



**ADDIS ABABA UNIVERSITY**

**COLLEGE OF SOCIAL SCIENCES, ART AND HUMANITIES**

**SCHOOL OF MEDIA AND COMMUNICATION**

**Final Thesis**

**Assessment of Effective Communication Strategies on Raising Awareness of  
Hate Speech and Fact-Checking: A Case Study of the Ethiopian Media  
Authority**

**By**

**Birshewa Seyoum**

**May, 2025**

**Addis Ababa, Ethiopia**

**ADDIS ABABA UNIVERSITY**  
**COLLEGE OF SOCIAL SCIENCES, ART AND HUMANITIES**  
**SCHOOL OF MEDIA AND COMMUNICATION**

**Assessment of Effective Communication Strategies on Raising Awareness of  
Hate Speech and Fact-Checking: A Case Study of the Ethiopian Media  
Authority**

**A Research Submitted to Addis Ababa University, Graduate School of  
Journalism and Communication in Partial Fulfillment of the Requirements  
for the Degree of Master of Arts in Journalism**

**By**

**Birshewa Seyoum (GSR/2815/16)**

**Advisor**

**Mekuria Mekasha (Assistant Professor)**

**Addis Ababa University, Addis Ababa, Ethiopia**

**May,2025**

## DECLARATION

Prepared by: Birshewa Seyoum

Advisor: Mekuria Mekasha (Assistant Professor)

This thesis, prepared by Birshewa Seyoum, titled " Effective Communication Strategies on Raising Awareness of Hate Speech and Fact-Checking: A Case Study of the Ethiopian Media Authority," is presented in partial fulfillment of the requirements for the degree of Master of Arts in Multimedia Journalism. This research is my original work and has not been presented by any other researcher to any university. I declare that all source materials used in this research have been properly recognized and acknowledged.

.

Approved by Board of Examiners

Advisor: Mekuria Mekasha (Assistant Professor) Signature \_\_\_\_\_ Date \_\_\_\_\_

Examiner \_\_\_\_\_ Signature \_\_\_\_\_ Date \_\_\_\_\_

Examiner \_\_\_\_\_ Signature \_\_\_\_\_ Date \_\_\_\_\_

## **Acknowledgments**

Foremost I am thankful to almighty God. I would like to say my most profound genuineness and conscious appreciation to my advisor, Mekuria Mekasha (Assistant Professor), for his veritable and amazing counsel.

I am also grateful to my mother, my sister Hiwot Getachew and friends, who has been a constant source of encouragement and support. Their presence and unwavering belief in me have been a driving force behind my accomplishments.

Finally, I would like to extend my sincerest gratitude to all the respondents who willingly participated in my study, their willingness to share their experiences and insights have been crucial in obtaining meaningful results. Their valuable contributions have greatly enriched this research

## ***Abstract***

*This study explores effective communication methods to raise awareness about hate speech and fact-checking, using the Ethiopian Media Authority as a case study. Both hate speech and harmful misinformation significantly threaten democratic dialogue and social unity, making media regulation and public education crucial. The research reviews the Ethiopian Media Authority's efforts to prevent hate speech, enhance fact-checking initiatives, and encourage media literacy through various communication channels. By examining the Ethiopian Media Authority's achievements, obstacles, and strategies, this research identifies optimal practices for fostering responsible journalism and engaging the public. The results aim to inform policy suggestions and practical measures to reduce hate speech and misinformation in Ethiopia, thereby fostering an informed and inclusive media environment. Qualitative approaches, such as stakeholder interviews and document analysis, are used in the study to deliver a thorough evaluation of the Ethiopian Media Authority's activities and their wider effects on media regulation. The study examines the various strategies the Ethiopian Media Authority employs to combat disinformation and foster media literacy, including social media campaigns, automated monitoring technologies, and public participation programs (such as the 9192 reporting hotline). Despite highlighting initiatives such as journalist training, fact-checking collaborations, and community workshops, the findings also expose difficulties including financial limitations, political prejudices, linguistic diversity, and internet deserts in rural areas.*

**Keywords:** *Hate speech, fact-checking, Ethiopian Media Authority, misinformation, communication strategies.*

<b>Table of Contents</b>	<b>Pages</b>
Acknowledgments.....	i
<i>Abstract</i> .....	ii
Contents .....	iii
List of Table.....	vi
CHAPTER ONE.....	1
Introduction.....	1
1.1. Background of the study .....	1
1.2. Statement of the problem.....	3
1.3. Objective of the study .....	4
1.3.1. General objective .....	4
1.3.2. Specific Objectives .....	4
1.4. Research Question .....	4
1.4. Significant of the study .....	5
1.5. Scope of the study.....	5
1.6. Definition of key Terms.....	6
1.7. Limitations of the study .....	6
1.8. Organization of the Study .....	7
CHAPTER TWO .....	8
REVIEW OF RELATED LITRATURE.....	8
2.1. Historical background and trends of hate speech.....	10
1. Gendered hate speech.....	12
2. Religious hate speech.....	13
3. Racist hate speech .....	13
4. Hate speech on disability .....	13
5. Ethnicity Hate speech.....	14
2.2. Disinformation .....	14
2.3. Historical background of Misinformation and disinformation .....	15
<b>2.4. Misinformation, Disinformation and Malinformation as the three types of information disorder.</b> .....	<b>16</b>
2.5. The Emergence of Fact Checking.....	18
2.6. The Role of Fact-Checking .....	18

2.7.	Communication.....	19
2.7.1.	Communication studies and freedom of speech.....	20
2.7.2.	Effective communication .....	20
2.7.3.	Creating Awareness .....	22
2.8.	Challenges and Barriers .....	23
	Theoretical Frameworks .....	24
I.	Agenda-Setting Theory:.....	24
II.	Elaboration Likelihood Model .....	26
III.	Social norm theory .....	27
	CHAPTER THREE .....	29
	Methodology of the study .....	29
3.1.	Introduction.....	29
3.2.	Reason for selection of the organization.....	29
3.3.	Research Methods.....	30
3.4.	Qualitative analysis.....	30
3.5.	Population and sampling Techniques .....	31
3.6.	Data Collection Tools .....	32
3.6.1.	Document Analysis.....	32
3.6.2.	Focus Group Discussion .....	33
3.6.3.	In-depth Interview.....	33
3.7.	Data Analysis and Interpretation.....	34
3.8.	Ethical Considerations .....	35
	CHAPTER FOUR.....	37
	Data presentation, Analysis and discussion .....	37
4.1.	Introduction.....	37
4.2.	Document analysis in relation to hate speech and fact checking .....	38
4.3.	Response Rate of Respondent.....	39
4.4.	Profile and status of Interview and Focus Group Discussion Informants.....	40
4.5.	Effective communication strategies employed by Ethiopia Media Authority for raising awareness about hate speech and fact checking.....	41
4.6.	The role of Ethiopian media authority in shaping conversation about hate speech and misinformation.....	44

4.7. Challenges and effectiveness in combating hate speech and misinformation: awareness strategies and impact measurement.....	49
<b>CHAPTER FIVE</b> .....	<b>53</b>
<b>CONCLUSIONS AND RECOMMNADATION</b> .....	<b>53</b>
5.1. Conclusion .....	53
5.2. Recommendation .....	54
Bibliography .....	57
<b>Appendices</b> .....	<b>62</b>
<b>Appendix I: Interview questions</b> .....	<b>62</b>

**List of Table**

**Table 1: Interview of informants coding ..... 40**

**Table 2: Informants of focus group discussion ..... 41**

**Figure 1: Gender of the informants ..... 41**

# CHAPTER ONE

## Introduction

### 1.1. Background of the study

In today's digital age, the rapid spread of hate speech and misinformation poses significant challenges to fostering informed and cohesive communities, particularly in Ethiopia, where ethnic and political tensions amplify these issues. Hate speech, defined by the United Nations as any communication that criticizes or employs negative language toward individuals or groups based on identity markers like religion, ethnicity, or nationality, can incite violence and deepen societal divisions (Nations, 2020). Similarly, misinformation, exacerbated by the interactive nature of social media, spreads rapidly as users share unverified content, often driven by the allure of sensational headlines (Williams, 2021).

In Ethiopia, the Ethiopian Media Authority, established under Proclamation No. 1238/2021, plays a pivotal role in regulating media to curb harmful content while promoting ethical journalism. This paper explores effective communication strategies employed by the Ethiopian Media Authority to raise awareness about hate speech and enhance fact-checking, addressing their impact on media practices and public perception in Ethiopia's complex sociopolitical landscape. Hate speech in Ethiopia, often intertwined with ethnic identities and political dynamics, has been linked to real-world violence and social unrest (K. Muller and C. Schwarz, 2020).

The FDRE Constitution of 1994, under Article 39(5), recognizes national groups based on shared culture, language, and territory, making the country particularly vulnerable to divisive rhetoric targeting these identities. The rise of digital media has further accelerated the spread of harmful content, with non-professional actors like bloggers and influencers contributing to what (Wardle C. &, 2017) term "information disorder." The Ethiopian Media Authority, tasked with regulating traditional and digital media, faces the challenge of countering these issues in a context of varying digital literacy levels (Wubetu Demilie and Ayodeji Salau, 2022). The Hate Speech and Disinformation Prevention and Suppression Proclamation No. 1185/2020 underscore

the government's commitment to addressing these concerns, mandating public awareness and media literacy campaigns.

However, the effectiveness of these efforts depends on strategic communication that engages diverse audiences and counters the sophisticated rhetorical tactics such as emotional appeals, coded language and memes used to propagate hate speech. The Ethiopian Media Authority's communication strategies focus on raising public awareness about the dangers of hate speech and promoting fact-checking to ensure access to accurate information. These strategies involve collaboration with government bodies, civil society organizations, and tech developers to create a resilient media landscape. Public awareness campaigns educate citizens on recognizing hate speech, defined by the Council of Europe as expressions that incite racial hatred, xenophobia, or discrimination (Agata de Latour, 2017).

The Ethiopian Media Authority also encourages media practitioners to adhere to ethical journalism standards, reducing the spread of harmful content. Fact-checking initiatives are critical, targeting rumors and false information that fuel hate speech. These efforts are supported by technological solutions, such as automated detection systems, though their implementation is challenged by low digital literacy among the populace. By leveraging media platforms and working with influential figures like government officials and religious leaders, the Ethiopian Media Authority shapes narratives that promote critical thinking and dialogue, fostering a more informed public.

The Ethiopian Media Authority's communication strategies have made strides in addressing hate speech and misinformation, but their impact requires ongoing evaluation. By analyzing the rhetorical strategies of hate speech, such as pseudo-scientific appeals or visual elements like memes, the Ethiopian Media Authority can develop advanced fact-checking tools to detect both overt and covert harmful content. These insights guide targeted education campaigns and improved content moderation algorithms, enhancing media literacy and public discernment. Collaborations between the Ethiopian Media Authority, civil society, and tech developers are crucial for creating to a media ecosystem that upholds democratic ideals while mitigating the risks of social unrest. Moving forward, the Ethiopian Media Authority must adapt its strategies to address evolving digital challenges, ensuring that communication efforts remain effective in promoting a tolerant and informed online environment. Through these multifaceted approaches,

the Ethiopian Media Authority contributes to a more cohesive and resilient Ethiopian society, capable of navigating the complexities of hate speech and misinformation.

## **1.2. Statement of the problem**

Hate speech poses a significant threat to social harmony, political stability, and national unity in Ethiopia, particularly in the age of social media. The unchecked proliferation of hate speech, often disguised as free expression, risks inciting civil unrest and deepening societal divisions. According to Craig Martin's study, *Striking the Balance: Hate Speech in Japan, The United States, and Canada*, hate speech targeting groups based on race, ethnicity, religion, gender, or sexual orientation through severe condemnation, contempt, or dehumanization causes considerable harm (Denu, 2021). In Ethiopia, the rapid growth of social media platforms like Twitter and Facebook has amplified the spread of hate speech, fueling animosity and polarizing communities (Fátima, 2023).

The 2018 liberalization of Ethiopia's media landscape, while expanding press freedom, unintentionally intensified ethnic divisions. (Workneh, 2019). Highlights how social media's "outrage communication" has deepened identity-based conflicts, with platforms serving as outlets for political dissatisfaction. (Skjerdal, 2021) notes that this period saw increased ethnification of media, where audiences trusted ethnically aligned sources while dismissing others, exacerbating polarization. This environment also spurred the creation of fact-checking organizations to counter misinformation. During the Tigray conflict, both government entities and the Tigray People's Liberation Front actively spread political disinformation, further complicating the information landscape (Workneh, 2021).

In response, Ethiopia's House of People's Representatives enacted Proclamation No. 1185/2020, the Hate Speech and Disinformation Prevention and Suppression Proclamation, tasking the Ethiopian Media Authority with monitoring and enforcing regulations to protect public interest, human dignity, and equality. The Ethiopian Media Authority employs awareness campaigns and fact-checking initiatives to combat harmful content. However, challenges persist, including the growing volume of user-generated content, which strains content moderation efforts. Legal frameworks often place the burden on communication platforms to address flagged comments, raising concerns about free speech when private companies, rather than democratic processes,

determine content removal (Alkiviadou, 2019). Moderators also face emotional burdens due to the volume and intensity of harmful content (S. Schäfer, 2023).

