hatred that constitutes incitement to discrimination, hostility, or violence.” The Camden Principles anchors an authoritative interpretation to describe the words "hatred", "advocacy" and "incitement" as;

⁵⁶ United Nations Treaty Collection, supra note 36

⁵⁷ CERD General Recommendation No. 15, *Measures to eradicate incitement to or acts of discrimination* (1993) para. 4

⁵⁸ OHCHR, Freedom of expression vs incitement to hatred: OHCHR and the Rabat Plan of Action <[https:// www.ohchr.org/EN/Issues/FreedomOpinion/Articles19-20/Pages/Index.aspx](https://www.ohchr.org/EN/Issues/FreedomOpinion/Articles19-20/Pages/Index.aspx)> accessed Nov 3 2020

⁵⁹ The Camden Principles on Freedom of Expression and Equality, Article 19 Global Campaign for Free Speech (2009)

- i. The words 'hatred' and 'hostility' refer to the strong and irrational feelings of opprobrium enmity, and hatred towards the target group.
- ii. The word 'advocacy' is to be understood as requiring an intention to publicly encourage hate towards the target group.
- iii. The word 'incitement' refers to statements about national, racial, or religious groups that establish an immediate risk of discrimination, aggression, or violence towards persons who belong to those groups.⁶⁰

The Durban Declaration and Programme of Action (DDPA)⁶¹, a soft law that is not legally binding, is also a detailed, action-oriented document that recommends specific steps to tackle racism, ethnic discriminat

ion, xenophobia, and related intolerance; it discusses a wide range of concerns and includes far-reaching guidelines and realistic measures. Acknowledging that no country can claim to be free from racism, that racism is a global problem, and that tackling racism is a global concern.

Furthermore, the Human Rights Committee's General Comment (GC) No. 34 on the right to freedom of expression and the Committee on the Elimination of Racial Discrimination's General Recommendation (GR) No. 35, Report to UN General Assembly, UN Strategy and Plan of Action on Hate Speech, Detailed Guidance on Implementation for UN Field Presences, Guidance Note on COVID-19 related Hate Speech has been further clarified the scope and content of these treaties hate speech relating provisions.

2.4. 2 The Regional Human Rights Frameworks

The major regional human rights treaties of the European Convention on Human Rights (ECHR), the American Convention on Human Rights (ACHR), and the African Charter on Human and Peoples' Rights (ACHPR) guarantee the fundamental right to freedom of speech in the same way as other international human rights instruments. And none of these instruments makes freedom of expression absolute; they contain provisions that allow States, in some circumstances, to restrict freedom of expression.

⁶⁰ Ibid, principle 12

⁶¹ Durban Declaration and Programme of Action of the World Conference against Racism, Racial Discrimination, Xenophobia and Related Intolerance (2001)

A. The European Convention on Human Rights (ECHR)

The 1950 European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR) provides the right to freedom of expression including the freedom to hold opinions and to receive and impart information and ideas.⁶² And like other instruments, the right is not absolute; it may subject to limitations, inter alia, protect the reputation or rights of others based.⁶³ The European Court of Human Rights (ECtHR), which has the authority to interpret the text of the Convention, has repeatedly stressed that freedom of speech must be respected because it is one of the fundamental foundations of a democratic society and one of the basic conditions for its development and self-fulfillment.

In its landmark decision in *Handyside v. the United Kingdom*, which is a case involving limitations on the right to freedom of speech in order to protect morals, the Court stated that “freedom of speech applies not only to information or ideas that are positively received or viewed as inoffensive or indifference but also to those that offend, shock or disturb the situation.”⁶⁴

Unlike ICCPR, ECHR under its grounds to limit freedom of expression does not explicitly include a prohibition of forms of hate speech and that is why there is no positive obligation on the States to develop legislation or other tools to prohibit hate speech.

The word hate speech is not recognized in the ECHR, the ECtHR has established a standard test to assess if Article 10 of the ECHR has been violated or not in its case law. Within the margin of appreciation doctrine, which has a major role in the interpretation of the ECHR at the national level, member states are granted a certain amount of freedom as to how they regulate speech. The scope of that discretion, which is subject to the oversight of the ECtHR varies depending on the type of the expression in question. Whereas States have only a limited margin of appreciation

⁶² ECHR Art 10

⁶³ Ibid Article 10(2),

⁶⁴ *Handyside v the United Kingdom*, judgment of 7 December 1976, para. 49, *Benjamin and Others v. Minister of Information and Broadcasting*, 14 February 2(1), Privy Council Appeal No. 2 of 1999, (Judicial Committee of the Privy Council), *Lingens v. Austria*, 8 July 1986, Application No. 9815/82, paras. 39-40 (European Court of Human Rights)

for political speech, they enjoy a broader margin of appreciation for public morality, moral values, and religion.⁶⁵

Despite this margin of appreciation, The ECtHR has developed a standard test to determine whether Article 10, ECHR, has been, especially in the form of a so-called three-part test. In order for the limitation on freedom of expression to be valid, all three sections of the test must be carried out; in the first place, the intervention must be defined by law (i.e., it must be adequately accessible and reasonably foreseeable in its consequences) Second, it must pursue a legitimate aim (i.e., correspond to one of the aims set out in Article 10(2)). Third, it must be necessary for a democratic society (i.e., correspond to a pressing social need) and be proportionate to the legitimate aim(s) pursued.⁶⁶

The outer extremity of that defense is defined by Article 17, ECHR, which is a classical prohibition of the violation of rights clause. It can be considered as a protective device intended to avoid misuse of the ECHR by individuals whose purposes are contrary to the letter and spirit of the Convention. In general, it appears to invoke it in order to ensure that Article 10 does not give protection to racist, xenophobia, or anti-Semitic speech; statements denying, contesting, minimizing, or condoning Holocaust ideas, or neo-Nazi ideas.⁶⁷

However, the European Court of Human Rights has not always applied Article 17 consistently in its case law. The court has sometimes excluded complaints concerning Holocaust-denial speech from the scope of Article 10 ECHR. In *M'Bala M'Bala vs. France*, Dieudonne M'Bala M'Bala, a comedian engaged in political activities, for public insults directed against an individual or a group of persons on account of their ethnicity or of belonging to a particular ethnic class, country, race, or religion, specifically, in this case, persons of Jewish origin or faith. At the end of the December 2008 show in Paris, the applicant welcomed Robert Faurisson, an academic who had obtained a variety of convictions for his revisionist convictions in France.⁶⁸ In

⁶⁵ Tarlach McGonagle, 'The Council of Europe Against Online Hate Speech: Conundrums and Challenges' MCM(2013)005

⁶⁶ *The Sunday Times v United Kingdom*, 26 April 1979, Application No. 6538/74, para. 49 (ECHR)

⁶⁷ *Ibid*

⁶⁸ ECtHR Fact sheet-Hatespeech *M'Bala M'Bala v. France*, 20 October 2015, Application No 25239/13 (EctHR)

accordance with Article 35 (admissibility criteria) of the Convention, the Court declared that the applicant was not entitled to the protection of Article 10 of the freedom of speech according to Article 17 of the prohibition on the violation of rights. The Court, therefore, concluded that the applicant had tried to distract Article 10 from its true purpose by exploiting its right to freedom of expression for reasons which were incompatible with the letter and spirit of the Convention.

While some cases were granted protection of Article 10. Case in point *Perincek v. Switzerland*,⁶⁹ the Armenian genocide, Mr. Perincek, a Turkish politician, publicly expressed his opinion in Switzerland that the mass deportations and massacres suffered by the Armenians in the Ottoman Empire in 1915 and the years that followed did not amount to genocide. In particular, the Swiss courts have held that his motivations seem to be racial and nationalist and that his comments have not contributed to the historical debate. The applicant argued that his criminal prosecution and sentence had infringed his right to freedom of association. The Court held that there had been a violation of Article 10 of the Convention.

The Council of Europe (COE) Committee of Ministers took a broader approach and came in 1997, with the adoption of Recommendations on hate speech⁷⁰ and on the media and the promotion of a culture of tolerance.⁷¹ The term hate speech was defined and taken to cover; “All modes of speech that propagate, incite, facilitate or justify racial hatred, xenophobia, anti-Semitism, or other modes of intolerance-based hatred, including intolerance expressed by violent nationalism and ethnocentrism, prejudice and animosity towards minorities, migrants and immigrants.”⁷²

While this definition is quite wide-ranging, the COE distinguishes between expressions which, although offensive, shocking, and insulting, are fully protected by the right to freedom of expression, and expressions that do not enjoy that protection.

⁶⁹ Ibid *Perincek v Switzerland* 15 October 2015 application no. 27510/08 (EctHR)

⁷⁰ Recommendation No. R (97) 20 of the Committee of Ministers to Member States on “Hate Speech”, (1997) <<https://rm.coe.int/1680505d5b>> Accessed Oct 26 2020

⁷¹ Recommendation No. R (97) 21 of the Committee of Ministers to Member States on the media and the promotion of a culture of tolerance, 1997. <<https://rm.coe.int/168050513b>> Accessed Oct 27, 2020

⁷² Council of Europe, Recommendation No. R 97 (20) *Supra* note 68

Moreover, with regard to the specific online hate speech, (CoE) issued Convention in Cyber Crime⁷³ and the Additional Protocol to the Convention on Cybercrime⁷⁴ which addresses online expression of racism and xenophobia. The Additional Protocol was signed by 40 members of the COE including Austria, France, Germany, the Netherlands, Poland, and Sweden. To date, 29 countries have ratified the instrument⁷⁵ and entered into force in March 2006.⁷⁶

The Protocol imposes a criminalization obligation on the member states Racist and xenophobic slurs online;

1. persons for the reason that they belong to a group differentiated by ethnicity, color, descent or national or ethnic origin, as well as religion if used as a pretext for either of these factors; or

2. a group of persons differentiated by any of these characteristics.⁷⁷

Additionally, the Protocol requires States to enact such legislative and other provisions as may be appropriate to create criminal offenses under their domestic law when they have committed the following intentionally and without rights; to publish, or otherwise make available, racial xenophobia content to the public via a computer system.⁷⁸

With regard to new technologies, the ECtHR, also decided on cases concerning the Internet company's responsibility for user-generated posts on their platforms;

⁷³ Council of Europe, Convention on Cybercrime, 2001< [https:// www. europarl. europa. eu/meetdocs /2014_2019/documents/libe/dv/7_conv_budapest_/7_conv_budapest_en.pdf](https://www.europarl.europa.eu/meetdocs/2014_2019/documents/libe/dv/7_conv_budapest_/7_conv_budapest_en.pdf)> Accessed Oct 27, 2020

⁷⁴ Additional Protocol to the Convention on cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems, 2003<<https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/189>>accessed Oct 27, 2020

⁷⁵ Council of Europe, Chart of signatures and ratifications of Treaty 189 Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems Status as of 01/11/2020<<https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/189/signatures>> accessed on Nov 1, 2020

⁷⁶ Ibid

⁷⁷ Ibid Art 5

⁷⁸ supra note 72

For example, in the case of *Delfi AS v. Estonia*,

The applicant company, which maintains a news portal on a commercial basis, alleged that it had been held responsible by the national courts for the offensive comments made by its readers below one of its online news posts on the ferry company. At the request of the lawyers of the owner of the ferry firm, the applicant company removed the derogatory remarks about six weeks after they had been posted. The Court held that there had been no violation of Article 10 (freedom of speech) of the Convention, holding that the finding of liability of the Estonian courts against the applicant company was a reasonable and proportionate restriction on the freedom of expression of the portal, in particular, because the comments at issue were serious and were made in reaction to an article published by the applicant on its professionally operated news portal on a commercial basis; the steps taken by the applicant to remove the derogatory comments without delay after their release were insufficient; and the fine of EUR 320 had by no way been disproportionate for the applicant, one of the largest internet provider.⁷⁹

B. The African Charter on Human Rights and Peoples' Rights (ACHRP)

The 1981 African Charter on Human and Peoples' Rights (Banjul Charter), guarantees the right to freedom of expression which includes the right to receive information and to express and share opinions.⁸⁰ Unlike other international and regional human rights instruments, such as UDHR, ICCPR, and ECHR the Charter does not mention a specific type of medium in which freedom of expression can be exercised. As Yohannes argues, Art 9(2) of the Charter covers all modern mediums, including the Internet; since the failure to mention any particular type of media could be a positive factor that could be interpreted to include any medium, including the Internet, and also the Charter's flexibility clause, which enables the African Commission to draw inspiration from international law on freedom of speech.⁸¹ In fact, The African Declaration on Speech and Information of 2019 clearly states that the right to freedom of expression must be

⁷⁹ EctHR, Fact sheet New technologies, *Delfi AS v. Estonia* .16 June 2015 App. No. 64569/09

⁸⁰ ACHPR Art 9

⁸¹ Yohannes (n24)

secured both offline and online.⁸² Importantly, like the Rabat Plan of Action, the Declaration stipulates a six-part touchstone that guide states to tackle hate speech.⁸³

The Banjul Charter does not expressly include the prohibition of hate speech, but it generally allows the restriction of rights as long as they are declared "under the law." This Charter has also been criticized for retaining such "claw-back" clauses that leave space for countries to enforce domestic laws and compromise the meaning of the right to freedom of speech.⁸⁴

Notwithstanding the clause, the African Commission on Human and Peoples' Rights (ACHPR) has provided a robust interpretation in various decisions that the term 'under the law' in Article 9(2) applies only to a national law that complies with international standards and does not allow States to override the obligations of the Charter.⁸⁵ The Commission also noted that States could only place the necessary limitations on rights protected by constitutional or international human rights instruments and that no circumstance justified a gross violation of human rights. Moreover, The 2002 Declaration of the Principles of Freedom of expression in Africa declares that the fundamental right to freedom of expression should not be restricted on public order or national security grounds unless there is a real risk of harm to a legitimate interest and there is a close causal link between the risk of harm and the expression.⁸⁶

C. The American Convention on Human Rights (ACHR)

Notwithstanding the strong protection of free speech of this convention, the right to freedom of expression is not an absolute right and can be subject to limitations. Article 13(2), of the

⁸² ACHPR, Declaration of Principles on Freedom of Expression and Access to Information in Africa, (2019) , <file:///C:/Users/yirga/Downloads/Declaration%20of%20Principles%20on%20Freedom%20of%20Expression_ENG_201> accessed Oct 28, 2020

⁸³ Ibid, Principle 23

⁸⁴ Makau Mutua, 'The African Human Rights The African Human Rights System a Critical Evaluation' (2000) <https://digitalcommons.law.buffalo.edu/other_scholarship/16/> accessed Nov 4, 2020

⁸⁵ Constitutional Rights Project (in respect of Lekwot & Others) v Nigeria (2000) Comm no. 148/96 , AHRLR 183

⁸⁶ African Commission on Human and Peoples' Rights (ACoHPR.) Declaration of Principles on Freedom of Expression in Africa, 32nd Session, Banjul, 17 - 23 October 2002.

convention, sets that freedom of expression may be limited to the extent necessary to ensure certain public interests or the rights of other persons.⁸⁷ Unlike other regional instruments, the ACHR is the only document specifically referring to the banning of hate-motivated speech. However, according to the convention, the hate speech to be banned must comply with two conditions, which are: (1) the speech directly incites violence and (2) the speech is likely to give rise to imminent violence.⁸⁸

Compared to other international instruments such as ICCPR and ICERD it is also a very narrow standard; even speech promoting violence and full of racist remarks would be protected in the absence of any evidence that violence or other similar action is likely to occur almost immediately.

As stated by its Special Rapporteur on Freedom of Expression, according to general principles of treaty interpretation, advocacy of hatred that is directed against individuals based on their sexual orientation, gender identity, or bodily diversity, and that constitutes incitement to lawless violence or to any other similar action, falls within the scope of this provision and is, therefore, a violation of Article 13 of the Convention.⁸⁹

Further Pursuant to Article 13(5) of the Convention States are required to adopt legislation to punish advocacy of hatred that constitutes “incitement to lawless violence or any other similar action.”⁹⁰

2.5. Freedom of Expression, the Right to Equality Versus Hate Speech

Freedom of expression is a fundamental human right guaranteed by major international and regional human rights treaties and is regarded as a dual right, first, the right to communicate, disseminate or express views and ideas of any kind (political, creative, commercial expression)

⁸⁷ AmCHR Art 13 (2)

⁸⁸ Ibid art 13

⁸⁹ IACHR. Annual Report 2009. Annual Report of the Office of the Special Rapporteur for Freedom of Expression. Chapter III (Inter-American Legal Framework of the Right to Freedom of Expression). OEA/Ser.L/V/II. Doc. 51. December 30, 2009, paras. 58-59. IACHR, Annual Report 2014: Chapter V: Follow-up to the 2012 Jamaica Report, para. 238.

⁹⁰ Ibid, para. 58

and, second, the right to seek and receive information in any form (orally, in writing, in the form of art, or through any other media, including new technologies such as the internet)⁹¹

The scope of the right to freedom of expression is broad. It includes the expression of opinions and ideas that others may find deeply offensive, and this may encompass scornful expression.⁹²The right to freedom of expression is not an absolute right, and states may, under certain exceptional circumstances, restrict it.

However, all the restrictions must meet the so-called three-part test. This means that restrictions must be; (1) provided by law; (2) pursue a legitimate aim as exhaustively listed under Article 19 (3) International Covenant on Civil and Political Rights (ICCPR), and (3) be necessary and proportionate to the aim pursued.

On the other hand, international human rights law ensures equality and non-discrimination for all people, and States are bound to guarantee equality in the realization of human rights and equal protection under the law⁹³; that is the domestic legislation required to ensure that all people are equal in the eyes of the law and free from discrimination on grounds of race, gender, ethnicity, religion or belief, disability, age, sex, national or social orientation, language, political or other grounds.

It is the protection of the dignity of all individuals without discrimination that motivates the majority of hate speech responses, including limitations on the right to freedom of expression.⁹⁴ In another word, the need to regulate hate speech is aimed at drawing a line between the right to equality and the protection of the fundamental right to freedom of expression.

Generally speaking, hate speech means a speech that could vilify a person or a group of people, because of who they are. This means that it's not hate speech to say, "I hate you," because there

⁹¹ Elena Mihajlova, Jasna Bacovska, Tome Shekerdjiev 'Freedom of expression and hate speech 'Skopje : OBSE, (2013).42

⁹² General Comment No. 34,(n21)para 11

⁹³ UDHR Art 1 and ICCPR Articles 2(1) and Article 26

⁹⁴ CPPCG Art. 3 (c) , RSICR),Art 6 and 25 (3) (e) and ICCPR Art 20(2)

is no connection to a group;⁹⁵ and to be considered as hate speech must be addressed to specific categories of groups, such as people of the same race, gender, ethnicity, religion or belief, disability, age, sex, national or social orientation, language, political or other grounds. It is also important to keep in mind that not all hate speech may legitimately be subject to restriction.

2.6. A Proposed Typology of Hate Speech

The word hate speech covers a very broad spectrum of both unlawful and lawful types of hate speech. In 2012, according to the level of severity, the annual report of the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression classified hate speech in three categories.⁹⁶

A. Hate speech that Must be Prohibited (Top Level of Hate Speech); Under international law, the severest form of hate speech is prohibited. International criminal law and Article 20(2) of the ICCPR requires States to prohibit certain severe forms of hate including direct and public incitement to commit genocide as defined in international criminal law; and any promotion of national, ethnic or religious hatred as a means of incitement to discrimination, hostility or violence as defined in Article 20(2) of the ICCPR.⁹⁷ Besides, ICERD obliges States to criminalize the propagation of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as any act of violence or incitement against any race or group of persons of any other color or ethnic origin.⁹⁸

The severity of the expression to be criminalized as incitement to discrimination, hatred, or violence depends on the fulfillment of the RAP six-part test. The Rabat threshold test provides the basis for a consistent understanding of the nature and severity of hate speech. And basically, incitement requires a triangular relationship between the three main actors (1) the "hate speaker" who promotes discriminatory hatred of the public audience; (2) the public audience who may

⁹⁵ Suzan Bensch, (n25)

⁹⁶ UN General Assembly , Promotion and protection of the right to freedom of opinion and expression (A/67/357 2012) para 2

⁹⁷ Supra not 92

⁹⁸ ICERD Art 4

participate in acts of discrimination, aggression, or violence; and (3) the target group towards which such acts may be committed.⁹⁹

B. Hate Speech That May be Prohibited (Intermediate level of Hate Speech): Certain types of hate speech may be prohibited under international law, even though they do not meet the Rabat Plan of Action threshold of incitement. States may prohibit other forms of hate speech provided they comply with the requirements of Article 19(3) of the ICCPR; that is, limitations must: (a) be provided by law; (b) seek a legitimate aim, such as respect for the rights of others, including the right to equality and non-discrimination, or the preservation of public order; and (c) be necessary and proportionate in a democratic society. Therefore, as long as the standards of the three-part test are respected, limitations on freedom of expression can be imposed in order to protect individuals from hate speech based on their protected characteristics to secure their right to equality and non-discrimination. For example, limits on freedom of speech can be placed to protect persons from threats of violence or harassment based on their identity (e.g. as women, youth, or migrants) if the three-part test is met.¹⁰⁰