Despite these efforts, hate speech on social media continues to threaten peaceful coexistence in Ethiopia (Muluken A.I, Mulatu A. and Biset A., 2021). The effectiveness of the EMA's communication strategies in raising awareness, shaping public discourse, and curbing harmful content requires evaluation. This study examines the Ethiopian Media Authority's current approaches, identifies gaps, and proposes actionable solutions to enhance its role in fostering a better-informed and cohesive society. It also explores the challenges and limitations the Ethiopian Media Authority faces in addressing hate speech and disinformation, aiming to strengthen its contribution to Ethiopia's social and political stability.

### **1.3. Objective of the study**

#### **1.3.1. General objective**

This study aims to investigate the Ethiopian Media Authority's strategies on raising awareness about hate speech and fact-checking, identifying effective practices, challenges, and opportunities to enhance media literacy, responsible speech, and regulatory measures against misinformation in Ethiopia.

#### **1.3.2. Specific Objectives**

The specific objectives of the study is;-

- To evaluate existing communication strategies used by the Ethiopian Media Authority in addressing hate speech and misinformation.
- To measure effective educational campaigns utilized through social media tools to inform the public about the dangers of hate speech and the significance of verifying information.
- To assess the effectiveness of the communication strategies in raising awareness and reducing the incidence of hate speech.

### **1.4. Research Question**

1. What is the most effective communication strategies used in Ethiopia Media Authority for promoting awareness about hate speech and fact checking?

2. What role does Ethiopia Media Authority play in shaping conversations about hate speech and misinformation?
3. What are the challenges combating hate speech and misinformation and how can we measure the effectiveness of communication strategies in reducing incidents and promoting fact checking?

#### **1.4. Significant of the study**

This study has significant implications for media regulators, policymakers, journalists, and civil society players who aim to combat hate speech and disinformation in Ethiopia and in general. It provides significant insights on how the Ethiopian Media Authority employs fact-checking and strategic communication as mechanisms to combat mis/disinformation that fuels ethnic conflict and political instability.

The study emphasizes that in order to create a more critical audience and stop harmful content from being amplified, media literacy and public awareness initiatives are crucial. By examining the tactics used by the Ethiopian Media Authority to combat online hate speech, such as partnerships with tech companies, community-based fact-checking projects, and quick reaction systems.

The finding can guide changes to media regulation policies, guaranteeing that fact-checking initiatives strike a balance between the need to prevent harmful disinformation and the right to free speech. The significance of institutional credibility is emphasized by the study.

#### **1.5. Scope of the study**

The aim of this research is to investigate the communication methods utilized by the Ethiopian Media Authority in their efforts to address hate speech and encourage a culture of fact-checking within Ethiopia. It explores Ethiopian Media Authority's multi-channel approach, which involves various digital media platforms, along with community outreach initiatives, media literacy training, and information verification. The research will also consider different cultural contexts, as well as the challenges faced by fact-checking in a rapidly evolving information landscape. The study's conclusions are intended to offer valuable information that could lead to improved

policies and enhanced strategic communication methods to combat hate speech and misinformation in Ethiopia and similar contexts

## **1.6. Definition of key Terms**

**Fact-Checking Mechanisms:** The processes employed by the Ethiopian Media Authority to verify the accuracy of information, utilizing digital tools, manual verification by trained personnel, and collaboration with journalists and community organizations to identify and counter misinformation.

**Communication Strategies:** The methods used by the Ethiopian Media Authority to disseminate verified information to the public, leveraging social media platforms, traditional media channels, and public awareness campaigns to ensure accurate information reaches diverse audiences.

**Hate Speech Reduction:** The initiatives undertaken by the Ethiopian Media Authority to identify and mitigate hate speech, including implementing regulatory policies, conducting public awareness efforts, and partnering with online platforms to address harmful content.

**Public Engagement & Media Literacy:** The Ethiopian Media Authority's programs designed to educate citizens on media literacy, emphasizing critical thinking skills and responsible information-sharing practices to reduce the spread of misinformation.

**Challenges & Recommendations:** The obstacles faced by the Ethiopian Media Authority, such as limited resources, political influences, and the proliferation of online misinformation, along with proposed strategies to enhance communication effectiveness and address these challenges.

## **1.7. Limitations of the study**

The study contains limitations and shortcomings of its own. The following are some of the main factors that influenced the study's limitations: It was quite difficult to get respondents in a short period of time, and conducting in-depth interviews presented difficulties for the study. It was challenging to organize and conduct the interviews effectively because the specialists weren't always present at the office. Consequently, after multiple tries, the in-depth interview was carried out.

Limitations of this research were the impact of unfavorable weather conditions on collecting data in rural communities, where it was not feasible to have structured one focus group discussions. Rural parts of Ethiopia are often exposed to extreme heat. Because of their concern, respondents showed little interest in taking part in group talks, and it frequently took a long time for them to come to an agreement.

Another limitation of the study is the difficulty in obtaining sufficient responses within a short period of time. This challenge affected the sample size and may have limited the representativeness of the findings.

## **1.8. Organization of the Study**

This study was organized into five chapters. The first chapter was made to have the introduction of the study that contains the problem statement of the study, objectives, research questions that were tried to be answered at the end of the research, the significant, scope and limitations of the study. The second chapter contains review of literature. Chapter three was made to include the methodology of the study. The findings, presentations and discussions of the research are presented in chapter four of the study. The summary, conclusions and recommendations of the study are the final part of the research encompassed in chapter five.

## CHAPTER TWO

### REVIEW OF RELATED LITRATURE

Hate speech appears in various social settings like political discussions, forms of art, competitive sports, or workplaces. Yet, the quick advancement of digital technology, particularly social media, has introduced new difficulties in grasping this severe issue. There is no single accepted definition of hate speech. Its interpretation is often debated among different perspectives, including many that are less understood and not from the Western world. Generally, hate speech targets an individual or a group, often focusing on those belonging to a social minority. As a result, it can be described as sexist, racist, xenophobic, ageist, fat phobic, or homophobic, among other labels. Those who engage in hate speech might, for instance, direct their hostility toward women, Black individuals, immigrants, the elderly, or people with disabilities, and so on. The (Nations, 2019) points out that hate speech involves attacks based on inherent characteristics like race, nationality, or gender.

Hate speech lacks a unified academic definition. According to (Matsuda, 1993), it represents messages filled with persecution, animosity, and derogatory views about low racial self-esteem, all aimed at groups that have been oppressed throughout history. To provide a broader understanding, Hylton (1996) describes hate speech as being rooted in factors like race, religion, nationality, gender, sexual orientation, disability, and other identity traits, often intentionally provoking violence and bias toward particular individuals or communities. It conveys animosity towards a person or group based solely on their identity traits. Typically, hate speech manifests through offensive and hostile language. It particularly targets vulnerable individuals or specific minorities. As a result, hate speech may act as a "mechanism for violent subversion," involving tactics such as fear, harassment, intimidation, and discrimination (Delgado, 1995). While those who are victimized can respond with their own words, the social environments fostered or sustained by hate speech pose serious risks and disadvantages to members of fragile social groups, making it difficult for their voices to be heard.

Hate speech serves as a harmful method for inciting widespread violence, including bloodshed, civil unrest, and genocide. It limits the rights of minorities to express themselves by creating an atmosphere that makes them hesitant to address the various forms of discrimination they endure.

When the scale of opinions shifts between violations of human dignity and hate speech, issues related to hate speech and global law arise. (Barendt, 2019) points out that hate speech is an especially provocative form of communication that is not aligned with the notion of free expression. (Delgado, 1995).

It is essential to clarify what "hate speech" means and the standards that may be applied to identify it. Limits on hate speech can potentially violate free speech rights, and certain standards may be enforced. Laws against hate speech that are overly broad may infringe upon the right to free expression, while insufficient regulation can result in bullying, harassment, and violence directed at minority and protected groups. It is important to emphasize that hate speech should not be mistaken for offensive speech, as the right to free expression includes robust, critical, or upsetting comments. (Elliott, 2016) Explain that "freedom of expression" involves communicative activities, often political, that may be shared or received without restriction or censorship, particularly from government entities. Therefore, the concept of freedom of expression is complex, involving various debates about key political principles and ideologies, along with their understanding and regulation. (Yong, 2011) explored whether hate speech falls under the umbrella of free expression and sought to distinguish among various forms of hate speech. Direct defamation is considered open speech, while scattered defamation and organized political advocacy for exclusionary or harmful actions are regarded as unprotected speech.

According to (Mafeza, 2016), the term hate speech usually describes remarks that are offensive, abusive, insulting, intimidating, or harassing, and/or provoke violence, hatred, or discrimination against groups defined by specific characteristics. Language that seeks to promote hate, bias, or violence should be restricted by international laws. Nevertheless, the challenge in controlling hate speech arises from the ambiguous nature of defining what is considered a crime, which can lead to the punishment of speech that does not intend to or truly is likely to foster hatred.

The ordinary concept is constituted by how ordinary people use the term 'hate speech' and this is likely to be influenced by how legal professionals and scholars of law use the term 'hate speech' and vice versa. So if the state decided to refrain from line-drawing, the ordinary concept hate speech might not exist in exactly the same form. But this concept it reflects ordinary people's sense of what is acceptable and unacceptable speech in relation to groups or classes of persons

identified by protected characteristics, based on ordinary people's moral values and principles. Now it is also true that the ordinary concept hate speech probably includes the idea that hate speech is regulating able speech (in the sense that there is a preliminary reason to regulate it). But at least as regards to the ordinary concept this feature of the concept does not depend on the state's ratification. Ordinary users of the term 'hate speech' could think that hate speech is regulate able speech even if the state itself decided to withdraw from line-drawing in relation to acceptable and unacceptable speech and/ or there was no immediate prospect of hate speech actually being regulated. (Brown, 2017)

The realm of social media platforms has recently emerged as a space for hate speech worldwide due to the lack of regulation on user-generated content. In contrast to traditional media, social media lacks a system for gatekeeping and editorial standards that review each submission before it is shared. Research indicates that, for these reasons and others, social media acts as a breeding ground for bullying, prejudice, discrimination, bigotry, and intolerance linked to various social identities such as ethnicity, religion, nationality, gender, and race (Carlson, 2017; Cohen-Amalgor, 2011; McGonagle, 2002). All of these factors support the idea that hate speech is an assault on individuals' identities.

Researchers also claim that hate speech does not exist in isolation, as the online and offline environments influence one another (Althoff et al., 2017; Tirrell, 2017). It often coincides with economic and political turmoil. Additionally, the ease of spreading hatred on social media platforms plays a role (Minority Rights Group International, 2014), which is evident in Ethiopian ethnic-focused media like the Oromo Media Network, Amhara Satellite Radio and Television, and Dimetsi Weyane TV. Therefore, historical, political, and modern socio-political factors serve as indicators for assessing both online and offline hate speech in the nation. As a result, during the shift from the Ethiopian People's Revolutionary Democratic Front government to the reformist Prosperity Party (PP), hate speech appears to be on the rise in social media. (Muluken Asegidew Chekol, 2023)

## **2.1. Historical background and trends of hate speech**

As indicated by (Brown, 2017), the term hate speech is think up by a group of legal scholars in the late 1980s in the United States in response to what they adage as the means diverse legal

systems attempted convinced classes of destructive bigoted speech. In the past, there has been a soaring acceptance of racially prejudiced expressions and discrimination, but after World War II, and in particular since the 1960s, the common trend is that European countries have brought hate speech under increasingly more strict regulation; the USA remains one of the very few countries to refuse to accept the trend to ban hate speech.

Hate speech found on social media can be expressed through words (like posts, comments, articles, etc.) and non-verbal means (such as emoji's, stickers, images, etc.). These multimedia assaults promote and strengthen stereotypes that arise from harmful language. They can vary from simple name-calling to extreme demands for violence and genocide. At times, these expressions originate from strong feelings and gain popularity online, spreading across different platforms (Fátima, 2023). As a result, they have negative impacts on both individuals targeted and society at large, weakening democratic spaces for discussion.

Regulating hate speech is not a simple issue. Sometimes it is driven by nationalist groups or far-right parties, going hand in hand with disinformation and conspiracy theories. In the name of combating hate, authoritarian states also have passed vague laws that censor the public sphere (Garbe, 2023).

Hate speech appears in more varied and intricate forms across social media platforms. Its rapid spread can influence actions far beyond its original location. Hatred is everywhere, engaging users in multiple ways through various media. It is accessible around the clock, influencing a much wider group of people. On social media, those who express hate can hide their identities and connect with others who share their hostile views. This summary offers a limited overview and undeniably has several theoretical shortcomings that require further development. (Fátima, 2023).

*“Hate speech is” any kind of communication in speech, writing or behavior, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor” (S. Pukallus, 2024).*

Social media algorithms facilitate the rapid spread of hate speech. This harmful content can include false information, gossip, insults, jokes, ridicule, and hatred toward women, intolerance

of religions, dehumanizing attitudes, stereotyping, broad statements, and bias. From a communication perspective, hate speech is most effectively seen as having three aspects: methods of public communication, rhetorical tools used in hate speech, and the channels through which it is shared.

The problem of hate speech, especially in relation to social media, can be examined using three main definitional perspectives. Initially, some definitions cover a broad spectrum of harmful expressions, which includes racism, homophobia, and cyberbullying (Carlson, 2021). The second perspective highlights the particular groups that hate speech targets, including those defined by characteristics such as race, religion, or gender. Lastly, the third perspective stresses the intention to cause harm, associating hate speech with intense, ingrained feelings directed towards the removal of a specific group (Halperin, 2011).