C. Lawful Hate Speech (Bottom level of Hate Speech): Under international law, expressions that may be inflammatory or offensive, but which do not reach any of the criteria at which limitations on speech are justified, may not be subject to legal sanctions. Legal limitations should not be put on the dissemination of expressions that are insulting, shocking or troubling, condoning or denying historical events, including crimes of genocide or crimes against humanity and blasphemous speech, including threats to religious sentiments, lack of reverence for religion or any other belief system, and defamation of religion. Unless, under Article 20(2) ICCPR, certain modes of speech also constitute an incitement to hostility, discrimination, or violence.¹⁰¹ However, this does not exclude States from taking legal and policy action to address the underlying biases that this category of hate speech is symptomatic of, or from maximizing

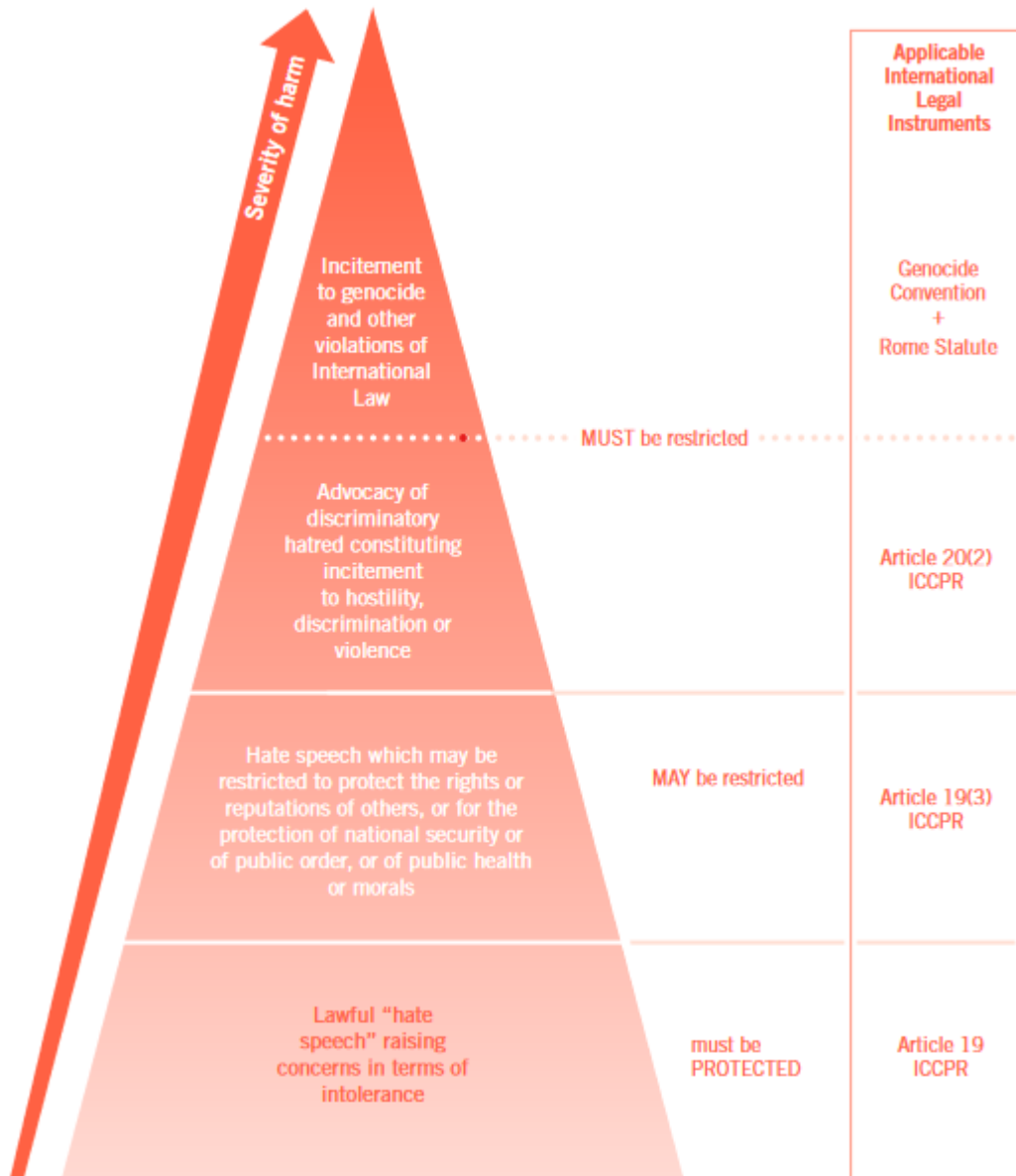
⁹⁹ Article19, “Hate Speech” Explained: A Toolkit, 2012

¹⁰⁰ United Nations Strategy and Plan of Action on Hate Speech Detailed Guidance on Implementation for United Nations Field Presences, 2020

¹⁰¹ Ibid

opportunities for all citizens, including elected officials and organizations, to participate in counter-speech.¹⁰²

Figure1. The ‘Hate Speech Pyramid’



SOURCE: Based on ARTICLE 19, “Hate Speech” Explained: A Toolkit, p.19

¹⁰² Article 19, Supra note 97

2.7. Hate speech vs Hate crime

Mostly, the words "Hate speech" and "hate crimes" are used interchangeably to refer to hatred base circumstances. But even though both words are symptomatic of prejudice and intolerance, they have divergent concerns and elements. The use of the word 'hate' may deceive people into assuming that every manifestation of 'hate,' including 'hate speech,' is a criminal offense; but hate speech would not always be a criminal offense without the explicitly forbidden content.¹⁰³

However, what is to be emphasized is that hate speech can constitute a criminal offense only if speech is unprotected or punishable, but, in general, speech irrespective of its offensiveness and can also be categorized as hate if it is protected by freedom of expression and is not a criminal offense.

On the other hand, hate crime refers to criminal activities in which the perpetrator is motivated by biases toward members of a group based on the racial, ethnic, religious, etc. identity of the victim.¹⁰⁴ To characterize a criminal offense as a hate crime, two combined elements must be fulfilled, i.e. the criminal base offense and the bias motive for committing the offense. It is the second factor that is a bias motive for committing an offense that makes hate crimes distinct from other common crimes¹⁰⁵ and hate crimes are not a single offense; they may be acts of violence, assaults, harm to property, assault, murder, or some other criminal offense.¹⁰⁶ Having a clear understanding between the terms of hate speech and hate crime is also essential for the protection of the right to a fair trial; the offender would be more likely to be held liable to the degree of the act that has been committed.

2.8 Hate speech on Social Media

Social media have become a more integral part of people's daily lives and have made it easier for many more people to enjoy their freedom of speech and to take part in public discourse. On the other hand, the atmosphere of social media discourse is blamed for polarizing, creating division,

¹⁰³ Ibid

¹⁰⁴ Elena Mihajlova ,supra note 89

¹⁰⁵ Ibid

¹⁰⁶HateCrime Laws: Practical Guide(2009) OSCE/ODIHR< <https://www.osce.org/files/f/documents/3/e/36426.pdf>>

Accessed Sep 6,2020

and promoting hatred. This has become a global concern and many experts have been forward their concerns with regard to these social media abuses.

For example, the UN Human Rights Council Special Rapporteur on Minority Issues has noted that social media platforms have also been used to disseminate hate speech messages towards groups that have fueled inter-communal tensions and, in some cases, led to violent clashes between communities¹⁰⁷

Beyond the international and regional legislative framework introduced to curb online hate speech, many social media intermediaries, such as Facebook, Youtube, and Twitter, have adopted their own laws on the regulation of hate speech in the content shared by users, but what constitutes hate speech varies from one intermediary to another.

For example, Facebook defines hate speech as a 'threatening language' that poses serious threats to public and personal safety¹⁰⁸ and YouTube defines hate speech as a material that encourages violence or hostility towards individuals or groups on the basis of characteristics such as race or ethnic origin, religion, disability, gender, age, veteran status or sexual orientation.¹⁰⁹ On the other hand, Twitter even it does not mention explicitly a prohibition of hate speech but it has a descriptive list which prohibits the promotion of hate content, sensitive topics, and violence globally.¹¹⁰

Despite the legislative mechanisms and intermediary laws that are implemented to curtail online hate speech control, online hate speech remains daunting.

¹⁰⁷ UN General Assembly, Report of the Special Rapporteur of Minority Issues (A/HRC/28/64 2015) para 76

¹⁰⁸ Facebook community Guidelines, 'What does Facebook consider to be hate speech?' <<https://www.facebook.com/help/135402139904490>> accessed 12 November 2020

¹⁰⁹ Youtube Community Guidelines, Hate speech <<https://support.google.com/youtube/answer/2801939?hl=en>> accessed 13 Sep 2020

¹¹⁰ Twitter, 'Hate content, sensitive topics and violence' <<https://support.twitter.com/articles/20170425>> accessed 13 Sep, 2020

CHAPTER THREE

REGULATING HATE SPEECH UNDER NATIONAL LAWS

3.1. Introduction

This chapter consists an overview of the national relevant frameworks and experiences of other jurisdictions relating to social media hate speech regulation.

3.2 The National Legal Frameworks

3.2.1. Legal protection of the Right to Freedom of Expression on Social Media

The Constitution of 1931, which was the first written constitution of Ethiopia, does not have a significant impact on human rights narratives in general, but the revised Constitution of 1955 recognized a number of human rights and the freedom of expression was entitled to protection for the first time under this revised constitution.¹¹¹ Subsequently, freedom of expression was also incorporated into the Constitution of 1987.¹¹²

Further, the 1995 FDRE Constitution guarantees protection of the right to freedom of expression in Article 29 in the following terms.

"Without interference, everyone has the right to freedom of expression. This right shall include freedom to seek, receive and share information and ideas of all kinds, irrespective of frontiers, whether orally, in writing or in print, in the form of art, or through any media of its choice." ¹¹³

The Constitution recognizes freedom of expression in the same way as the ICCPR, which includes freedom to seek, receive, and impart information through any media of its choice. With this articulation, one may ask does freedom of expression on social media is protected under the constitution?

¹¹¹ The revised Constitution of the Empire of Ethiopia, 1955,Article 41.