On social media, hate speech proliferates due to a two-fold structure: the mechanical design of platforms, which enables rapid circulation of harmful content, and the purposeful use of communication strategies that amplify hate. Hate speech is communicated through various rhetorical styles and devices, such as memes or coded language, which make it difficult to detect and regulate. These strategies contribute to the pervasive nature of online hate speech, posing challenges to social harmony and fostering potential violence (S. Pukallus, 2024).

Hate speech does not target based on only single identity. It can target on the basis of gender, religion, race, and disability (Seglow1, 2016). These are some of hate speech (Naganna Chetty, 2018):-

### **1. Gendered hate speech**

This refers to a type of expression based on gender or sex. Women and girls are often the targets of this hate speech. Due to their gender identity, there is a deliberate violence directed at women and girls around the world. This is referred to as sexist hate speech, a form of social humiliation that aims to insult women and create fear and insecurity among them within society. The Internet's widespread access, the rapid development of information and communication technologies, and the frequent use of social media have made it much easier to portray violence against women and girls. These developments are being utilized as means to harm women and

girls. The issue of online violence against women and girls is recognized as a worldwide concern.

## **2. Religious hate speech**

This form of expressing hatred targets religions like Islam, Hinduism, and Christianity. Since religion consists of groups, the harmful impact of hate speech is greater when directed at a religion rather than at an individual. Online, Muslims face demonization and negative portrayals through discrimination, stereotypes, physical assaults, and harassment, all intended to incite violence. The rise of anti-Muslim hate online highlights the need to tackle the problem of Islamophobia on social media platforms. By looking at the activities in online communities, such as the content they publish, share, and endorse, we can analyze their behaviors (Awan, 2016). Muslims are often presented as a uniform out-group associated with violence, conflict, and extremism (Tornberg, 2016). The internet serves to amplify and reinforce the existing narratives, leading to more extreme divisions. .

## **3. Racist hate speech**

Hate speech that targets the looks of individuals or groups is referred to as racist hate speech. Typically, this type of communication occurs on an international scale. The rate and effect of this kind of speech rely on the motives and views of a country's government, which can vary depending on the leadership in place. Tatum has suggested that racism functions as a system that includes cultural messages, institutional policies, and practices, along with individual beliefs and actions (Tatum, 2001) .According to (Wodak, 1999), p. 181), racism is seen as both a syncretic ideology and a social practice that discriminates, which can be supported and established by dominant social groups. This shows that within a system or environment, members of one group may exert their dominance over another group or person based on visible traits like skin color.

## **4. Hate speech on disability**

Hate speech directed at an individual's physical and mental state is known as hate speech regarding disability. Disability is seen as a social category, similar to race and gender, rather than a purely medical issue. It refers to any health condition that hinders a person's ability to engage in certain daily activities. Thanks to improved medical diagnostics and treatments, individuals

can live longer lives with the aid of support tools, which can lead to disabilities. Any person can experience disability at any stage in life, encompassing all protected identities like race, gender, nationality, and generational aspects. Those without disabilities are often seen as temporarily able-bodied (Kudlick, 2003). Disabled individuals are at a higher risk of encountering hate speech compared to those who are able-bodied. The hate speech aimed at disabled individuals stems from the offender's viewpoint on disability, rather than the actual disabilities themselves (Hollomotz, 2013). Numerous structural barriers exist that legally restrict parental rights and undermine sexual freedom, impacting the sexual autonomy of disabled individuals (Stevens, 2011).

## **5. Ethnicity Hate speech**

Ethnicity is a multifaceted and changing phenomenon and can be viewed in different ways. Various scholars have considered different aspects like country of birth, nationality, language spoken at home, skin colour (an aspect for consideration for some and not for others), national/geographical origin and religion (Kimotho, 2016). Ethnically affiliated and regional state media tend to nurture themselves alongside the ethnic groups they presume to represent. This has been intensified by political contestations among regional state government politics ascribed to identity politics (Nigussie, 2024). Inter-ethnic induced hate speech resulted in insidious consequences, ranging from psychological harm on the target individuals and groups, to the incitement of genocide, displacements, civil war and other mass violence. (Ahmed, 2021)

## **2.2. Disinformation**

In the current digital age, the distribution of false information poses a significant issue that can negatively impact individuals, communities, and society as a whole. Disinformation refers to the intentional sharing of misleading or false information to deceive or manipulate others. This differs from misinformation, which occurs when incorrect information is shared unintentionally. Although both types can have harmful effects, disinformation is more damaging as it is deliberately created and disseminated for a specific purpose. This can involve topics related to politics, finance, or simply inciting chaos.

In the quest for understanding in the midst of misleading information, (Atif Ahmad, 2019) offer a clear definition and a structured model, based on the theory of situation awareness. Their findings highlight how disinformation is strategically used to undermine awareness and affect decision-making processes, creating a significant obstacle for organizations in various fields. As discussions about misleading information progress, Ruiz and Nilsson in 2022 suggest a detailed framework that explains the complex ways disinformation spreads on social media. Their definition captures the hostile aspects of disinformation efforts, utilizing a range of persuasive techniques to take advantage of controversies related to identity.

### **2.3. Historical background of Misinformation and disinformation**

The amount of false and misleading information available today has reached an all-time high in 2020. "Actually, misinformation has existed for a long time, with roots in the battles of false information in ancient Rome." However, in the twenty-first century, this knowledge has been used as a powerful tool like never before (Posetti, 2019). Among the many instances of misinformation and disinformation throughout history, Posetti (2019) pointed out that during ancient Rome, misinformation, disinformation, and propaganda were utilized when the Roman general and politician Antony encountered Cleopatra. Additionally, the Roman leader Octavian launched a campaign filled with propaganda aimed at damaging Antony's reputation, employing sharp phrases on coins that resembled early versions of Tweets.

In old times forms of writing inscribed on materials like stone, clay, and papyrus appeared several thousand years ago, during that period the information these writings was usually limited to the leaders of the group (emperors, pharaohs, Incas, religious and military leaders, and alike), (<https://journals.ala.org/index.php/ltr/article/view/6497/8636>.)

These slogans painted Antony as a womanizer and a drunk, implying he had become Cleopatra's puppet, having been corrupted by his affair with her who ultimately created distrust and division (Posetti, 2019). In the same way, Posetti referred the 1947-1991 – Cold War period, and how the international broadcasters were harnessed to influence populations to take sides.

However, in the late twentieth century, the internet has been provided new means for disseminating fake news on a vastly increased scale; according to world economic forum

describes a massive digital misinformation as is becoming a pervasive in online social media to the extent as one of the main threats to our society.

Distortion of facts, production of fake information, biased information, advertising, and information intended to denigrate views and ideals have all been part of society in the past, even on traditional platforms (Posetti, 2019).

Disinformation, misinformation, contaminate the information space worldwide and the trend of influencing facts continues to disrupt public communication, distort public confidence and, consequently affects democratic processes in societies.

#### **2.4. Misinformation, Disinformation and Malinformation as the three types of information disorder.**

The inability of citizens to tell apart truth from false information can weaken the political system, especially during elections in democratic nations. This confusion can stop voters from making choices based on accurate information, while a rise in skepticism can lead to long-lasting issues in society by decreasing trust in reliable and credible public institutions. Additionally, by taking advantage of the weaknesses within the digital media landscape, the spread of false information can deepen social divides. This division, driven by disordered information, has become a major hurdle for democratic communities. With rising polarization, individuals are growing more disconnected from each other and forming “truth publics”, each with their own separate realities and stories online. Therefore, it is essential to view digital and social media platforms as places where suitable strategies and policies can be created to combat information disorders that result in various harmful social and political consequences. However, establishing a thorough strategy requires proper categorization and classification of information disorders. Many researchers argue that without a full categorization, identifying the problem becomes significantly harder, making it nearly impossible to implement effective solutions that address the core issues. Moreover, if all forms of false or misleading information are treated as dangerous and needing intervention, it can severely threaten the right to free speech (Ekin Balkan, 2023).

Misinformation refers to information that is incorrect but not done on purpose; it can happen due to a journalist’s error, for instance. In contrast, disinformation is intentionally created, shared, and spread to inflict damage or make money. The term malinformation describes truthful

information that is shared without consent or aimed at causing harm. Instances of malinformation can involve leaks, bullying, and offensive speech. (Wardle C. &, 2017).

The spread of false or misleading information often has a dynamic nature, leading to a mix-up of various kinds of information confusion. On one side, disinformation might turn into misinformation. For instance, someone who produces disinformation might purposely share false claims on social media. Individuals who encounter this information might not realize it is wrong and could spread it within their circles, presenting it in their own way. Conversely, misinformation can also change into disinformation. For example, an article that satirizes news may be spread deliberately out of context to trick audiences.

(Wardle C. , 2019), describes major types of Mis and Disinformation as follows:-

**Satire or parody:** It refers to the use of humor or exaggeration to criticize or condemn a person, organization, or policy. This is also a method of criticizing society, a certain group of individuals, or using comedy, irony, or scorn. When satire or parody is made with the intention of entertaining the public through fiction or parody, it has the potential to deceive and mislead the audience.

**False connection:** indicate using of headlines, visuals or captions that doesn't support and articulate the content.

**Misleading content:** it refers to the application of ambiguous of information to frame an issue when facts/information is consumed and misrepresented or skewed.

**False context:** When Information spreaders takes an image, a quote or other piece of content and puts it into a new, false context to change its meaning, when a genuine content is shared with false contextual information.

**Imposter content:** it refers when well-known name, brand or logo to fool people into believing that it is authentic content uses falsely. Meaning genuine source of news outlets or government agencies, are impersonated.

**Fabricated content:** when entire message compiled information that is 100% false and made-up designed to deceive people as well as do harm.

**Manipulated content:** when a certain genuine information or imagery is changes through another material either by adding or deleting the first original material intended to manipulate and to deceive through tools of deep fakes or other kinds of manipulation of audio and/or visuals.

## **2.5. The Emergence of Fact Checking**

The rise of fact-checking as a specialized editorial function demonstrated how essential accuracy became within the field of journalism after the Progressive Era. In the 1920s, Time magazine and The New Yorker established the first dedicated fact-checking departments. This approach eventually spread to other national magazines, which started to compete with newspapers and each other in their efforts to present a reliable and thorough understanding of current events. By the late 20th century, it had become normal practice for magazines in North America to have a formal fact-checking phase; one study indicated that in the 1980s, well-known newsweeklies had one fact-checker (often referred to as a researcher or associate editor) for every three or four reporters. In addition to its original meaning, fact-checking has increasingly come to describe the assessment of statements or texts that are already public, rather than just those intended for publication. This type of external or "political" fact-checking involves not just verification but also the possibility of confrontation, where the fact-checker either supports or disputes the accuracy of a claim made by an individual or organization. Although those engaged in this new role express their aims in various ways, their general objective is to fight against misinformation, such as by convincing readers not to accept false claims or encouraging politicians to communicate more cautiously (Amazeen, 2019).

## **2.6. The Role of Fact-Checking**

Fact-checking is described as the practice of seeking factual information in order to promote its truth value. (Miller, 2020). Thus, taking fact-checking as practice related to teaching practices. Integrating fact-checking into digital literacy teaching practice may encourage learners to practice with digital fact checking tools and to identify information in the digital platforms. The main focus of the workshop on fact-checking was this integration. The tools for digital fact-checking and the media environment were examined. Participants analyzed media content by employing strategies and tools for fact-checking. They also produced fact-checking reports. Additionally, there was a reflection on their views and critical understanding of the media, encouraging them to take steps to address issues affecting their lives and careers (Tekoniemi, 2022).

Fact-checking has often been identified as a key antidote to misinformation (Azoulay, 2024), and its study has coincided with a growing scholarly interest in disinformation. Recent research has noted a shift in the way fact-checking is done, referred to as a "debunking turn". " This change means that fact-checking now prioritizes tackling false information that spreads quickly on social media instead of concentrating on political statements. One reason for this shift is the increasing influence of social media platforms.

Nowadays, social media platforms are affected by manipulative online communication which consists of negative and destructive contents. Its contents deliberately generated by different participants. (Quandt, 2018) described this manifestation as "Dark Participation". "Misinformation and propaganda can also take the form of hate campaigns that attack specific groups or individuals that symbolize these groups". Organized hate speech in online forums applies demeaning language intentionally, and hate in this context is not to be confused with situational rage (although it seems to be emotionally loaded, which is part of a strategy to appeal to specific target groups).

The impact of technology and social media right now, there is not much organized work in Ethiopia focused on creating strong fact-checking systems to address the harmful effects of false information on the community. The way news and information are received is greatly shaped by government-related media, along with political and cultural influences. Thus, understanding the situation that fact-checkers in Ethiopia encounter is crucial, and it is necessary to come up with fact-checking plans that fit the unique aspects of the Ethiopian environment (Endert, 2021).

## **2.7. Communication**

Communication plays an important role in human life. According to (MOGHADDAMA, 2022)Thao (2005), communication is introduced as a process in which a concept or a message is converted from senders to receivers. It is explained that the sender transforms their thoughts and ideas into messages, while the receiver interprets those messages back into thoughts. Communication allows individuals to share messages successfully and to exchange understanding (MOGHADDAMA, 2022).