¹¹² The Constitution Of The People's Democratic Republic of Ethiopia, 1987,Article 47(1)

¹¹³ The Constitution of the Federal Democratic Republic of Ethiopia, 1995, Article 29

Although the Constitution does not provide a direct reference to freedom of speech through the internet and social media, the term 'through any media of his choice' can be interpreted as including all mediums, including the Internet and social media. Besides, the Constitution also includes a laudable interpretation provision regarding the chapter on human rights and fundamental freedoms. Article 13(2) demands that the third chapter of the Constitution be read in a manner consistent with the principles of the (UDHR), the International Covenants on Human Rights, and the international instruments adopted by Ethiopia.¹¹⁴ The reading of this Article should therefore be compatible with the specific provisions of international instruments such as the ICCPR and the ACHPR. And freedom of expression on the internet is protected in both ICCPR and ACHPR.¹¹⁵ Thus the constitution guarantees freedom of expression through all mediums including social media.

On the other hand, the constitution allows the limitation of the right to freedom of expression in the following terms.

These rights can be restricted by law that is governed by the principle that freedom of speech and information can not be limited by the content or effect of the point of view expressed. Legal limits may be laid down in order to preserve the well-being of Youth and the honor and reputation of individuals. Any propaganda for war, as well as a public expression of opinion intended to harm human dignity, shall be prohibited by law.¹¹⁶

3.2.2 The Legal Protection on the Right to Equality

The FDRE constitution guarantees protection of the right to equality under Article 25. It states that all persons are equal before the law and are entitled to equal protection of the law without discrimination: and the law shall guarantee equality and strong protection without discrimination on grounds of ethnicity, nationality, nationality or other social origins, color, sex, language,

¹¹⁴ The 1995 constitution, Supura note 109, Article 13

¹¹⁵ General comment 34 and Declaration of Principles on Freedom of Expression and Access to Information in Africa

¹¹⁶ FDRE constitution supra note 111, Art 29(6)

religion, political or other opinions, property, birth or other status.¹¹⁷ In contrast to the protection of freedom of speech, the right to equality is not subject to any restriction and is non-derogable even if during a state of emergency.¹¹⁸

In the same vein, the Ethiopian Criminal Code forbids discrimination on the basis of individual social circumstances, race, nation, nationality, social origin, color, sex, language, religion, political or other opinions, property, birth, or other status.¹¹⁹

The hate speech regulation can be seen as a context in which two of the constitutional principles listed above, namely freedom of speech and equality, are in dispute.¹²⁰ In other words, the regulation on hate speech is a tool for striking a balance between the two basic rights of freedom of speech and the right to equality. As long as Ethiopia is a State party to ICCPR, ICERD, and ACHPR, it is legally entitled to limit freedom of speech, including online expressions. Besides, Article 29 of the Constitution allows limitation on the right to freedom of speech.

3.2.3. Ethiopia's Legal Framework for Tackling Social Media Speech

Prior to the recent Hate Speech and Disinformation Prevention and Suppression Declaration proclamation, hate speech law in Ethiopia was founded under a scattered piece of legislation.

For example, Article 486, which is also repealed by the current Hate Speech law, and Article 816, of the Criminal Code, are among the clauses that are relevant to the law of hate speech regulation. Article 486(b) forbids expressions that cause hatred or incite acts of violence or political, ethnic, or religious disturbances and Art 816 also criminalizes blasphemous utterances or attitudes that are grossly insulting to the religious sentiments or convictions of others, or the divine being or to the religious symbols, rituals, or religious characters of others.

Prior to the adoption of the latest hate speech, there are already various cases related to hate speech;

¹¹⁷ Ibid Art 25

¹¹⁸ Ibid Art 93 (4) c

¹¹⁹ The Criminal Code of the Federal Democratic Republic of Ethiopia 2004, Art 4

¹²⁰ Yared Legesse Mengistu, ' (n22) 352

For example.

- **Public Prosecutor v. Hailu Shawel**,¹²¹ the Public Prosecutor, charged the founders and representatives of the Coalition for Unity and Democracy (CUD), with a statement made by prominent opposition figure Ato Bedru Adem, according to a video recording of his speech in Assela, a small town in the Oromia region. Ato Bedru said, "The power of the Federal Government is completely in the hands of the Tigrayans and the EPRDF, so they should then be forced back into their land by the united power of the people." The prosecutor also said as a result of the speech that two houses of individuals who were Tigrayans had been burning and another Tigrayan were attacked. The Federal High Court ruled that it was simply a "hate speech" and that it was not a genocide
- **Prosecutor v Elias Gebru Godana**¹²² ; Elias was charged with a violation of Article 257(e) of the revised Penal Code – an incitement of hatred relating to an article published in Enqu Magazine in March 2012 about a controversial monument to ethnic Oromos massacred in the 19th century, which claims "the opposition and challenge caused to Emperor Menelik's to strengthen the country that was weakened was unexpected and a betrayal and the actions that Emperor Menelik accused of committing, even if true, what choice did he have? What is the injustice of the act? "

The prosecutor argued that the article contained statements targeting a particular group and, as a result of the statements, student demonstrators smashed windows and other property worth 39, 408 Birr.

Moreover, though it does not specifically include the ban of hate speech, Ethiopia has a law that tackles illegal content on the Internet, the so-called 'Computer Crime Proclamation 958/2016',¹²³ which criminalizes various contents based on cyber crimes such as obscene or indecent crimes against minors, cybercrimes against liberty and the reputation of individuals, crimes against public protection and crimes against minors. The Telecom Fraud Offense

¹²¹ Mesenbet A Tadege, 'Making Space for Non-Liberal Constitutionalism in Free Speech: Lessons from a Comparative Study of the State of Free Speech in Ethiopia and Thailand' (2018) 30 J Ethiopian L 27

¹²² Ibid

¹²³ A Proclamation to Provide for the Computer Crime 958/2016 Federal Negarit Gazette 22th Year, Addis Ababa 2016

Proclamation is also relevant legislation aimed at regulating some unlawful content related to the use of telecommunications services.¹²⁴

Since these legislative requirements lack consistency and comprehensiveness which could not adequately address hate speech, the government enacted special legislation called the 'Hate Speech and Disinformation Prevention and Suppression Proclamation proclamation' which came into force on 23 March 2020.

3.3 Lessons from other Jurisdictions Relating to Social Media Hate Speech Regulation

3.3.1 Germany:

Enforcement on Social Networks '(NetzDG) 2017

Germany's hate speech legislation stems from both the old laws defending "honor" and "social status" and comprehensive legislation passed after World War II for the Protection of minorities.¹²⁵ The existing German social media hate speech legislation called 'Enforcement on Social Networks '(NetzDG) came into effect in the country in Oct 2017.¹²⁶ NetzDG is written to regulate the dissemination of hate speech and fake news online.¹²⁷ In determining the illegality of the substance, the legislation refers to the numerous criminal offenses provided in the Criminal Code, such as incitement to hatred, propagation of depictions of violence, defamation of religions, religious and ideological alliances, insult, deliberate defamation, and public incitement to crime¹²⁸ and it is only applicable to social media networks that have two or more million registered users in the country¹²⁹ such as Facebook, Youtube, and Twitter.¹³⁰

¹²⁴ Telecom Fraud offence Proclamation No .761/2012 Art (6)

¹²⁵ Rebecca Zipursky ' Nuts About NETZ: The Network Enforcement Act and Freedom of Expression' (2019) Fordham International Law Journal Volume 42, Issue 4 Article 7, 1328

¹²⁶ Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act), 2017

¹²⁷ Ibid

¹²⁸ Ibid Sec. 1(3)

¹²⁹ Ibid Sec 1(2)

¹³⁰ Alkiviadou, Natalie ' Hate Speech on Social Media Networks: Towards a Regulatory Framework?' (2019) Information & Communications Technology Law, 28:1, 19-35

The Act defines social media networks as Telemedia service providers which for profit-making purposes, operate internet platforms designed to allow users to share any content with other users or to making such content accessible to the public.¹³¹ However, platforms providing journalistic or editorial content and platforms designed to allow individual communication or distribution of particular content are specifically excluded.¹³²

NetzDG requires the provider of Social Networks to have an efficient and clear procedure for handling complaints about illegal content intended to provide users with a clearly identifiable, directly accessible, and permanently available procedure for filing complaints about illegal content.¹³³ After taking immediate notice of the complaint, they are required to ensure whether the contents of the complaint are illegal and subject to removal or whether access to the contents must be blocked. The Act obliges these social media networks to delete clear illegal content within 24 hours after receiving a user complaint.¹³⁴ In such exceptional cases, the removal of illegal material can take up to seven days or more.¹³⁵ The decisions taken with respect to the complaint and the reasons for approving or denying it must be reported to the complainant and the user affected without unnecessary delay.¹³⁶

social media network falling within the scope of the Act is required to appoint a domestic agent to facilitate the process and to reply to the competent authorities.¹³⁷ Furthermore, social media networks that receive more than 100 complaints about illegal content in a calendar year are forced to disclose biannual reports in German on the handling of these complaints via the Federal Gazette and on the homepage of the social media network one month after the end of each half-

¹³¹ supra note 124, Sec 1(1)

¹³² Ibid

¹³³ Ibid Sec 3 (1)

¹³⁴ Ibid Sec. 3(2)

¹³⁵ If the illegality of the material is not clear, the social network has seven days to investigate and remove the content. The seven-day duration could be extended if the social network employs an outside agency to carry out the verification process. See Sec 3(3) of the Act

¹³⁶ Supra Note 121, Sec 3(2)5

¹³⁷ Ibid Sec 5(2)

year cycle.¹³⁸ Social media failure to comply with the above-mentioned obligations imposes a regulatory fine plating of between 5 hundred thousand and 5 million Euros.¹³⁹

3.3.2 The United State of America

The Communication Decency Act (CDA) 1996

The United States is unique in its conviction that the damage caused by the prohibition of hate speech is more severe and more harmful than the harm done by hate speech.¹⁴⁰ It grants constitutional protection against hate speech, except when it includes to give rise to imminent violence¹⁴¹.

For example, in the landmark case, *Brandenburg v. Ohio*, the Supreme Court reversed the conviction of the leader of the Ku Klux Klan, who had been prosecuted under the Ohio Criminal syndicalism Law for promoting violent political and industrial change when the local news outlet made statements during the Klan meeting. The Supreme Court declared that the state cannot restrict activism unless it is directed to inciting or bringing about imminent lawless action and is likely to bring about such action.¹⁴²

The Communications Decency Act (CDA) that passed in 1996 is the applicable legal standard for Internet computer service (social media) in the United States. The Act defines Internet computer service as any information service, system, or software access provider that provides or allows several users to access the computer server; Specifically, service or system offering access to the Internet and those systems run or programs provided by libraries or educational Institutions¹⁴³ which social media normally fall under this definition.

¹³⁸ Ibid Sec 2(1)

¹³⁹ Ibid, Sec 4(2)

¹⁴⁰ Raohel Cohen-Almagor 'Hate And Racist Speech In The United States a Critique(2015) *Philosophy and Public Issues (New Series)*, Vol. 5, No. 3 77

¹⁴¹ IACHR art 13

¹⁴² *Brandenberg v. Ohio*, 395 U.S. 444, 447 (1969)

¹⁴³ 47 U.S. Code § 230 - Protection for Private Blocking and Screening of Offensive Material.” Legal Information Institute.

Under Section 230 of the DCA, it is clearly stated that

" No provider or user of an interactive computer service shall be considered as the publisher or speaker of any information provided by another information content provider."¹⁴⁴

And this articulation is being interpreted that Internet service providers (ISP) shall not be legally responsible for the words of third parties using their services.¹⁴⁵ In other words, this article provides a haven for social media outlets such as Facebook, Youtube, not to be held accountable for illegal content uploaded by third parties to their websites. Besides the Act provides more protection to ISPs, removing civil liability for acts that ISPs may take to block access to content that they consider objectionable.¹⁴⁶ But the CDA does not exempt IPS from intellectual property lawsuits or federal criminal laws, state laws, and communications privacy law.¹⁴⁷

Yet the global nature of social media keeps failing to completely protect ISPS from responsibility, as it cannot avoid foreign regulations. For example, in the case of *Yahoo! v. La Ligue Contre Le Racisme Et L'Antisemitisme*¹⁴⁸, in 2006, Yahoo! found itself subject to French laws banning Holocaust denial when it was forced to delete the Nazi memorabilia auction site from every URL accessible in France.

Even though Yahoo! has immunity in America, the Ninth Circuit held that Yahoo! had an obligation to take "all reasonable (or available) steps" to block content in France, but that the ban was not enforceable in the United States.

¹⁴⁴ Ibid, 47 U.S. § 230 (c)1

¹⁴⁵ *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997)

¹⁴⁶ *Supra* note 138, 47 US § 230(c) 2

¹⁴⁷ Ibid, 47 U.S. § 230 (e)

¹⁴⁸ *Yahoo! Inc. v. La Ligue Contre Le Racisme Et L'Antisemitisme*, 433 F.3d 1199, 1202 (9th Cir. 2006)

CHAPTER FOUR

Assessing the Adequacy of the HsDPSP to Regulate Hate Speech on Social Media in Ethiopia

4.1 Introduction

This chapter seeks to answer the central research question, which is how adequate is Ethiopia's prevailing legal framework to regulate social media hate speech? It does so by assessing and examining the new legislation on hate speech with respect to regulating hate speech on social media in Ethiopia. The evaluation takes in the context of essential parameters, including compliance of the document with the international principles of freedom of expression, compatibility of its scope of protection with the international principle of non-discrimination, legal enforcement, and feasibility of the Social Media Service Providers (SMSPs) duty, and procedural protections.

4.2 The Duty of the Ethiopian government to tackle hate speech on social media

Freedom of expression is one of the fundamental rights guaranteed by the Constitution of Ethiopia. However, like many other international treaties, the protection of the right is not absolute, and the Constitution also provides for exceptions that the right may be restricted. These include the protection of the well-being of youth, the honor and integrity of individuals, the prohibition of war propaganda, and public expression of views aimed at destroying human dignity.¹⁴⁹ Furthermore, Ethiopia has ratified international and regional treaties, such as the ICCPR in 1993¹⁵⁰ and the ACHPR in 1998¹⁵¹, which guarantee the right of freedom of expression. Thus, as a state party to these treaties to fulfill the obligations under international

¹⁴⁹ FDRE Constitution Art 29(6)

¹⁵⁰ UN Treaty Collection, ICCPR, New York, 16 December 1966 https://treaties.un.org/pages/ViewDetails.aspx?sr c= TREATY&mtdsg_no=IV-4&chapter =4&clang= en accessed on 22 Nov, 2020

¹⁵¹ List Of Countries Which Have Signed, Ratified/Accessed to The African Charter on Human and People's Rights, https://au.int/sites/default/files/treaties/36390-sl-african_charter_on_human_and_peoples_rights_2.pdf accessed on 22Nov,2020

human rights law. This means the Ethiopian government has a duty to respect, protect, and fulfill the right to freedom of expression.

At the same time, the above international and regional treaties the ICCPR and ACHPR make it exceptionally possible to restrict freedom of speech. However, that does not mean that a limitation can be imposed on the right with no restriction. Accordingly, a State may restrict the right if only the restriction passes the three-part test laid down in Article 19(3) of the ICCPR. These are limitations placed on the freedom (1) should be provided by law (2) in pursuit of a legitimate aim and (3) necessary in a democratic society.¹⁵²

Many states have introduced various legislation that restricts freedom of expression, including legislation on the prohibition of hate speech. There are, of course, international treaties that mandate a state to ban forms of hate speech. For example, the CPPCG and the ICCPR forbid direct and public incitement to genocide and any advocacy of discriminatory hatred that constitutes incitement to discrimination, hostility, or violence. Thus, including the State Party's prohibitions under these Conventions, any prohibition on 'hate speech' must fulfil every component of this three-part examination.¹⁵³

The Ethiopian Government as a state party to ICCPR is obliged to prohibit by law any advocacy of national, racial, and religious hatred that constitutes incitement to discrimination, hostility, and violence. And as hate speech is a pressing issue in Ethiopia that demands appropriate regulation, it is appealing that the government of Ethiopia has recently enacted a special law that can counter hate speech in the country both online and offline. Because, beyond its part in resolving the alarming problem in the country, it shows that the government of Ethiopia is committed to fulfilling the international obligation emanating from the ratification of the treaties.

Therefore, while the Ethiopian government has a duty emanating from the adoption of international instruments to protect freedom of expression, these instruments have also made it possible to limit such freedom exceptionally. And the following section has analyzed the

¹⁵² General Comment 34 ,(n21) para 22

¹⁵³ Supra note 97

prevailing hate speech legislation, which limits freedom of expression in light of the Ethiopian government international obligations.

4.3 The Hate Speech and Disinformation Prevention and Suppression Proclamation (HsDPSP)

Ethiopia's legal framework for regulating hate speech was adopted in March 2020. It is clearly stated in its preamble that it aims to prevent the spread of hate speech and fake news in Ethiopia, both online and offline.¹⁵⁴ It is intended to minimize hate speech mainly by criminalizing individuals who disseminate hate speech by "broadcasting, printing, or social media using text, pictures, audio, or video." It also puts duties on SMSP to takedown the related contents that including various forms of hate speech.

In the case of hate speech, the law explicitly prohibits speech that intentionally promotes hatred, discrimination, or attacks against a person or a discernible identity group on the basis of ethnicity, religion, race, sex, gender, or disability through broadcasting, printing, or social media¹⁵⁵ The law imposes fines up to 100,000 Ethiopia Birr and prison sentences on up to 3 years for any person who disseminates hate speech.¹⁵⁶ The law imposes fines of up to 100,000 Ethiopia Birr and prison sentences of up to 3 years for any person who disseminates hate speech. Besides, it obliges social media networks to remove illegal content within 24 hours. The proclamation also imposed obligations on the Human Rights Commission to carry out public awareness programs to curb hate speech.¹⁵⁷

¹⁵⁴ Hate Speech and Disinformation Prevention and Suppression Proclamation 1185/2020, Federal Negarit Gazette 26th Year, Addis Ababa, 23 March 2020

¹⁵⁵ Ibid Art 2(2) and 4

¹⁵⁶ Ibid Art 7

¹⁵⁷ Ibid Art 8(6)

4.3.1 Compatibility of the Proclamation with International human rights law Standards on Freedom of Expression

Under international human rights law, a measure limiting hate speech demand to reconcile two sets of values: demands for free discussion and individual liberty and progress, with an equally persuasive duty, to avoid attacks on vulnerable communities and to guarantee the equitable and non-discriminatory participation of all people in public life.¹⁵⁸ And since freedom of expression is a fundamental right of a democratic society, the imitations on it including the prohibition of hate speech, should be properly examined so as not to override the enjoyment of the right.

As such, specific guidelines are provided for States to restrict freedom of expression in their domestic laws in a manner compatible with international human rights law. The Human Rights Committee found that if a state adopts a law that limits freedom of speech and is to be defined as a "law," it must be drafted with adequate clarity to allow a person to control his or her actions accordingly.¹⁵⁹ Moreover, although the ACHPR does implicitly list the grounds for limitation, the 2019 Declaration of Principles on Freedom of Expression and Access to Information in Africa explicitly stipulates States can only restrict the exercise of freedom of speech and access to information in the following ways. First, the restriction should be prescribed by law meaning that every legislation restricting the right to freedom of expression and access to information: should be clear, precise, accessible, and foreseeable.¹⁶⁰ Second, the measure needs to uphold a legitimate aim and third, the measure must be necessary and proportionate means of achieving the stated objective in a democratic society.

Thus, Ethiopia, as a party to the ACHPR and the ICCPR, while passing laws that restrict freedom of expression, such as the Hate Speech Proclamation, has a duty to draft laws clearly and narrowly that enables a person to control his or her acts in line with the law without requiring additional discretion on the part of government actors.

¹⁵⁸ UNGA, Promotion and protection of the right to freedom of opinion and expression, Note by the Secretary-General <https://www.ohchr.org/Documents/Issues/Opinion/A_74_486.pdf> accessed on 2 Dec 2020

¹⁵⁹ General comment No. 34(n21) Article 19: Freedoms of opinion and expression, 2011 para 25

¹⁶⁰ Declaration of Principles on Freedom of Expression and Access to Information in Africa, supra note 80, principle 9

In Article 2(2) of the Proclamation, hate speech is defined as "a speech that intentionally promotes hatred, discrimination, or attack against a person or a discernible group of identity based on ethnicity, religion, race, gender, or disability." This definition is vague in its formulations to determine the scope of hate speech. Despite the Proclamation attempts to give a clear definition of the keywords such as "discrimination and 'attack' but it, however, fails to define just what the term "hatred" means.¹⁶¹

The text leaves room for the discretion of law enforcement authorities, such as prosecutors and courts of what is to be viewed as 'hatred', which could lead to violations of the right to freedom of expression. In other words, it opens the door to arbitrary application of the law and may be misinterpreted to remove legal content. Besides that, the formulation of the legislation has not been adequately precise for people to regulate their conduct i.e it is difficult for a person who has written or who shares it to identify whether the content is perceived to be as hate speech and may be liable. For instance, perception and interpretation may vary from one to the other, what is perceived to be hatred for some group of society may not necessarily be hatred to others. So this ambiguous language will make it impossible for people to control their affairs, and they may choose to remain silent in fear of criminalization. Such vague terminology does not conform with the requirements of freedom of expression and information set out in Article 19(3) of the ICCPR and Article 9 of the ACHPR.