### **2.7.1. Communication studies and freedom of speech**

The modern communication discipline traces its origins to the early 20th Century (Keith, 2008). From the outset, communication scholars were slow to appreciate the important role that freedom of speech plays in a democratic society. During the initial years of the field, there was surprisingly a lack of interest in the subject, as shown by the few courses available on freedom of speech and the minimal discussions found in the prominent communication journals of that time (Haiman, 1964). In the early 1960s, a small group of academics from the Speech Association of America, which later changed its name to the Speech Communication Association and then the National Communication Association, suggested setting up a Commission on Freedom of Expression (Haiman, 2003) . After some reluctance, the leadership of the Speech Association of America acknowledged this group, leading to the Commission organizing programs during the annual convention and launching the Free Speech Yearbook, which was later renamed First Amendment Studies, in 1962.

While hate is described as strong negative emotions towards others based on certain aspects of their identity (Waltman M. S., 2011), hate speech refers to communication that aims to condemn or criticize another person's identity (Waltman M. S., 2017). Although hate speech shares similarities with the Supreme Court's definition of "fighting words", it also has some distinct features. Hate speech is more than name calling. It is an effort to damage someone else's identity to the point where their humanity and worth are doubted. Comparing African Americans or undocumented immigrants to animals or savages rhetorically constructs them as uncivilized and unworthy. And then, when these same groups are constructed as threats to white people or traditional American values, calls for their elimination (to remove their threat) are often not far behind. Put differently, hate speech is often used to construct someone as worthy of killing before actual hate crimes and ethno violence take place.

### **2.7.2. Effective communication**

The goal is to foster positive change, empower communities, and enhance the well-being of individuals. By leveraging effective communication strategies and approaches, development efforts can be better planned, implemented, and evaluated. Communication plays a pivotal role in

creating awareness, mobilizing resources, fostering collaboration, and empowering individuals and communities to actively participate in the development process.

Many authors have explained and defined communication in various ways. Among these authors are (Kumar, 2011) each has their take on how to define this term. As a result, communication is defined and understood from the start. Communication, as a process, significantly aids in making people understand their surroundings. People need to be educated about the advantages of literacy, the strengthening of Messages sent via communication channels will have a positive impact on them, and the role of the media in this regard is critical.

A system of checks and balances, which includes communication as a key element, is designed to foster accountability within government agencies, manage conflicts of interest, increase public participation, and prevent corrupt practices. Effective management in the public sector relies on these accountability systems (Worldbank, 2006).

(Hargie, 2016), explained that Effective communication plays a crucial role in boosting employee motivation by providing clear information about their tasks, performance, and ways to enhance their work if needed. Moreover, communication serves as a valuable source of information for organizational members, aiding the decision-making process by identifying and evaluating different options and courses of action.

Communication is instrumental in shaping an individual's attitude, as being well-informed tends to result in a more positive outlook compared to those who are less informed. Various means of communication such as organizational publications, meetings, and both written and spoken forms of communication contribute to this process of molding attitudes (Morley, 2002). In addition, Communication also helps in socializing. In modern life, being around another person encourages communication. It is also said that one cannot survive without communication (Paul, 1991). As discussed earlier, communication also assists in controlling the process. It helps control organizational members' behaviors in various ways.

Effective communication plays a vital role in controlling the functions of management. Communication serves as the foundation of institutions and enables the understanding of people, groups, and organizations. It facilitates the exchange of information between the organization and its environment, influencing the organization's structure, comprehensiveness, and scope.

Effective communication is crucial for delegation of tasks, achievement of objectives, and implementation of discussions. Without structured communication, an organization would simply be a collection of individuals, resources, and finances (Lunenburg, 1991). Silver (1987) also emphasizes the significance of communication and states that it is essential for an effective manager in any organization. It is a managerial skill that is a necessary foundation for effective leadership because it encompasses all management functions. No manager can do well without planning, organizing, and controlling.

### **2.7.3. Creating Awareness**

Communication involves a two-way exchange between a sender and receiver, with feedback from the receiver being an essential component. (Rimal, Why health communication is important in public health, 2009) Emphasize the significance of communication, describing it as central to our human nature, involving the exchange of symbolic information that defines us. Various strategies, such as books, photographs, songs, movies, posters, flyers, and seminars, can be employed to package and disseminate information with the aim of creating awareness, informing, educating, and promoting behavioral change (Rimer, 2006).

This study focuses on examining communication strategies, channels, and messaging techniques that been use to prevent hate speech. It highlights the importance of evidence based communication practices in raising awareness, educating individuals about the risks and consequences of hate speech

The effectiveness of communication interventions depends on how recipients engage with the issue being addressed, considering its relevance, importance, and intention to take action. (Rimal, 2009) Stress that communication interventions involve individuals receiving and processing information, influencing what they are exposed to (selective exposure) and the meaning they derive from communication (selective perception), which is influenced by various factors. Communication interventions have a significant impact on behavior, shaping opinions, beliefs, and knowledge on specific topics. While the communication field is dynamic and subject to criticism, its overall influence has increased over the years, especially with technological advancements.

## **2.8. Challenges and Barriers**

Additionally, tackling misinformation and hate speech is made more difficult by the influence of prominent individuals in Ethiopian society, who frequently use their platforms to share divisive messages. Recent research shows that political figures and religious leaders often spread negative language that heightens social tensions. This situation highlights the importance of focused efforts directed at these influential personalities, urging them to practice more responsible forms of communication. By involving these leaders in workshops and awareness initiatives, it could be possible to guide public conversations towards more positive discussions that emphasize unity and empathy instead of separation. In the end, dealing with this complex issue demands both technological approaches and cultural changes that encourage responsibility among those who influence public views.

There are several factors that can impede effective communication; including various factors such as filtering, selective perception, information overload, emotional disconnection, reliance on unfamiliar or untrustworthy sources, workplace gossip, semantic barriers, gender differences, differences in interpretation between the sender and receiver, and biased language can all impact communication. (Alessandra, 1993). These barriers largely depend on the organization's structure, with complex structures involving multiple layers of management often causing communication breakdowns or distortions. Information can get lost as it moves up the hierarchy, and formal structures can cause delays and rigidity. Poor instructions and high work pressure can also hinder communication.

Overcoming communication barriers requires strategies such as improving language skills, promoting cultural understanding, and using clear and concise language, active listening, providing feedback, utilizing appropriate technology, and fostering a supportive and inclusive communication environment. However, communication is a complex task that is susceptible to interference and distortion, with only 50% of communicated information being retained by subordinates and 60% by managers (Koontz & O'Donnell-Trujillo, 1983). Oral communication can result in a loss of 30% of information in each communication procedure. As messages travel, they may be transformed, resulting in the recipient receiving something different from what was originally sent. Therefore, addressing communication barriers is crucial for improving communication effectiveness. Therefore, Organizational communication is not perfect due to

numerous barriers. Throughout the procedure many factors obstruct a message's free flow to its destination when it is transmitted. The message may not reach its intended destination. Any factor that impedes the exchange of information between a Communications is hampered by the presence of a sender and a receiver. Barriers are elements that interfere with the receiver's reception. Comprehension of the intended meaning of the sender's message.

## **Theoretical Frameworks**

### **I. Agenda-Setting Theory:**

In the initial Chapel Hill research, as well as in several studies that came after, there were objects and public matters that made up both the media's agenda and the public's agenda. But Weaver et al. (1981) introduced a new agenda to the literature, the agenda of personal concerns, on which politics is but a single entry. This agenda consists of various items. Communication involves a process. It can relate to any group of items or just one item trying to gain attention. The news media establish a list of characteristics for each competitor in how they show the campaign. Voters develop a perception of each candidate, which includes a list of characteristics, through exposure to news media and information gathered from family and friends (L.Shaw, 1993).

The Internet dramatically changed the communication landscape with the introduction of myriad new channels. Email, internet newspapers, chat rooms, and websites that cover various beliefs, businesses, and personal interests have transformed how millions communicate globally and have created extensive new areas for communication researchers to explore. In modern society, numerous agendas exist, and a significant portion of them is now accessible to a wide audience. As a result, some experts suggest that agenda setting may decline since audiences are becoming more divided. Almost everyone now has their own specific media agenda, created from the abundance of news and information available online. The outcome of these unique personal goals, as the argument goes on, will lead to a public agenda that shows great variety and a spread of public interest (McCombs, 2005).

The agenda-setting theory looks at how the public decides which issues are significant at a given time. It suggests that while the media does not directly tell individuals what to think about an issue, it does shape their views on what matters and should be discussed. When a person is frequently exposed to news coverage about a particular topic, they are likely to regard that topic

as important. Thus, the impact does not flow in the anticipated way from the person to the media. It is not the individual that sets the media agenda, it is the media coverage of an issue that sets the public agenda. This situation becomes more troubling when we consider the standards that media outlets apply to decide if a story is important enough to report (Stephen W. Littlejohn Karen A. Foss, 2009).

The power of agenda setting that McCombs and Shaw describe may be on the wane. In an innovative study, researchers Scott Althaus and David Tewksbury from the University of Illinois suggested that old print media might be better than modern electronic media at influencing what readers think is important. They reasoned that people who are reading a newspaper know that editors consider a long, front-page article under a banner headline more important than a short story buried on an inside page. These comparative signals are missing from the computer display, and users online can click on links to related articles without encountering the events that readers of print media find while flipping through the pages. . (GRIFFIN, 2012).

The Agenda-Setting Theory is useful for understanding how the Ethiopian Media Authority can influence discussions about hate speech and the importance of fact-checking among the public. In the past, media has shaped public opinion by highlighting certain topics, thereby making them more prominent this is where Ethiopian Media Authority has the opportunity to elevate awareness about hate speech and fact-checking to lead news stories. However, in today's diverse digital landscape, where individuals access a variety of online content, the power of agenda-setting may be diminishing. To address this, Ethiopian Media Authority needs to effectively promote messages against hate speech and the significance of fact-checking across different platforms such as mainstream media, social networks, and community conversations to highlight their relevance. By continuously presenting these issues as key public concerns (descriptive norms) and viewing hate speech as socially unacceptable (injunctive norms), Ethiopian Media Authority can shape the priorities of their audience. Therefore, a communication strategy based on these norms that spans multiple platforms can enable Ethiopian Media Authority to positively influence public discussions, foster responsible media usage, and combat misinformation.

## II. Elaboration Likelihood Model

The elaboration likelihood model, a framework developed by psychologists Richard Petty and John Cacioppo in 1981. This model of cognitive processes gets its name from the chance that an individual reflects thoroughly (elaborates) on an advertisement when they see it. Its main focus is on altering attitudes by using persuasive techniques. The basic premise of this model is that the route by which a message persuades consumers depends on their involvement with the message. Two routes exist: the *central route* and the *peripheral route*. In the first situation, individuals are both eager and capable of judging the message, so they will carefully consider it. This means they will seek out and react to strong points that support the message while challenging what they see as weak points. In the second situation, individuals might not feel motivated or may not have the ability to assess the message, leading them to pay more attention to aspects related to the message, like how entertaining it is or if a famous person is promoting it, instead of focusing on logical arguments. In short, high involvement leads to central processing resembling traditional hierarchy models, whereas low involvement leads to peripheral processing. Important to note is that attitude change in terms of central processing is the result of thoughtful reflection on information. As such, change does not occur just as an outward compliance (behavior change) but indicates a change in beliefs (personal acceptance or a shift in values or opinion). (O'keefe, 2008)

In elaboration likelihood model, individuals are seen as either motivated to process messages or not; also, they are seen as having different abilities to process messages. When motivation and ability are high, individuals will use a central route for processing messages, giving carefully considered cognitive effort to the message. Under this condition elaboration is high; the information is considered cognitively and rationally. Attitude change and learning under such conditions are predicted to be enduring. When motivation and/or ability are not present, individuals will process the information heuristically using other non-rational cues. Under such conditions, elaboration is low, and any attitude change or learning from such persuasion is considered to be less enduring. A similar dual process model of persuasion and learning is known as the heuristic-systematic model of persuasion.

The Elaboration Likelihood Model serves as a theoretical basis for exploring effective ways to communicate about hate speech and fact-checking, especially within the Ethiopian Media

Authority context. Given the diverse and frequently divided media landscape in Ethiopia, where both ethnic conflict and false information are widespread, the Ethiopian Media Authority needs to craft its messages carefully to cater to different audience motivations and their capacity to grasp information. For audiences that are highly motivated and engaged, such as journalists, activists, and well-educated city dwellers, using the central route for persuasion is likely to be the most beneficial. This approach involves utilizing strong arguments, credible evidence, and detailed fact-checking reports to bring about lasting changes in attitudes. On the other hand, for audiences that may be less aware of media issues or show less interest, like rural residents or those influenced by biased narratives, the peripheral route could prove to be more advantageous. This method would involve using straightforward information, well-known local figures, and emotionally appealing stories to inspire initial changes in behavior. Additionally, Elaboration Likelihood Model highlights the importance of dividing audiences into segments when creating convincing campaigns against hate speech and promoting fact-checking, enabling the Ethiopian Media Authority to enhance both immediate compliance and enduring changes in beliefs, ultimately fostering media integrity.

### **III. Social norm theory**

Much of human behavior is influenced by social norms. Although communication represents one important mechanism through which norms are formed and disseminated. The integration of communication with referent others and observation of referents as norm-building processes in normative social influences allows us to consider developments regarding the media environment that are becoming more and more relevant for normative social influences. Due to technical convergence, interpersonal communication not only takes place face to face but also digitally and via mobile devices (CIALDINI, 1998). From a comprehensive communication perspective, not only communication and observation among group members but also media should contribute to perceptions about the prevalence of and attitude towards a behavior by presenting individuals who are perceived as similar to oneself or members of the reference group. From a social norms perspective, the individual's assumption that media influence social referents' behaviors should affect the formation of descriptive norms, whereas the expectation that referents' attitudes are affected by media should affect perceptions of social approval and thus injunctive normative perceptions (Hefner, 2019).

Social norms play a crucial role in how humans behave. Previous studies describe norms as declarations about what should or should not occur, as expectations linked to certain roles, or as evident when actions lead to negative consequences. Generally, norms represent what is deemed appropriate behavior within a society and its smaller groups, thus influencing individual actions. For norms to be considered social, they must be accepted by others and rely partially on their approval or disapproval. The perspective of public good does not necessarily mean that social norms are advantageous for everyone involved. In reality, many social norms can leave specific groups out of public benefits because they favor one subgroup, meaning they perform roles of both inclusion and exclusion. Simple punishments uphold social norms, which can evoke feelings of guilt and shame when individuals have internalized these social expectations (Katja Rost, 2016).

The Social Norm Theory highlights the way media can influence behavior through descriptive norms (what is commonly done) and injunctive norms (what is considered acceptable). For the Ethiopian Media Authority, this implies that it should implement communication methods to normalize fact-checking (descriptive norm) and to denounce hate speech (injunctive norm). Influencers and community leaders can help strengthen these norms and spread them to foster responsible dialogue. By aligning media messages with this theory, the Ethiopian Media Authority can encourage a culture of honesty and unity in Ethiopia's emerging digital environment.

## **CHAPTER THREE**

### **Methodology of the study**

#### **3.1. Introduction**

The study on effective communication strategies for raising awareness of hate speech and fact-checking within the context of the Ethiopia Media Authority is crucial given the increasing prevalence of misinformation and hate speech in digital and traditional media. The methodology employed in this study is designed to systematically explore, analyze, and evaluate various communication strategies that can effectively address these issues.

This chapter presents study design, method of the study, sample and sampling techniques, data collection tools, data collection procedures and method of data analysis.

The qualitative research approach was used to answer the research questions and achieve the study objectives. This involves conducting interviews and focus group discussions with key stakeholders, including media professionals, civil society. The aim is to gather insights into their perceptions of hate speech, its impact on society, and the effectiveness of current communication strategies.

#### **3.2. Reason for selection of the organization**

As a study conducted about the Ethiopian Media Authority, it has focused on past year to raise awareness on hate speech and fact-checking across the highly polarized media in Ethiopia.

The Ethiopian Media Authority was selected as the subject of this case study because it has a pivotal position in the regulation and guidance of media practices in Ethiopia, a country with a highly diverse and at times polarized media landscape. The Ethiopian Media Authority is primarily the regulating body and has a privileged position in defining media standards, promoting responsible journalism, and combating harmful content such as hate speech and disinformation. Ethiopia has experienced great challenges that are related to ethnic conflict, political instability, and spread of disinformation, therefore the effort by the Ethiopian Media Authority in combating disinformation and hate speech is timely critical. In addition, the Ethiopian Media Authority's collaboration with media stakeholders, civil society, and technology

platforms provides valuable lessons on multi-sectoral responses to countering harmful speech without compromising freedom of expression.

### **3.3. Research Methods**

Qualitative research encompasses non-quantitative analysis. It involves collecting, examining, and interpreting information derived from people's actions and words, which are referred to as qualitative research. This type of study looks into the meanings, definitions, traits, symbols, metaphors, and descriptions related to situations to gain a deeper understanding of them.

The qualitative method portrays a problem or situation through the eyes of those who are living it and often relies on personal evaluations of the human aspects of an issue to explore behaviors, beliefs, opinions, and feelings about specific topics (Kothary, 2004).

Qualitative research focuses on how data sources give meaning to events around them. A qualitative approach facilitates an understanding of reality from the perspective of the meanings. This study employed the qualitative method of study and analysis to enhance the study's trustworthiness (Babbie, 2010).

### **3.4. Qualitative analysis**

In contrast to the quantitative technique, a qualitative research method is centered on uncovering the reasons and mechanisms behind the phenomena being studied. This approach is deemed more significant because it allows the researcher to gain a deeper understanding of the topic, yielding better, trustworthy, authentic, and in-depth insights. In this research, the qualitative approach was utilized. It serves as a suitable method for exploring questions of what, who, and why (Patton and Cochran, 2002). Qualitative research is better suited for exploring words and concepts instead of merely tallying numbers, and a qualitative interview allows both the interviewer and the participants to engage in detailed discussions on various subjects.

This study employs a qualitative methodological approach to conduct a context evaluation of the opportunities and challenges in institutionalizing hate speech awareness and fact-checking in the Ethiopian Media Authority. Through in-depth interviews, focus group discussion and document analysis.

### **3.5. Population and sampling Techniques**

In social scientific research, a population is the cluster of people, events, things, or other phenomena that you are most interested in; it is often the "who" or "what" that you want to be able to say something about at the end of your study.

The target population is the specific, conceptually bounded group of potential participants to whom the researcher may have access that represents the nature of the population of interest (Shukla, 2020).

Sampling refers to the method of choosing specific observations to be examined for research. A sample is a segment of a population that fully represents it and is called a sample. This indicates that the selected units from the population, serving as a sample, should embody all varieties of traits of different units within the population (Shukla, 2020).

Since Ethiopian Media Authority is a federal entity based in Addis Ababa, most of the interviewees were based on the training given. In this study, the researcher employs purposive sampling and snowball sampling techniques to select key informants for the interview section.

A researcher used snowball sampling, also known as chain-referral sampling, to identify key informants from the Ethiopian mass media professionals' association.

Purposive sampling is a method that does not rely on probability, and it takes place when "the researcher selects the elements for the sample based on their judgment." Because gathering data is essential in research, the information gathered aims to enhance comprehension of a theory. Thus, it is important to carefully consider how to collect data and from whom to obtain it, particularly since flawed data collection cannot be compensated for by any level of analysis (Ilker Etikan, 2016).

These sampling methods were chosen to ensure diverse perspectives and gather comprehensive information about effective communication strategies for raising awareness hate speech and fact checking. In total, 16 key respondents participated in the study, with 9 individuals selected through purposive sampling for in-depth interviews and 7 participants selected through snowball sampling for focus group discussions.

The researcher conducted individual in-depth interviews with four Ethiopian Media Authority officials and media experts, three public sector experts, and two secondary school students to gather diverse perspectives on awareness of hate speech and fact checking in Ethiopia. The selection of individuals was based on purposive sampling.

A purposive sampling strategy is utilized to select participants who have relevant experience or knowledge about hate speech and media practices in Ethiopia. This may include journalists, media professionals, educators, community leaders, and civil society. The aim is to gather diverse insights that reflect various viewpoints within the Ethiopian context.

In this particular study, participants were chosen based on predefined inclusion criteria to ensure their relevance and suitability for providing valuable insights into the awareness of hate speech and fact checking.

### **3.6. Data Collection Tools**

Qualitative data collection methods typically involve direct interaction with individuals either individually or in group settings. These approaches provide the advantages of obtaining rich and in-depth data, offering deeper insights into the phenomenon being studied. In line with this, the current study incorporated both primary and secondary sources to gather relevant data. Primary data was generated through in-depth interviews and focus group discussions. As data collection tools, the researcher applied focus group discussion, in-depth Interviews, content analysis

#### **3.6.1. Document Analysis**

In this study, media proclamations and National Report on Hate Speech and Disinformation on Social Media and documents of the Ethiopian Media Authority regarding misinformation and other useful documentation of the organization were discussed and reviewed

Document analysis is a methodical approach for assessing various documents, including those that are either printed or electronic, such as materials from the computer or internet. Similar to other qualitative research techniques, document analysis involves examining and interpreting data to extract meaning, deepen understanding, and build empirical knowledge. Documents include words and images that have been captured without involvement from a researcher. This discussion will not consider other forms of silent or indirect evidence, like cultural items.

Documents that may be used for systematic evaluation as part of a study take a variety of forms. They consist of ads; schedules, lists of participants, and meeting notes; guides; reference documents; publications and leaflets; planners and logs; event schedules (such as printed guides); correspondence and notes; diagrams and graphs; news articles (clippings/stories); media statements; project suggestions, application documents, and overviews; scripts for radio and television shows; reports from organizations or institutions; data from surveys; and multiple public documents. Scrapbooks and photo albums can also furnish documentary material for research purposes. These types of documents are found in libraries, newspaper archives, historical society offices, and organizational or institutional files. (Bowen, 2009).

### **3.6.2. Focus Group Discussion**

A focus group discussion brings together individuals who share similar backgrounds or experiences to talk about a particular topic. This type of research is qualitative and involves asking questions about their opinions, attitudes, beliefs, and ideas. In focus groups, participants are encouraged to engage in conversation with one another; this sets it apart from other research methods. Typically, it consists of a group interview with a small number of participants, often between 8 and 12. A moderator (interviewer) guides the conversation in a relaxed manner, covering different topics of interest.

The makeup of the group and the way the discussion is led need to be thought out to ensure a welcoming atmosphere where everyone can share their thoughts freely and honestly. Participants are motivated to not just share their views but also to engage with others and answer questions from the facilitator. Because of this, focus groups provide a richness, subtlety, and diversity in conversation that surveys cannot achieve (Suaahara, 2016).

### **3.6.3. In-depth Interview**

The in-depth interview method used in qualitative research stands out from other interview types in two main ways: the objectives of the interview and the approach taken while interviewing. A journalist aims to collect information for a news article, a psychologist seeks to reduce a person's emotional pain, and a police detective interviews individuals to secure a confession. In contrast, a qualitative researcher carries out in-depth interviews to gather detailed insights from a small set of individuals within a specific group, using a carefully planned research framework designed to

ensure reliable and analyzable results. The main purpose of research interviews is to facilitate improvements in the lives of the selected group and others in similar situations. Because of these differing goals, it's not surprising that qualitative researchers adopt interviewing techniques that differ notably from those used by journalists and detectives.

In qualitative research, the way interviews are conducted is not meant to be aggressive or confrontational. The goal is not to cause conflict to challenge the person being interviewed. Instead, the focus is on creating a trusting connection where all feedback is respected, and honest sharing can flourish. This happens because it is made clear that the information will be kept private unless the interviewee agrees to share it. The relationship between the interviewer and the interviewee is fundamental to the in-depth interview process, which is why this method is considered one of the most intimate forms of qualitative research design (Roller, 2012-16).

Additionally, various secondary data sources were utilized to enhance the study by incorporating different perspectives and viewpoints. These materials encompassed books, research papers, articles, journals, electronic resources, and other relevant documents. The researcher conducted a comprehensive review of these secondary sources to gain a deeper understanding of different authors' perspectives on the research problem. The aim of this review was to establish connections between the current study and the existing literature on the subject matter. By analyzing and referencing pertinent secondary sources, the researcher sought to contextualize the study within the broader body of knowledge on the topic.

### **3.7. Data Analysis and Interpretation**

All data gathered from the study units using various qualitative collection methods have been transcribed, with most of it recorded on audio tapes. Following this, the important data have been organized into categories for easier analysis; this organization was based on their connection to the main themes of the research questions posed in this study. According to Wimmer and Dominick (2003), qualitative researchers adopt an inductive approach: data pertinent to the research topic is collected and sorted into fitting and significant categories, allowing explanations to arise naturally from the data.

To extract insights from the qualitative information obtained through detailed interviews with key informants, as well as from secondary data sourced from relevant documents, the researcher

employed thematic analysis and content analysis for each type, respectively. According to Lacey et.al, (2009) analysis of qualitative data usually goes through some or all of the following stages (though the order may vary):

**Familiarization:** with the data through review, reading, listening etc.: it is an essential stage in which the researcher will listen to tapes and watching video material, reading and re-reading the data, making memos and summaries before the formal analysis begins. **Transcription of tape recorded material:** as the data might be gathered in either of tape recorded interviews, focus groups, video recordings, or handwritten field notes it has to be transcribed orderly.

**Organization and indexing of data for easy retrieval and identification:** it is a process of organizing data into easily retrievable sections. In each interview a number or code, or to break up field notes into sections identified by date, or by context and some of the Interviewees might need to be given pseudonyms with confidentiality or referred to by a code number. **Coding (may be called indexing)** It is a method of arranging data based on its subject matter, which allows the researcher to begin the process of carefully evaluating it, determining what the data is saying us as well as the links and patterns in the study data.

Finally, it is a systematic presentation of the data to give a report that can be easily utilized to conclude the analytical process. The researcher's technique is supposed to be stated in the final report.

### **3.8. Ethical Considerations**

It's important to continuously evaluate communication strategies' effectiveness in raising awareness about hate speech and fact-checking initiatives. Gathering feedback from community members can provide insights into what works well and what needs improvement.