The RPA, which is soft law, recommends that as states enact domestic laws on the regulation of incitement to hatred, they should consider providing robust definitions of keywords such as hatred, discrimination, violence, hostility, and in doing so, legislation should have been drawn from the standards and definitions set out in the Camden Principles.¹⁶² The terms 'hatred' and 'hostility' are defined in the Camden Principle as "a strong and irrational feelings of opprobrium, enmity, and hatred towards the target group."¹⁶³

¹⁶¹ Addis Zeyeye media conversation with Yohannes Eneyew Ayalew PhD Candidate, Monash Law at in this link <<https://www.youtube.com/watch?v=1SqceP7RS8Y&t=420s>>

¹⁶² Rabat Plan of Action (n14). para 21

¹⁶³ The Camden Principle (n57)

RPA and the Camden Principle are both non-binding instruments that can be used internationally as guidelines by states to effectively regulate incitement to hatred in accordance with international norms of freedom of speech. Therefore the Government of Ethiopia should have been learnt from the Camden Principles and define the term hatred which provides the capability to have clear and narrowly formulated legislation.

Second, the legislation fails to provide a clear guideline for law enforcement authorities on how to determine the severity of hate speech i.e it does not set down requirements that should be followed to decide whether or not statements are deemed to be criminal offenses. The RPA and the 2019 Declaration of Principles on Freedom of Expression and Access to Information in Africa¹⁶⁴ contain six parts threshold to be taken in order to decide for an expression to be considered as hate speech. These are

(1) The context of the expression; Context analysis should bring the speech act within the social and political context existing at the moment the speech was made and disseminated since it would have a clear effect on both intent and/or causation.¹⁶⁵ In Ethiopia, the issues of ethnicity and ethnic slurs must be properly considered to assess whether the expression is hateful or not.

For example; statements on social media that could have intensified ethnic tensions across the country

(2) Status of the speaker; the position or status of the speaker in the community in particular the position of the person or organization in the context of the audience to which the speech is addressed, should be considered.¹⁶⁶ i.e the speaker power or influence in society should be taken into account

For Example, a speech of national leaders, politicians, public officials, religious or faith leaders, or social media influencers has a great influence on one society

¹⁶⁴ Declaration of Principles on Freedom of Expression and Access to Information in Africa, supra note 80, principle 23

¹⁶⁵ Rabat Plan of Action (n14) para 29(a)

¹⁶⁶ Ibid para 29(b)

(3) intent; Negligence and recklessness are not enough for an act to be a crime under Article 20 of the ICCPR, since this article provides for "advocacy" and "incitement" rather than mere dissemination or circulation of content. In this respect, it involves the activation of a triangular relationship between the object and the subject of the speech act as well as between the audience.¹⁶⁷

(4) content and form; Content analysis may include the degree to which the speech was offensive and straightforward, as well as the form, style, nature of the arguments used in the speech or the balance between the arguments used;

(5) extent and magnitude; Extent includes elements such as the reach of the speech act, its public existence, its duration and the size of its audience, whether the audience had the means to act on the incitement i.e whether the speech (or work) was distributed in a restricted atmosphere or was freely open to the general public;

(6) the likelihood, including imminence; Some degree of risk of harming must be identified. This means that the courts would have to decide that there was a reasonable probability that the speech would result in inciting a direct action against the targeted group recognizing that such a cause should be quite direct.

If the HsDPSP was guided by the RPA and the 2019 Declaration of Principles on Freedom of Expression and Access to Information in Africa and incorporated the above-mentioned valuable six-part test, the law enforcement authority such as courts would easily determine speeches that are considered as hate speech and also ensure the consistent enforcement of the law. However, the cumulative of both the ambiguous definition of hate speech and the lack of explicit guidelines on how to assess a speech that is deemed to be hate speech by law enforcement authority may intensify the possibility of arbitrary implementation of the law which has a chilling effect on freedom of expression.

Concerns are also being raised about the new hate speech law implementation, owing to the fact that it has been used to detain journalists since its inception. For Example, Bekalu Alamrew, a

¹⁶⁷ Ibid para 29(C)

broadcast journalist, was arrested on November 4 /2020 on various accusations under the new hate speech law, including incitement of ethnic violence and disseminating false information¹⁶⁸. Furthermore, because of a Facebook post, journalist Yayesew Shimelis was also the first person to be charged under the new hate speech law. He stated on his personal Facebook page that the government has ordered the preparation of 200, 000 burial sites in anticipation of the danger that COVID-19 poses in Ethiopia.

This is not a positive sign, because we know that in the past years the government has been used repressive laws such as the 2009 Anti-Terrorism Proclamation's broad provisions to arrest, detain, and prosecute scores of journalists, bringing a chilling effect on freedom of expression.

The RPA further explained that criminal sanctions relating to illegal forms of speech can be used as last resort interventions to be enforced only in strictly justifiable case scenarios. It proposes that States should consider civil sanctions and remedies, including pecuniary and non-pecuniary damages, the right to correction and the right of reply, as well as administrative sanctions and remedies, such as those identified and enforced by various professional and regulatory bodies.¹⁶⁹ Similarly, the 2019 Declaration of Principles on Freedom of Expression requires States to criminalize illegal speech as a last resort and only in the most serious cases.¹⁷⁰

Besides, the Human Rights Committee pointed out that in any event, criminal law can be enforced only in the most severe instances and that imprisonment is never an appropriate penalty.¹⁷¹

However, contrary to the above-mentioned useful guidance instruments, the HsDPSP has introduced criminal sanctions, including imprisonment. These criminal sanctions may have a chilling effect that could unduly limit the enjoyment of freedom of speech on the internet.

¹⁶⁸ CPJ, Ethiopian journalist Bekalu Alamrew arrested, accused of disseminating false news '<https://cpj.org/2020/11/ethiopian-journalist-bekalu-alamrew-arrested-accused-of-disseminating-false-news/>' accessed on March 7, 2021

¹⁶⁹ Rabat Plan of Action (n14)para 34

¹⁷⁰ Declaration of Principles on Freedom of Expression and Access to Information in Africa, supra note 80, Principle 23

¹⁷¹ General Comment 34 ,(n21), para 47

4.3.2 The scope of protected groups

Beyond the international duty demanding the regulation of hate speech states may have their own cause and aim to adopt hate speech legislation. Owing to this the substance and scope of the law on hate speech differ across countries. For example, in respect to protected groups, where some States such as Poland and Italy ¹⁷² choose for a closed list i.e. explicitly and exhaustively listing protected groups, and some states like Germany and Hungary ¹⁷³ prefer an open list by adding the words 'such as' or 'and other characteristics' to the text of the law.

Ethiopia is just one of the countries that have a closed list of protected groups, Article 2.2 of the existing hate speech law (HsDPSP) stipulates that ;

Hate speech is a speech that intentionally promotes hatred, discrimination, or attack against a person or a discernible group of identity based on ethnicity, religion, race, gender, or disability.¹⁷⁴

The above provision gives an exhaustive list of protected groups that are ethnicity, religion, gender, race, or disability. It is clear from the above clause that hate speeches against other protected characteristics, such as color, language, political or other opinions, property, birth, indigenous origin or identity, migrant or refugee status, etc., are not protected under this law.

However, in the section which defines the prohibited consequences of hate speech, the term 'discrimination' is defined with open list, as

Any act of denial towards a person or a group on the grounds of ethnicity, religion, gender or disability; or other protected characteristics.¹⁷⁵

Unlike the former provision, the later provision defining the prohibited consequences of hate speech provides an expanded list of protected groups because, in addition to the list of protected

¹⁷² Judit Bayer and Petra Bard, *Hate speech and hate crime in the EU and the evaluation of online content regulation approaches*, Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies, 2020

¹⁷³ Ibid

¹⁷⁴ HsDPSP Art 2(2)

¹⁷⁵ Ibid Art 2(4)

groups in the former provision, it incorporates the word 'other protected characters' which could apply to the protection of other protected characters which are not specified in the former provision. The law has a significant inconsistency problem because one of the consequences of hate speech i.e 'discrimination' is defined with an open list that requires the protection of 'other protected characters' also while the other effects of hate speech that are 'hatred' and 'attack' are protected only to the selected groups that are referred in article 2(2).

In the context of Ethiopia, such codification could be inappropriate. Because when we see the issue of hate speech on the ground, it is not just only about the selected groups under the proclamation, but other protected characteristics have been also targeted by hate speech and worth protection.

For example, there are derogatory terms in social media exchanges such as 'Sefari' or 'Mettie' referring to 'settlers' and 'newcomers' in other regions, which in fact may constitute hateful content. And over the past few years, various dangerous statements have been disseminated via social media that can give rise to violence against groups and individuals who have been deemed newcomers to the specific region. These groups are not covered by the HsDPSP because their group characteristic is beyond the selected that are protected by the under the hate speech law. In general, the legislation fails to encompass all protected characteristics recognized under international human rights law, including color, language, political or other opinions, national or social origin, property, birth or other status, including indigenous origin or identity, migrant or refugee status.

The legislatures do not state whether the selected group characteristics have also been deserving of extra protection, taking into account situations of vulnerability due to entrenched or historic stigmatization, discrimination, long-standing conflicts, and exclusion and marginalization from the political, economic and social life of the society. Instead, its preamble clearly claimed that the motivation for enacting this law was to deter the propagation of hate speech and misinformation from being a threat to civil peace, political cohesion, national unity, human dignity, plurality, and equality. Since the legislation does not cover all group characteristics, arguably it would be inadequate to accomplish its goal.

According to the UN Human Rights Committee, the laws restricting the right to expression, including hate speech, must not only comply with article 19(3) of the ICCPR but must adhere to non-discrimination provisions and must also be compatible with the aims and objectives of the Convention.¹⁷⁶ As with any legal measure to curb hate speech, includes all of the protected characteristics recognized under international Human rights law, but not only limited to the protected characteristics of the HsDPSP, i.e. ethnicity, religion, race, gender, or disability.

Thus the proclamation does not conform with Ethiopia's international obligations of the principle of non-discrimination and national provisions on equality, such as the Constitution Art 25 and the Criminal Code Art 4. Besides it would not be appropriate to address the problem on the ground i.e to regulate the advocacy of hate speech both online and offline under this act, particularly social media hate speech, which is a very pressing issue that needs to be adequately regulated.

4.3 .4 Legal Enforcement towards Social Media Service Providers

Quite few states have adopted laws requiring SMSP to remove speech that is considered illegal within a specific period of time. For example, the HsDPSP and the Network Enforcement Act of Germany impose obligations on companies to remove from their platforms speech that is unlawful.