In line with ethical principles of informed consent, all participants were provided with comprehensive information regarding the research objectives, procedures, and potential risks, and their written consent was obtained before their involvement in the study. Moreover, measures were in place to safeguard the confidentiality and anonymity of participants, with assurances provided that their identities and responses would remain strictly confidential. To mitigate the risk of misrepresentation and ensure the accuracy of findings, validation sessions

were conducted with participants, allowing them the opportunity to review and confirm the accuracy of their contributions before final reporting.

This research underscores the foundational importance of ethical principles in guiding the entire research process. Upholding the rights and welfare of participants is not just a procedural requirement but a fundamental ethical imperative. The meticulous steps taken to obtain clearance from regulatory authorities and secure informed consent from participants reflect a commitment to transparency and respect for autonomy. Additionally, the emphasis on confidentiality and anonymity serves to protect participants from potential harm or repercussions stemming from their involvement in the study. The validation sessions further exemplify a participatory approach, empowering participants to validate the accuracy of their contributions and ensuring their voices are accurately represented. Moreover, the stringent adherence to data protection legislation underscores a commitment to safeguarding sensitive information and preserving the trust and confidentiality bestowed upon researchers. Ethical considerations extend beyond mere compliance with regulations; they embody a moral responsibility to conduct research with integrity, respect, and accountability. By prioritizing ethical conduct, this research not only upholds the integrity of its findings but also fosters a culture of trust and credibility in the broader research community

Participants will be fully informed about the purpose of the study, their role, and their right to withdraw at any time. Written or verbal consent will be obtained. All participant data will be kept confidential, and their identities will be anonymized in the reporting of findings. Given the potentially sensitive nature of discussions around hate speech, interviews will be conducted with care to ensure participants feel safe and comfortable sharing their experiences.

## **CHAPTER FOUR**

### **Data presentation, Analysis and discussion**

#### **4.1. Introduction**

This study evaluates the effectiveness of the Ethiopian Media Authority's community training programs in reducing the harm caused by hate speech and misinformation, with a focus on fact-checking mechanisms. The primary objective is to assess how these training initiatives influence community awareness and behavior regarding harmful content, examining the perspectives of key stakeholders. The research aligns its findings with theoretical frameworks from the literature review, addressing specific research questions and objectives to understand the Ethiopian Media Authority's role in mitigating these issues.

Data were collected using qualitative methods, including in-depth interviews, focus group discussions, and secondary document reviews, targeting purposefully selected Ethiopian Media Authority officials, media professionals, civil society members, and practitioners who participated in the training. All interviews were recorded and transcribed to ensure accuracy during analysis. The researcher thoroughly reviewed the transcriptions and recordings, creating detailed notes and write-ups to organize the data in relation to the study's goals and research questions, ensuring a comprehensive understanding of the stakeholders' insights on hate speech and fact-checking processes.

The analysis involved identifying and coding common themes, patterns, and expressions across the data sources. Data from interviews and documents were synthesized when they shared similarities and compared when differences emerged, allowing for a nuanced understanding of the Ethiopian Media Authority's efforts. This approach enabled the study to evaluate the impact of the training programs and their alignment with theoretical perspectives, providing insights into the effectiveness of the Ethiopian Media Authority's strategies in addressing hate speech and misinformation.

## 4.2. Document analysis in relation to hate speech and fact checking

This researcher has obtained from the Ethiopian Media Authority national report on social media hate speech and dissemination of misinformation in Ethiopia

Ethiopian Media Authority's legal action, as a legitimate federal institution, is grounded in the established legal and policy frameworks. In this sense, the proclamations pertaining to media, advertisements, hate speech and misinformation prevention and suppression, and other instructions are all subject to the power to actively address the media and media practices. There is no directive available for information searchers who wish to learn about the law and regulation regarding misinformation and disinformation, or any associated benefits, even though the proclamation has been published and made available online on its website. Similarly, the available documents are published only in Amharic.

Proclamation No.1185/2020, which aims to prevent and suppress hate speech and disinformation, appears to have good intentions based on its stated goals: "protecting free expression while suppressing all forms of hatred and discrimination; promoting tolerance, civil discourse, and dialogue; fostering mutual respect; strengthening democratic governance; and controlling and suppressing the dissemination and proliferation of hate speech, disinformation, and other related false and misleading information."

*““The Ethiopia Media Authority has the duty and responsibility to monitor and report hate speech and false information spread by the mass media in Ethiopia. The spread of hate and false information has been highlighted since 2021. when the conflict between the Tigray People’s Liberation Front and the Federal government broke out. According to the United Nation report, digital communication networks are used to exacerbate conflicts and affect other part of the society in addition to the target group”” (Authority, 2024).*

Currently, with the advent of new media (social media), the spread of hate speech and false information is worsening the problem more than ever. Hate speech, messages that inflame conflict, and false information have become a source of concern globally and in Ethiopia, due to their unique characteristics. The Internet has fostered an environment conducive to hate speech due to its low-cost distribution, widespread accessibility, and anonymity, posing significant challenges for Ethiopia. The minimal expense of posting on social media or other platforms

allows malicious actors to spread harmful content widely, while global accessibility amplifies its reach across diverse communities. Anonymity shields perpetrators, enabling them to disseminate divisive messages from anywhere without accountability, straining the Ethiopian Media Authority's efforts to monitor and mitigate such content effectively, as highlighted by their training initiatives to combat misinformation and hate speech. The Ethiopian Media Authority is tasked with keeping an eye on and reporting hate speech and misinformation propagated by Ethiopian media outlets. Its purpose is to evaluate the actions taken by social media to stop hate speech and misinformation from spreading and to offer solutions.

The Ethiopian Media Authority has found it necessary to make the youth feel responsible in terms of information use, as the youth constitute the largest number of social media users in our country. The involvement of religious media owners and editors has its own contribution to preventing the problem. The participation of religious leaders in terms of their influence on education makes the training effective in terms of reaching a large audience. The training of Ethiopian Media Authority leaders and employees creates awareness that hate speech and false information can be prevented quickly and without causing serious harm in the media monitoring work they carry out.

The data, originally collected in Amharic, was first transcribed into English. The researcher then conducted a data-cleaning process to prepare the transcripts for the coding phase

### **4.3. Response Rate of Respondent**

This section outlines the results of the thematic analysis, focusing on the characteristics and response rates of the study's participants, including youths and experts from the Ethiopian Media Authority. The age range of participants varied, with youth informants aged 16–35 years and other informants, primarily experts, aged 28–48 years. Participants came from diverse educational backgrounds, with some being students, offering insights into younger perspectives, while others held qualifications above a diploma, providing academically advanced viewpoints.

The study involved 16 informants, with data collected through in-depth interviews with nine individuals and one focus group discussion with seven participants. This combination of qualitative methods allowed for a comprehensive exploration of perspectives on hate speech and

fact-checking initiatives by the Ethiopian Media Authority. The varied age groups and educational levels of the informants enriched the analysis, ensuring a broad range of insights aligned with the study's objectives.

#### **4.4. Profile and status of Interview and Focus Group Discussion Informants**

The researcher has tried to ask the interviewee of experts and status of the informants of focus group discussion who participates in the awareness of hate speech and fact checking.

The purpose of the study was to gather in-depth insights and expert opinions on a specific topic, which necessitated direct engagement with knowledgeable individuals in the field. To achieve this, the researcher conducted structured or semi-structured interviews with nine carefully selected experts, ensuring a diverse range of perspectives and expertise.

The researcher conducted in-depth interviews with experienced trainers specializing in these fields. By engaging directly with the trainers and the responses from the interviewees were transcribed and encoded in an appropriate manner. Along with the interview, the researchers has analyzed documents which the obtained from Ethiopian Media Authority and related media establishment documents and research in the field.

**Table 1: Interview of informants coding**

No	Position	Sex	Level of education	Code	Remark
1	Capacity Building Desk Manager	M	MA	E-1	
2	Communication Director	M	MA	E-2	
3	Senior Capacity Building Expert	M	BA	E-3	
4	Senior Expert	M	MA	E-4	
5	Teacher	F	BA	E-5	
6	Student	F	Grade 9 student	E-6	
7	Student	M	Grade 10 student	E-7	
8	Expert	M	BA	E-8	

9	Senior Expert at public sector	F	Diploma	E-9	
---	--------------------------------	---	---------	-----	--

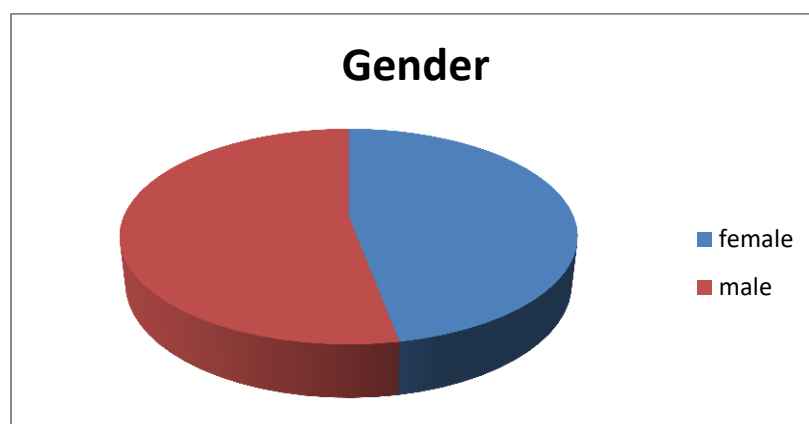
Source: - interviewee

**Table 2: Informants of focus group discussion**

Participants of Focus Group Discussion 1		
No	Code	Education Level /status
1	T-1	BA Degree
2	T-2	Grade 10 student
3	T-3	Grade 9 student
4	T-4	Grade 10 student
5	T-5	BA Degree
6	T-6	Grade 10 student
7	T-7	Grade 9 student

Source:-Informants of focus group discussion

**Figure 1: Gender of the informants**



#### **4.5. Effective communication strategies employed by Ethiopia Media Authority for raising awareness about hate speech and fact checking**

The Ethiopian media authority establishment proclamation No. 1238/2021 gives the authority a responsibility to ensure the expansion of media services that are accessible and of high quality,

which enables the creation of an informed society, which is believed to be a key to accelerating sustainable and holistic political, social, and economic development. According to the proclamation, it is crucial to establish a legal framework and system that ensure the media operate responsibly by respecting public peace, security, and competing rights and interests as per the constitution.

The Ethiopian Media Authority conducted its annual assessment by collecting and processing information from across multiple social media platforms to analyze hate speech diffusion trends and success of counter-actions. Employing a combination of automated watchdog technology, user-complaint reports.

Ethiopian Media Authority adopts a multi-channel strategy to fight hate speech and misinformation using digital platforms to maximize reach and impact. Noting the rapid spread of hazardous content on social media, the Ethiopian Media Authority targets sites like Facebook, Tiktok and YouTube to post fact-checked materials, respond to misinformation stories, and engage citizens through interactive campaigns, hashtags, and viral study materials. Furthermore, town halls and outreach programs at the community level facilitate grassroots dialogue about identifying and countering hate speech.

The growing awareness of mis- and disinformation contributed to the rise and spread of fact-checking websites run by journalists or/and activists debunking fake news and political manipulation and providing the citizens with non-partisan, objective and verified facts (Kumar.A, 2022). The content presented by the fact-checkers should come from open, generally available sources, allowing the readers to reach the same data and conclude on their own. Today, fact-checking is considered a tool to verify claims from all types of social actors: politicians, journalists, activists as well as social media users (Humprecht.E, 2020).

The primary way of debunking visual fake news is the use of reverse image search, which helps find the place an image appeared for the first time and show the history of how it was used for mis- and disinformation purposes. The fact-checkers, however, rarely limit themselves to using this technical method only. The majority of fact-checks in which reverse image search was used also contain other sources of information providing the context and facts on what the manipulated photograph really refers to. Sometimes, when the photograph was edited or remade,

it is crucial to provide an exact comparison between the original and fake picture. In a case like this, a useful tool is a slider, thanks to which the reader may immediately see what changes were incorporated and where they can be seen (Szafrńska, 2024).

The Ethiopian Media Authority employs the multi-channel technique to raise awareness on hate speech and fact-checking, applying digital channels in order to realize wider reach and engagement. Different types of fact-checking tools have been developed to help users identify and evaluate the accuracy of news and information that they encounter online. Social media platforms like Facebook, Tiktok, and YouTube are critical for real-time fact-checking, with the Ethiopian Media Authority refuting falsehoods using advanced verification processes like reverse image searches (Google Lens, Tin Eye, Yandex), video verification tools (InVID), and cross-validation against credible databases (Google Fact Check Explorer, FactCheck.org, PolitiFact, Snopes). The authority also collaborates with news sites and blogs to publish investigative reports and formatted fact-checks through platforms like the Fact Check Markup Tool and Claim Review to achieve higher search engine ranking. While being inclusive. Workshops, media training courses, and partnerships with civil society organizations boost fact-checking skills. By integrating these platforms with the newest digital technologies, the Ethiopian Media Authority receives a dynamic, multi-pronged approach to countering disinformation and building a culture of media literacy in Ethiopia.

According to interview E-2

The "9192" initiative by the Ethiopian Media Authority is a communication strategy designed to combat hate speech and disinformation through a dedicated hotline call center, with community engagement as a central component. This policy gives the citizens an active role to play in content monitoring and reporting injurious content by developing a collective effort between the public and regulating bodies”.

Social media platforms, while often conduits for the rapid spread of hate speech and disinformation due to their accessibility, algorithmic amplification, and anonymity, are strategically utilized by the Ethiopian Media Authority as tools to counteract these issues. The Ethiopian Media Authority leverages social media’s vast reach and immediacy to disseminate accurate information, promote media literacy, and engage communities in countering harmful

content. As social media consumers in Ethiopia are so highly engaged frequently sharing, commenting, and reacting to posts. The utilization of sites like Facebook and TikTok enables the Ethiopian Media Authority to reach viewers where they are most engaged, with greater breadth and impact.

“Public engagement is important in ensuring a responsible and honest media culture. In order to fight hate speech and disinformation, citizens must actively report such content to 9192, the Ethiopian Media Authority's special hotline” said T-3.

To maximize engagement, the Ethiopian Media Authority creates content that will engage diverse audiences, in the form of short videos and infographics dissecting complex issues like media literacy and hate speech risks.

While social media remains a primary focus due to its vast reach. On other hand it is also actively developing strategies to extend these efforts to broadcast media including television, radio, and other traditional platforms interview E-2 added.

#### **4.6. The role of Ethiopian media authority in shaping conversation about hate speech and misinformation**

It seems important to underline that today’s digital platforms characteristics facilitate the creation and dissemination of hate speech. Since they allow for rapid, effective, permanent and inexpensive dissemination of thought, they have built an open road for the publication of any kind of message, without the presence of any structure (formal or informal) capable of exercising a mediating or controlling function. Indeed, social media permits the message to be extended to a wide audience, but they also allow the acquisition and maintenance over time of that message, which remains tracked and therefore retrievable, possibly leading to a continuous harm to the victims.

For sake of completeness, it should be remarked that technology can not only contribute to hate speech spread, but it can also be useful to fight and reduce it. In this vein, IT research has focused on creating database of hate speech to recognize it through algorithms that assimilate comments on the web to those in the database, with the aim of being able to promptly eliminate such comments .With such tools, social networking sites are able to autonomously identify the

use of inappropriate words or content in the area of hate speech. However, it seems very difficult to create a complete and constantly updated database, and for these reasons part of this work is delegated directly to social media users (Benvenuto, 2023).

In this research, the key informants alarming nature and seriousness of hate speech, viewing it as a widespread and dangerous problem that threatens our social fabric and national stability. Many voiced their deep worries about how hate speech drives division, incites violence, and erodes democratic values, especially in societies that are already fragile or polarized. They pointed out that hate speech often targets vulnerable groups like ethnic minorities, religious communities, and political dissidents intensifying discrimination and conflict. Some informants noted how quickly hate speech spreads on digital platforms, where misinformation and inflammatory language only make its harmful effects worse. They also emphasized that hate speech isn't just an online issue; it has real-world consequences, including mob violence, systemic oppression, and even genocide in the most extreme cases. Given these dangers, the informants called for stronger legal measures, media literacy programs, and proactive counter-speech initiatives to help lessen its impact.

"Ethiopian Media Authority needs to hold media and social media users who spread hate speech to account, but they need to be fair and not silence critics" said T-5.

All the informants have pointed out that hate speech has turned into a major problem, spreading quickly not just in our country but around the world, thanks to social media and digital platforms. As technology keeps evolving, the damage it causes only gets worse, allowing false and exaggerated claims to masquerade as the truth while specifically targeting certain communities.

"Instead of just blocking content, Ethiopian Media Authority should teach us how to spot fake news through school programs and social media campaigns" said T-7.

To combat the growing threat of hate speech and disinformation, the Ethiopian Media Authority has actively engaged by providing specialized public awareness campaigns to media practitioners, content producers, and journalists, secondary school students involving in mini media since 2024 in different regions like Dire Dawa, Jimma so on. The public awareness campaigns on detecting, examining, and countering hate speech and promoting ethical reporting and fact-checking techniques. Through equipping media practitioners with fundamental skills in

digital literacy and responsible journalism, the Ethiopian Media Authority aims to reduce the spread of harmful content that incites violence or discrimination. The program also emphasizes verifying information before dissemination, fostering a culture of accuracy and accountability in media reporting.

As the regulator, the Ethiopian Media Authority regulates broadcast, print, and online media to determine compliance with anti-hate speech and disinformation laws, such as Ethiopia's Hate Speech and Disinformation Prevention Proclamation.

Most of the informants shared hate speech as any form of communication verbal, written, or visual whose purpose is to attack, degrade, or invite harm on someone or group because of markers such as ethnicity, religion, gender, or politics. They underlined its role in disseminating misinformation, enforcing stereotypes, and fostering estrangement, whose practical consequences are violence or discrimination. A few noted that hate speech thrives on emotional manipulation, drawing upon social tensions and masquerading as legitimate opinion or fact-based speech.

“If you are asking me about the issue in Ethiopia, Hate speech encompasses any form of expression whether written, verbal, visual, or artistic that is deliberately used to demean, threaten, or incite violence against individuals or groups based on their identity, such as race, ethnicity, religion, gender, or political beliefs. Its core intention is to dehumanize the targeted individuals, portraying them as inferior, dangerous, or unworthy of dignity, thereby fostering exclusion and division within society” interview E-1 said.

However, despite the minor impact, there are people and media platforms spreading inauthentic content on social media platforms, which remains a challenge for us, he added.

The role of youth empowerment and media literacy in addressing hate speech. Educating young individuals on responsible online behavior, critical thinking, and media analysis can equip them with the skills to discern and counter hate speech effectively (Rahmi, 2024).

“The quick and widespread spread of hate speech, especially on social media and digital platforms, poses significant challenges for those trying to reduce its impact. Unlike traditional media, the internet's decentralized structure lets harmful content circulate at lightning speed,

often outstripping our ability to detect and manage it. Plus, the anonymity that online spaces provide gives a sense of security to those spreading hate, making it tough to hold them accountable” interview E-3 said.

“The training has successfully raised awareness among individuals who were previously uninformed, teaching them valuable skills and information that they did not know before. This heightened awareness indicates that the training has been effective as it bridges gaps in knowledge and provides relevant facts” interview E-7 said.

Interview E-8 articulates about the issue for now, I see the problem of hate speech in Ethiopia is connected to ethnic divisions, political polarization, and the swift growth of digital media.

Hate speech can appear in different forms, and one especially damaging type is the distribution of modified or private pictures of individuals, particularly when these images are intended to incite anger, prejudice, or violence towards them. Take, for instance, the act of posting a carefully edited image of someone along with false claims that portray them as a danger to society. This is a clear example of hate speech, as it strips away the person's humanity and fuels public animosity” interview E-8 added.

The role of community engagement, arts education, and philosophical discussions in countering hate speech. Promoting dialogue, tolerance, and understanding through community initiatives and educational programs can foster a culture of respect and empathy, mitigating the prevalence of hate speech (Rahmi, 2024).

Interview E-6 articulates about the issue about hate speech

Hate speech is a harmful occurrence in which information that is false, overstated, or distorted is spread deliberately to vilify certain groups and present itself as the truth. Through targeted training schemes, people who had no idea about these strategies before have gained crucial awareness regarding how misinformation is tactically crafted to appear credible and convincing” said E-6.

E-4 he articulates about the awareness of hate speech:-

The Ethiopian Media Authority plays a key role in overseeing the media and shaping discussions to reduce hate speech and misinformation in Ethiopia. As the primary regulatory organization, it observes how media operates, adhering to legal and ethical guidelines while keeping an eye on content that might provoke conflicts. The authority develops policies and protocols to identify and address misinformation and hate speech, as well as conducts public campaigns to inform citizens about responsible media consumption. The authority also offers training for journalists to enhance ethical journalism and fact-checking skills, promoting professionalism in the media. Working together with civil organizations, international partners, and other stakeholders, the authority encourages initiatives that strengthen social unity. Furthermore, it enforces laws by applying penalties on violators and monitors social media platforms, where hate speech and misinformation can spread rapidly. Through these actions, the Ethiopian Media Authority aims to create a media culture in Ethiopia that is more responsible and honest.

“They are particularly teaching people not to be deceived by lies and disinformation, and they are explaining in schools that they should not accept what they come across on social media as facts” interview E-5 said.

This tells as how social media is playing in framing hate speech issue amplifier and shaper of hate speech because it provides an environment where divisive content can instantaneously go viral and reach vast audiences with minimal oversight. The anonymity and ease of sharing on social media allow malicious actors to distort facts, manipulate images, or spread lies that target some groups while hiding hate speech as opinion or news.

Emphasizes the significance of improving digital literacy among the population to combat hate speech. Educating individuals on identifying and responding to hate speech can empower them to navigate online spaces responsibly and contribute to creating a more respectful online discourse (Rahmi, 2024).

“Fact-checking is a way of checking various websites to verify the accuracy of various false information and the channels or channels through which it was disseminated” interview E-9 said.

Most of the informants also mentioned the indispensable role of fact checking for verification of information on the digital age where information spreads at unprecedented speed and scale. Considering that the public is a larger consumer of information, Ethiopia Media Authority has

done little to educate the public or build significant skill and capacity in hate speech and fact checking.

#### **4.7. Challenges and effectiveness in combating hate speech and misinformation: awareness strategies and impact measurement**

Almost all informants agreed on there are various Challenges in promoting awareness about the hate speech and fact checking. These challenges are:-

The most significant challenge is limited access and coverage, particularly in rural and excluded communities whose media and internet coverage is unpredictable and hinders effective dissemination of teaching content. The lack of societal awareness about the dangers of these issues in its efforts to counter hate speech and disinformation interview E-4 said.

According to interview E-1 the challenges we face in promoting awareness of hate speech and fact checking

Political bias presents a significant challenge since raising awareness tends to be perceived as the goals of the government instead of focusing on what the public truly needs. Additionally, awareness is often regarded mainly as a community or local matter, rather than something that affects the entire nation, while it is presumed that the Ethiopian Media Authority operates independently.

In addition, scarcity of resources, digital literacy deficits, and extensive spreading of false information through social media hinder the ability of the authority to conduct successful awareness campaigns, interview E-1 added.

Financial limitations:-which restrict the scope and sustainability of awareness campaigns, training programs, and technological tools.

“It takes a lot of resources, knowledge, and high budget to communicate effectively, which prove to be barriers in scaling up and scaling out awareness programs. On other hand cultural and linguistic diversity in Ethiopia also poses difficulties, since content must be translated into various regions, dialects, and settings to resonate with various audiences” interview E-2 said.

Influence of technology and social media: Rapidly evolving technology and social media platforms provide both opportunities and challenges. While they offer avenues for reaching a wide audience, they also expose youth to misleading information.

“The common use of social media by youths has become a double-edged sword. On the one hand, it offers learning, networking, and self-expression opportunities, but increasingly youths' users struggle with unrestrained use and addiction” interview E-5 said.

Grassroots campaigns, education for media literacy, and coordination with civil society are slowly altering perceptions, instilling growing awareness of the dangers of hate speech and the importance of seeking the truth.

“Due to their lack of formal training in fact-checking and hate speech recognition, the majority of students may be more vulnerable to false information” said T-2.

One of the major limitations in creating awareness on hate speech and disinformation in Ethiopia lies in the limitations of AI-driven fact-checking tools, particularly due to linguistic diversity and the paucity of localized data. Ethiopia is a multilingual country with over 80 languages, including Amharic, Oromo, Tigrinya, and Somali, with various dialects and contextual meanings for each.

Linguistic tools play a crucial role in developing methods to identify fake news and hate speech. Yet, languages that are considered "low-resource," especially those found in Africa, are lacking these vital resources. Ethiopian languages fall into the category of "low-resource" languages as well, facing a shortage of necessary tools and resources for tasks related to natural language processing and other tech-based language activities. The absence of suitable datasets and effective word embedding's has hindered the ability to create reliable detection methods. However, advancements in natural language processing and comprehension have enhanced the ability to identify and respond to fake news and hate speech in text streams more accurately, employing various techniques (Salau2, 2022).

The importance of leveraging advanced technologies such as natural language processing, deep learning models, and language detection algorithms to detect and combat hate speech on social

media platforms effectively. Implementing these technologies can enhance the efficiency of hate speech detection and enable prompt responses to problematic content (Rahmi, 2024).

Interview E-4 articulates about the AI technology

“The AI-based fact-checking models require large, high-quality datasets to train but lack such datasets in Ethiopian languages, which makes it difficult to develop accurate and trustworthy detection models. The majority of AI tools are fine-tuned for widely spoken global languages like English, and low-resource languages are underrepresented”.

This gap not only limits the scalability of fact-checking initiatives but also leaves non-English speaking communities to greater risk of harmful content, undermining more general awareness efforts. Addressing these challenges involves investment in localized AI research, consultation with Ethiopian linguists, and the creation of open-source datasets to improve algorithmic fairness and effectiveness in the Ethiopian environment, interview E-4 added.

A major drawback of professional fact-checking is its lack of scalability, as it involves a considerable amount of time and effort to examine and confirm a specific claim (Danula Hettiachchi, 2023).