Article 8.1 of the HsDPSP includes a provision that mandates social media networks to remove hate speech that is being circulated through their platforms in less than twenty-four hours. It requires SMSP to act within 24 hours to remove the circulation of hate speech after receiving notices of certain content.¹⁷⁷ This provision, however, is problematic in that it mandates social networks to delete or remove hate speech from dissemination following receipt of complaints without the requirement for the content to be recognized as illegal by the courts. Furthermore, as this law does not specify who can provide notification, anyone who is not competent to make an authoritative determination on the legality of the content can make a complaint.

¹⁷⁶ General Comment 34 ,(n21), para 26

¹⁷⁷ Ibid Art 8

SMSPs are granted the power to make decisions on what should be considered unlawful content and must be removed. And the real problem is that the proclamation does not provide clear guidelines to social media platforms about how to determine illegal content in the same manner that it does not include guidelines to law enforcement authorities. While business enterprises are expected to uphold human rights under the Guiding Principles on Business and Human Rights.¹⁷⁸ However, business organizations regulate hate speech on their platforms almost exclusively without regard to the human rights implications of their services.¹⁷⁹ Thus in domestic legislations, such as HsDPSP, it is important to offer clear guidance on respecting freedom of expression and other human rights in the determination of illegal content.

In Ethiopia, it is complex and impractical to allow Social Networks to determine speeches that are considered as hate speech, which requires not only to take into account the material and mental state of the individual but also to understand the context of the content. Because to decide whether a speech is unlawful or not to know and understand a language is the essential thing, and then to learn more than 70 languages spoken across Ethiopia is a daunting task by itself.

In addition, it can also be difficult to grasp the context of speech in the communities where the information is disseminated. For example, 'Sefari' or 'Mettie' who have the literal meaning of 'settlers' and 'newcomers' using these terminologies have different contexts in Ethiopia which may constitute hate contents, but which may not have such a sense in the rest of the world.

There are also experiences of using "code" instead of using actual texts in social media communication, which is difficult to grasp as hate speech but can constitute hatred. For instance, there are codes given to describe different groups such as: 'Nefteгна', 'Tebab' and 'Yeken Jeb'. It is common to use these terms by various groups and individuals on a social media platform in such a way that it can lead to violent actions, hostility, and discrimination towards that coded group.

¹⁷⁸ Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework 2011 , principle 11

¹⁷⁹ UN General Assembly ,promotion and protection of the right to freedom of opinion and expression, (A/74/486 2019), para 42

The other major issue, although the HsDPSP can be recognized as a positive attempt to deal with widespread concern with social media platforms and also impose obligations on SMSP to regulate the propagation of hate speech adequately, the legislation neglects to mention anything about what the punishments would be if the SMSP failed to meet their legal obligations.

The German NetzDG Act, which imposes obligations on SMSP, imposes fines on social media that fail to comply with the prescribed obligations, imposes a regulatory fine of between 5 hundred thousand and 5 million Euros.¹⁸⁰ For example, while the company argued that the new legislation lacked clarification, the government imposed its first fine under the new law, NetzDG, to Facebook in July 2019. The organization had to pay €2m for under-reporting of criminal activities on its platforms in Germany.¹⁸¹

However, under the HsDPSP, even the government has a legal ground to ask SMSP to delete illegal content that is posted through their platforms but there is no any legal ground under the law to place penalties on them for failure to meet that responsibility, so the SMSP may be negligent in performing their obligations, which could obviously make it difficult for the legislation to accomplish its purpose i.e. to combat the propagation of hate speech on social media. Even so, It has been over a year since Ethiopia's hate speech and disinformation law was passed, and nothing has changed in terms of the spread of hate speech both online and offline, as well as its negative impact. As a result, it is fair to say that the law is merely a promise that doesn't change anything in terms of the real issue, which is the spread of hate speech.

Besides the majority of social media sites, including the most commonly used site in Ethiopia, Facebook is located in America. And as stated in Chapter three, America has legislation, Section 230 of the Communication Decency Act that indemnifies SMSP in the United States from being found responsible for illegal content on their platforms. And since there is no punishment for SMSP for failure to comply with the obligation under the HsDPSP, the companies may seek to take advantage of the law where they are founded, i.e. the immunity from being responsible for unlawful content on their platforms. And the SMSP may not be motivated to

¹⁸⁰ NetZDG section 4(1)8(2)

¹⁸¹ BBC News:Social Media: How do other governments regulate it it '<https://www.bbc.com/news/technology-47135058>' Accessed on Oct 3,2020

meet their responsibilities, as there are no punishments that they can avoid by fulfilling their duties. Therefore, the legislation should have been drawn inspiration from the German NetzDG Act in order to impose this duty in practice or be supported by enabling regulation clarifying the *modus operandi* of fines for contravention

4.3. Procedural safeguards

While the HsDPSP allows SMSP to remove from circulation hate speech within 24 upon receipt of notices of such communication or post¹⁸², the legislation has not yet made clear the procedure to be established for complaining about illegal content. Rather it sets forward a general provision that stipulates that SMSP shall have policies and procedures in order to meet its responsibilities under the law.¹⁸³ And this provision is ambiguous, which does not provide for a clear procedure to be adopted by the SMSP and availability of information and transparency for complainants or users whose content is subject to removal.

But when we see the German NetzDG Act, it explicitly points out the procedures to be adopted by the SMSP for bringing complain about illegal content. The Act requires the SMSP to maintain an efficient and transparent procedure i.e a procedure that is clearly identifiable, directly accessible, and permanently usable.¹⁸⁴ In addition, it requires the SMSP to confirm receipt of a complaint without undue delay¹⁸⁵ and to notify both the applicant and the author of the content of any decision taken and of the reasons for the decision taken.¹⁸⁶

The existence of such a clear procedure for complaining about illegal content allows complaints to quickly send their complaints, which enables to regulate the dissemination of hate speech on social media as well as its harmful effects. And as far as the required procedure is explicitly established, the SMSP will then know what is actually required to fulfill their obligations, and, similarly, the government would also make them accountable while they have failed to comply with the clear and specific obligations.

¹⁸² HsDPSP Art 8(2)

¹⁸³ Ibid Art 8(3)

¹⁸⁴ Section 3(1) of the OF NetzDG Act

¹⁸⁵ Ibid Section 3(2)(1));

¹⁸⁶ Ibid Section 3(2)(5)

Thus with regard to the procedure for submitting complaints about hate speech on social media, the HsDPSP lacks clarification. And in the absence of clear and specific obligations, it will be difficult for SMSP to satisfy their obligations. Moreover, if the decision-making process by SMSP is not transparent, it may be shielded from public scrutiny,¹⁸⁷ which offers an adequate opportunity for public feedback and discussions to ensure the proper enforcement of the legislation.

The other concerning issue is that the Proclamation does not require SMSP to provide users with the means to appeal towards the content-removal decisions when the author thinks that the decision is illegitimate and is contrary to the right to freedom of speech. As well as concerning the harmful effects for the targeted groups, the law neglected to provide adequate remedies for the infringement of their protected rights such as the right to equality and dignity.

¹⁸⁷ Supra note 171, para 35

CHAPTER FIVE

Conclusion and Recommendations

5.1 Conclusion

The internet, in particular social media, has greatly changed the way people exchange all kinds of messages, both positive conversations and incendiary ones that destroy tolerance and unity among the people. And over the last few years, the inappropriate use of social media in Ethiopia has inflamed ethnic conflict in various parts of the country, leading millions of citizens to be displaced from their habitual residence.

And since it has become apparent that social media has contributed to significant offline damages, the government has adopted a hate speech law called Hate Speech and Disinformation Prevention and Suppression Proclamation (HsDPSP) that regulate various kinds of hate speech both online and offline. In addressing the problem of hate speech, it is necessary to strike a balance between eliminating a speech that actually harms targeted groups and encouraging freedom of expression in society.

This research has examined the adequacy of the existing hate speech legal framework, the HsDPSP, for the regulation of social media hate speech in Ethiopia and finds out the legal framework is inadequate to combat the dissemination of hate speech via social media. Because first of all, the definition of hate speech lacks clarity, is not narrowly defined and there is a significant possibility that its implementation will result in subjective interpretation, which could also infringe the freedom of speech. Moreover, this vague text of the law can make it impossible for people to control their actions.

Second, the legislation does not set out specific requirements for law enforcement authorities that can be met to decide whether or not statements are perceived to be criminal offenses, meaning law enforcement authorities may have the full power to determine a speech that could fall into the definition of hate speech which could even pose a serious threat to inconsistent law enforcement.

Third, the study has found that the scope of protected groups is not consistent with the law, while the definition of hate speech provides an exhaustive list of protected groups that are ethnic, religious, gender, race, or disability, but still in the following section of the law, gives extended protection to other protected groups. Such inconsistencies create uncertainty regarding the application of the legislation. But if it is assumed that the law guarantees protection to the selected groups only the law may not resolve the problem on the ground and yet is contradictory to the international principle on non-discrimination and national provisions on equality.

The study has navigated issues regarding the obligation of SMSP to remove expressions that are considered illegal within a fixed period of time. SMSPs are empowered to make decisions about what should be considered illegal content and it is not clear whether the unlawful content is decided in the context of Ethiopia. Besides, it does not provide users with the means to appeal against the content-removal decisions if the action is considered unlawful.

5.2 Recommendations

As the study has found, the prevailing hate speech legal framework work is not adequate to comprehensively address the spread of hate speech across social media and its negative effects in the offline environment. Consequently, a stronger commitment is required by the Ethiopian Government to make the legal framework effective in halting hate speech on social media, as well as to achieve its purpose and to put it into line with international obligations that strike a balance between the protection of freedom of expression and the prohibition of hate speech. Thus, the researcher proposes the following suggestions that the Government of Ethiopia should consider in order to resolve the pressing issue of the country's social media hate speech.

- To review and amend some of the vaguely formulated hate Speech legislation provisions in order to formulate a definition of hate speech precisely and narrowly.
- To draw inspiration from the RPA and to enact a regulation requiring law enforcement authorities to use the RPA as a framework for deciding whether or not an expression is deemed to be hate speech.
- Reviewing and amending the legislation to ensure that the scope of protected groups inconsistent and compatible with the government's international obligation to the principle of non-discrimination and equality provisions of domestic laws.
- Revising the law or adopt a regulation requiring SMSP to set down a clear procedure for complaining about illegal content and means to pose a challenge against content-removal decisions when a decision is illegitimate.
- As part of a broader effort to combat hate speech, reducing the threat of violence and social tensions by prioritizing programs that consider the contextual causes of hate speech, encourage intergroup communication, and recognize the internal diversity of groups.
- Finally, the government should consider the root causes and generators of hate speech to take appropriate steps to better address the problem on the ground and also to minimize its negative effect.

BIBLIOGRAPHY

Books and Articles

- Alexander Thesis, 'Hate in Cyberspace: Regulating Hate Speech on the Internet' (2001), 38 SAN DIEGO L. REV. 817,
- ARTICLE 19, "Hate Speech" Explained: A Toolkit, 2015 Edition
- Barendt, Eric, *Freedom of Speech*, (2nd ed. Oxford: OUP, 2007) 18-21
- Bhikhu Parekh, 'Is There a Case for Banning Hate Speech' in Michael Herz and Peter Molnar (eds) *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge University 2012)
- C. Edwin Baker, 'Hate Speech' in Michael Herz and Peter Molnar (eds) *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge University 2012)
- Elena Mihajlova, JasnaBacovska, Tome Shekerdjiev 'Freedom of expression and hate speech' 'Skopje : OBSE, (2013)
- F .Cassim , 'Regulating Hate Speech And Freedom Of Expression On The Internet: Promoting Tolerance and diversity ' (2015) SACJ 3, 303
- HateCrime Laws: Practical Guide(2009) OSCE/ODIHR
- Iginio Gagliardone, 'Countring Online Hate Speech' (2015)UNESCO
- Jacob Amedie, 'The Impact of Social Media on Society'(2015) Pop Culture Intersections.
- James Banks , 'Regulating Hate Speech Online ' (2010) I R LCT 233
- James Weinstein,' Hate Speech Bans, Democracy, and PoloticalLegitimacy' (2017). Constitutional Commentary. 465
- Jan H. Kietzmann, Kristopher Hermkens, Ian P. McCarthy, Bruno S. Silvestre, 'Social Media? Get serious! Understanding the Functional Building Blocks of Social Media ' (2011)
- Jennifer Rosenberg, Timeline of World War II From 1939 to 1945
- Johannes Morsink, 'The Universal Declartion of Human Rights: Origions, Drafting, And Intent' 66 (1999).
- Judit Bayer and Petra Bard,*Hate speech and hate crime in the EU and the evaluation of online content regulation approaches*, Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies, 2020
- Jyoti Suraj Harchekar, 'Empact of Social Media on Society ' (2017) IJERT Vol. 6 381

Lashel Shaw, 'Hate Speech in Cyberspace: Bitterness without Boundaries' (2012), Notre Dame Journal of Law, Ethics & Public Policy, 282

Makau Mutua,' The African Human Rights The African Human Rights System a Critical Evaluation'(2000)

Mari J. Matsuda , ' Public Response to Racist Speech: Considering the Victim's Story' (1989) 87 Mich.L.R ,2320

Pesins Antonis ' The Regulation of "Hate speech"; the Meaning of"Incitement" Under the Case-law of European Court of Human Rights and the Jurisdictions of European Union; the United Kingdom andGreece ' (2015)Philosophy and Public Issues (New Series), Vol. 5,

Raohel Cohen-Almagor 'Hate And Racist Speech In The United States a Critique(2015)

Rebecca Zipursky,' Nuts About NETZ: The Network Enforcement Act and Freedom of Expression' (2019) Fordham International Law Journal Volume 42, Issue 4 Article 7, 1328

Report of the Special Rapporteur of Minority Issues (A/HRC/28/64

Ritika Patni; KasturikaKaumudi, 'Regulation of Hate Speech',(2009) 2 NUJS L. Rev. 749

‘Say no to hate speech and xenophobia’, urges Guterres, marking 26 years since genocide in Rwanda

Suzan Bensch, Catherine Buerger, ToneiGlavnic,' Dangerous Speech: a Practical Guide', 2018

TarlachMcGonagle,'The Council of Europe Against Online Hate Speech: Conundrums and Challenges'MCM(2013)005

Waldron .J ’ Dignity and Defamation: The Visibility of Hate’(2010) Harv.L.Rev 1596

Waldron. J ,'The Harm In Hate speech' (2012) London: Harvard University Press 2-3

Yared Legesse Mengistu,' Shielding Marginalized Groups from Verbal Assaults Without Abusing Hate Speech Laws' in Michael Herz and Peter Molnar (eds) *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge University 2012)

Yohannes Eneyew Ayalew ‘Assessing the limitations to freedom of expression on the internet in Ethiopia against the African Charter on Human and Peoples’ Rights’(2020) 20 *African Human Rights Law Journal* 315-345

Laws

- Additional Protocol to the Convention on cybercrime, concerning the criminalization of acts of a racist and xenophobic nature committed through computer systems, (2003)
- African charter on human and peoples right (1982)

- American Convention on Human Rights(1969)
- Criminal Code of the Federal Democratic Republic of Ethiopia(2004)
- Constitution of the Federal Democratic Republic of Ethiopia, (1995)
- Constitution Of the People's Democratic Republic of Ethiopia, (1987)
- Convention on the Prevention and Punishment of Crime of Genocide, (1948)
- Council of Europe, Convention on Cybercrime, (2001)
- Declaration of Principles on Freedom of Expression and Access to Information in Africa,(2019)
- Durban Declaration and Programme of Action of the World Conference against Racism, Racial Discrimination, Xenophobia and Related Intolerance (2001)
- Enforcement on Social Networks '(NetzDG) 2017
- European Convention for the Protection of Human Rights and Fundamental Freedoms(1950)
- Hate Speech and Disinformation Prevention and Suppression Proclamation NO. 1185 /2020,
- International convention on civil and political rights, 1966
- International Convention on the Elimination of All Forms of Racial Discrimination (1965)
- Recommendation No. R (97) 20 of the Committee of Ministers to Member States on “Hate Speech”, (1997)
- Recommendation No. R (97) 21 of the Committee of Ministers to Member States on the media and the promotion of a culture of tolerance, (1997)
- The Camden Principles on Freedom of Expression and Equality(2009)
- The Communication Decency Act (1996)
- The Rabat Plan of Action on the prohibition of Advocacy of National, Racial, or Religious that constitutes incitement hatred to discrimination, hostility or violence(2012)
- The Revised Constitution of the Empire of Ethiopia, (1955)
- The Rome Statute of the International Criminal Court, (1998)
- Universal Declaration of Human Rights,(1948)

Cases

- Benjamin and Others v. Minister of Information and Broadcasting [2001] UKPC 8 (14th February, 2001)
- Brandenburg v. Ohio, 395 USSC 444, 447 (1969)
- Ferdinand Nahimana, Jean-Bosco Barayagwiza, Hassan Ngeze v. The Prosecutor, Case No ICTR-99-52-A, Judgment of 1 October 1946
- Handyside v. the United Kingdom, Application No. 5493/73, judgment of 7 December 1976
- Lingens v. Austria, ECHR ((1986) 8 EHRR 407, 9815/82
- M’Bala M’Bala v. France, 20 October 2015 Application Number 25239/13
- Reno v. American Civil Liberties Union, 521 U.S. 844 (1997)
- The Sunday Times v. United Kingdom, 26 April 1979, EHRR 245, 26 APRIL 1979 case No. 6538/74

Guidance, General comments and Reports

- UN, 'United Nations Guidance Note on Addressing and countering Covid-19 Related Hate speeches', 2020
- United Nations Strategy and Plan of Action on Hate Speech: Detailed Guidance 2020
- ICCPR, Human Rights Committee, General comment No. 34 Article 19: Freedoms of opinion and expression, 2011
- CERD General Recommendation No. 15 on article 4 of the Convention
- Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework

Internet Sources

- Christopher McFadden, 'The Chronological History of Social Media' (2010) IE <<https://interestingengineering.com/a-chronological-history-of-social-media> accessed> Accessed 21 Oct 2020
- OHCHR /The Rabat Plan of Action, 'Freedom of expression vs incitement to hatred' <[https://www.ohchr.org/en/issues/freedom of opinion/articles19-20/](https://www.ohchr.org/en/issues/freedom%20of%20opinion/articles19-20/)> accessed 10 Oct 2020

- DATAREPORTAL, 'Digital 2020 Ethiopia' <https://datareportal.com/reports/digital-2020-ethiopia> accessed 4 Mar 2020
- IDMC. 'Global Report on Internal Displacement' <https://www.internal-displacement.org/countries/ethiopia> accessed 1 Oct 2020
- Abdi Latif Dahir, Nobel Peace Laureate Says Social Media Sows Hate in Ethiopia, *The New York Times*, 10 Dec 2019 < <https://www.nytimes.com/2019/12/10/world/africa/nobel-peace-abiy-ahmed.html> > accessed 5 Oct 2020
- CPJ, Ethiopian journalist Bekalu Alamrew arrested, accused of disseminating false news '<https://cpj.org/2020/11/ethiopian-journalist-bekalu-alamrew-arrested-accused-of-disseminating-false-news/>' accessed on March 7, 2021
- David Kaye, Visit to Ethiopia, 2-9 December 2019 End of mission statement. Ohchr.org <<https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25402&LangID=E>> accessed 8 Oct 2020
- Elias Meseret, Hate speech and disinformation concerns escalate in Ethiopia, *Devex*, 06 May 2020 <https://www.devex.com/news/hate-speech-and-disinformation-concerns-escalate-in-ethiopia_97095?fbclid=IwAR1ZUqpjSMQD6z_cFh0kCjXcGF9_AWAfL3CPR7S6mBUDgbGEJbD62SaUqbnY> accessed October 21, 2020.
- Facebook community Guidelines, 'What does Facebook consider to be hate speech?' <<https://www.facebook.com/help/135402139904490>> accessed 12 November 2020
- Youtube Community Guidelines, Hatespeech <<https://support.google.com/youtube/answer/2801939?hl=en>> accessed 13 Sep 2020
- Twitter, 'Hate content, sensitive topics and violence' <<https://support.twitter.com/articles/20170425>> accessed 13 Sep, 2020
- UN Treaty Collection, ICCPR, New York, 16 December 1966 'https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-4&chapter=4&clang=en' accessed on 22 Nov, 2020
- List Of Countries Which Have Signed, Ratified/Acceded to The African Charter on Human and People's Rights, 'https://au.int/sites/default/files/treaties/36390-sl-african_charter_on_human_and_peoples_rights_2.pdf' accessed on 22 Nov, 2020

- UNGA, Promotion and protection of the right to freedom of opinion and expression, Note by the Secretary-General <https://www.ohchr.org/Documents/Issues/Opinion/A_74_486.pdf> accessed on 2 Dec 2020
- BBC News: Social Media: How do other governments regulate it it '<https://www.bbc.com/news/technology-47135058>' Accessed on Oct 3,2020

Conversation

- Conversation with Yohannes Eneyew Ayalew on hate speech at Addis Zeybe media in this link <<https://www.youtube.com/watch?v=1SqceP7RS8Y&t=420s>>