To effectively evaluate the effectiveness of communication efforts aimed at combating hate speech and supporting fact-checking efforts, it is necessary to create a sophisticated evaluation system with qualitative and quantitative measures. For instance, opinion polls that survey the awareness levels of hate speech definitions and the importance of fact-checking could provide good data on trends in societal attitudes over time. In addition, comparing the rate and nature of reported incidents of hate speech prior to and following Ethiopian Media Authority initiative introduction would yield empirical evidence of impact. This approach aligns with evidence indicating that disinformation not only deepened conflict but also furthered social cleavages in Ethiopia, making the targeted education campaigns appealing to multiple communities of particular importance (Assen, 2023). Moreover, encouraging media outlets to cooperate with civil society can support outreach efforts in an effort to get messages across to vulnerable populations most exposed to misinformation, hence boosting community resilience to polarizing narratives

Interview E-4 articulates about the measurement

Several key metrics should be considered. First, tracking the incidence rate of hate speech by analyzing reports from the Ethiopian Media Authority and other organizations before and after implementation can reveal trends. Public sentiment analysis, conducted through surveys or social media monitoring, helps gauge shifts in attitudes toward hate speech. Additionally, engagement metrics such as likes, shares, and comments on anti-hate speech content can indicate how effectively messages resonate with audiences. Monitoring reporting rates of hate speech incidents by the public to authorities or online platforms provides insight into changes in community responsiveness. Gathering community feedback through qualitative methods offers deeper understanding of public perceptions regarding hate speech prevalence. Finally, assessing awareness levels via pre- and post-campaign surveys on media literacy and hate speech definitions helps measure educational impact. Together, these metrics provide a comprehensive view of the strategies' effectiveness in reducing hate speech and fostering a more inclusive public.

On other hand fact checking should be measured. Verification behavior rate tracks the percentage of individuals who actively fact-check information before sharing it on social media, indicating a shift toward more cautious online habits. Source credibility assessment measures improvements in the public's ability to distinguish reliable sources from unreliable ones, which can be evaluated through surveys or practical tests. Monitoring engagement with verification tools, such as fact-checking websites or browser extensions, reveals whether audiences are utilizing available resources to combat misinformation. Training participation metrics, including attendance and completion rates of workshops on media literacy, help determine the reach and effectiveness of educational initiatives. Additionally, behavioral change surveys can capture self-reported improvements in information verification practices, offering insight into personal accountability. Finally, tracking the reduction in misinformation spread, particularly regarding hate speech, provides concrete evidence of whether these strategies are successfully limiting the circulation of false or harmful content. Together, these metrics offer a comprehensive evaluation of how well communication efforts are fostering a more informed and discerning public.

## CHAPTER FIVE

### CONCLUSIONS AND RECOMMENDATION

This section includes a concise summary of the study's main findings, conclusions, recommendations, and suggestions for future research. These components are presented based on the specific objectives and corresponding research questions that were formulated to address the overall aim of assessing Effective Communication Strategies for Raising Awareness of Hate Speech and Fact-Checking: A Case Study of the Ethiopian Media Authority.

#### 5.1. Conclusion

In Ethiopia's complex socio-political landscape, the rapid spread of hate speech and misinformation on digital platforms poses significant challenges, particularly given the country's ethnic diversity. The Ethiopian Media Authority plays a critical role in addressing these issues by regulating media practices and promoting ethical journalism. Through automated monitoring software, partnerships for fact-checking, and interactive social media engagement on platforms like Facebook and TikTok, the Ethiopian Media Authority has made notable strides in identifying and countering harmful content. Initiatives such as the "9192" hotline, grassroots outreach, and journalism training further enhance media literacy and encourage responsible reporting, fostering a culture of critical engagement among citizens.

Despite these efforts, the Ethiopian Media Authority faces substantial obstacles that hinder its ability to fully combat hate speech and misinformation. Limited internet access in rural areas restricts the reach of digital campaigns, while Ethiopia's linguistic diversity complicates AI-driven detection, as tools tailored to local languages are scarce. Political concerns about bias and insufficient funding further exacerbate these challenges, underscoring the need for localized AI solutions, broader digital literacy programs, and deeper collaboration with civil society and international organizations. These systemic barriers must be addressed to ensure the Ethiopian Media Authority's strategies are effective and inclusive across Ethiopia's diverse population.

The Ethiopian Media Authority's multifaceted approach, integrating reverse image searches, verification applications, policy enforcement, and informative campaigns, demonstrates significant potential for reducing hate speech and improving content verification. By leveraging

social media platforms and community outreach, the Ethiopian Media Authority not only counters false information but also promotes a more informed public discourse. However, achieving lasting success requires sustained investment in adaptive technologies and comprehensive media literacy initiatives to bridge gaps in access and understanding, particularly in underserved regions.

Ultimately, this study underscores the urgent need for robust communication frameworks to mitigate hate speech and enhance the credibility of media institutions in Ethiopia. By adopting a fact-based approach, utilizing both quantitative metrics (such as reduced hate speech complaints) and qualitative measures (like public opinion surveys), the Ethiopian Media Authority can refine its strategies for greater impact. Strengthening partnerships with global and local stakeholders will further bolster these efforts, creating a resilient media landscape that empowers citizens and fosters accountability in an increasingly digital world.

## **5.2. Recommendation**

Based on the findings, the following recommendations are proposed to strengthen the Ethiopian Media Authority's efforts in combating hate speech and misinformation while promoting fact-checking and media literacy:

- **Expand Digital Campaigns:** Increase the use of interactive content (short videos, infographics, and quizzes) on platforms like Facebook, TikTok, and YouTube to engage younger audiences.
- **Collaborate with Fact-Checking Organizations:** Partner with global fact-checking platforms (PolitiFact, Snopes) and local initiatives to enhance verification capabilities.
- **Utilize the strategic leverage of broadcast media,** particularly television and radio, which remain the most accessible and trusted sources of information in these areas. As internet penetration is low among rural areas, radio programs in local languages can readily disseminate messages through talk shows, dramas, and public service announcements that advise listeners on how to identify and report hate speech.
- **Collaborate closely with the Ministry of Education** to embark on integrating media literacy and fact-checking abilities into the national curriculum of Ethiopia's schools.

This would instill lifelong abilities in students to interact with the digital world responsibly from an early stage

- Provide Target Vulnerable Groups specialized training for students, journalists, and social media influencers on identifying and countering hate speech.
- Improve the efficacy and reach of its "9192" hotline by embracing a multi-pronged strategy encompassing awareness campaigns, incentives for users, and easier reporting mechanisms. Because the public may be unaware of the hotline or may be hesitant to call, Ethiopian Media Authority must roll out targeted outreach campaigns e.g., radio jingles, Short Message Services (SMS) blasts, and social media campaigns to educate the public on how and why to report offending content.
- To ensure that communities in the regions of the country, especially those who cannot access alternative media, understand the various options for protection so that they do not become easy victims of hate speech.
- The Ethiopian Media Authority needs to conduct a fair and open sanction process, which may involve warnings, penalties, temporary suspensions, or, in serious cases, revoking licenses. Furthermore, the Ethiopian Media Authority ought to create a searchable database available to the public that lists violations and enforcement actions to promote accountability and deter future wrongdoings.
- When implementing the Hate Speech and Disinformation Prevention Proclamation, the Ethiopian Media Authority should create specific and unbiased standards to distinguish between harmful material, such as incitements to violence or intentional misinformation, and protected speech, which includes political analysis, satire, and investigative reporting.
- The Ethiopian Media Authority needs to establish a method grounded in evidence to evaluate its influence on society. This entails examining current data and collecting people's thoughts via surveys. Consistent surveys across the country should monitor shifts in attitudes toward hateful speech and false information, confidence in media sources, awareness of fact-checking tools, and perspectives on the Ethiopian Media Authority's neutral approach to regulation.

- To improve its efforts in fighting against false information and hateful speech, the Ethiopian Media Authority needs to work more closely with universities, tech firms, and civil society groups. This collaboration would help boost public awareness programs and fact-checking projects. Additionally, the Ethiopian Media Authority should look for lasting financial support by combining private partnerships, international funding, and government backing.

## Bibliography

- Agata de Latour, N. P. (2017). taking action against hate speech through counter and alternative narratives. *CouncilOfEuropePublications*.
- Ahmed, A. (2021). *Regulating Inter-Ethnic Induced Hate Speech in Ethiopia* .
- Alessandra, T. (1993). *Communicating at work*. New York: Fireside publishers.
- Alkiviadou, N. (2019). *Hate speech on social media networks: Towards a regulatory framework?Information & Communications Technology Law*.
- Amazeen, L. G. (2019). *Fact-Checking as Idea and Practice in Journalism* .
- Assen, M. D. (2023). *Challenges of governance in an age of disinformation and hate speech in Africa: A lesson from the post-2018 Ethiopian experience*.
- Atif Ahmad, J. W. (2019). *Strategically-Motivated Advanced Persistent Threat*::.
- Authority, E. M. (2024). *Ethiopian Media Authority national report on social media hate speech*.
- Awan, I. (2016). Islamophobia on Social Networks: A Qualitative Analysis of the Facebook's Walls of Hate. *International Journal of Cyber Criminology*, 10(1), 1–20. DOI: 10.5281/zenodo.58517.
- Azoulay, L. (2024). *Truth in the Crossfire: The Case of Ethiopia and Fact-Checking in Authoritarian Contexts*.
- Babbie, E. (. (2010). *The practice of social research*. Wadsworth.
- Barendt, E. (2019). *What Is the Harm of Hate Speech ?*
- Benvenuto, F. M. (2023). *Hate Speech Recognition: The Role of Empathy and Awareness of Social Media Influence*.
- Bowen, G. A. (2009). *Document Analysis as a Qualitative Research Method*.
- Brown, A. (2017). *What is hate speech? Part 1: The Myth of Hate*. Law and Philosophy.
- Carlson, C. (2021). *Hate Speech. Massachusetts*.
- CIALDINI, R. B. (1998). *social influence: social norm, conformity and compliance* .
- Cover, R. (2022). *Digital hostility, subjectivity and ethics: Theorising the disruption of identity in instances of mass online abuse and hate speech*.
- Danula Hettiachchi, K. J. (2023). *Designing and Evaluating Presentation Strategies for Fact-Checked Content*.
- Delgado, R. &. (1995). *Images of the outsider in American law and culture: Can free expression remedy systemic social ills?*.

- Denu, T. (2021). Appraisal of Ethiopian Hate Speech and.
- Ekin Balkan, S. Ü. (2023). *A PRIMER ON MISINFORMATION, MALINFORMATION AND DISINFORMATION*.
- Elliott, C. C. (2016). Hate Speech, Key concept paper. Media, Conflict and Democratisation.
- Endert, K. P. (2021). Developing Fact-Checking strategies with Human Centered Design — insights of an approach in Ethiopia.
- Farid, A. S. (2023). Local Government Communication Strategy in Overcoming the Spread of Hoaxes and Hate Speech in the General Election in Mandailing Natal.
- Fátima, B. D. (2023). *Hate Speech on Social Media*. LabCom Books in co-edition with OdeCom and EdiPUCE.
- Garbe, L. S. (2023). How African countries respond to fake news and hate speech. *Information, Communication & Society*, 26(1), 86-103. <https://doi.org/10.1080/1369118X.2021.1994623>.
- GRIFFIN, E. (2012). *A FIRST LOOK AT COMMUNICATION THEORY, EIGHTH EDITION*. Avenue of the Americas, New York, NY 10020: David Patterson.
- Haiman, F. S. (1964). Some reflections on the teaching of freedom of speech. *Free Speech Yearbook*, 3, 26–30.
- Haiman, F. S. (2003). Introduction. In R. A. Parker (Ed.), *Free speech on trial: Communication perspectives on landmark Supreme Court decisions* (pp. 1–8). Tuscaloosa, AL: University of Alabama Press.
- Halperin, E. (2011). *Emotional Barriers to Peace: Emotions and Public Opinion of Jewish Israelis about the Peace Process in the Middle East*.
- Hargie, O. (2016). *Skilled Interpersonal Communication: Research, Theory and Practice*. London: Routledge Publishers.
- Hefner, S. G. (2019). *Social norms as communicative phenomena*.
- Hollomotz, A. (2013). Disability, oppression, and violence: Towards a sociological explanation. *Sociology*, 47(3), 477-493.
- <https://journals.ala.org/index.php/ltr/article/view/6497/8636>. (2017). Retrieved 1 28, 2025, from <https://journals.ala.org/index.php/ltr/article/view/6497/8636>.
- Humphreht.E. (2020). How Do They Debunk “Fake News”? A Cross-National Comparison of Transparency in Fact Checks. *Digital Journalism*, vol. 8(3), pp. 310–327.
- Ilker Etikan, S. A. (2016). Comparison of Convenience Sampling and Purposive Sampling.

- K. Muller and C. Schwarz. (2020). FANNING THE FLAMES OF HATE: SOCIAL MEDIA AND HATE CRIME.
- Kamga, O. (2019). Difficile émergence d'un journalisme objectif en Afrique: Origine du drame et pistes de.
- Katja Rost, L. S. (2016). *Digital Social Norm Enforcement: Online Firestorms in Social Media*.
- Keith, W. (2008). On the origins of speech as a discipline: James A. Winans and public speaking as practical democracy. *Rhetoric Society Quarterly*, 38, 239–258.
- Kimotho, S. G. (2016). Digitized Ethnic Hate Speech: Understanding Effects of Digital Media Hate Speech on Citizen Journalism in Kenya .
- Koontz & O'Donnell-Trujillo, N. (1983). Organizational communication as cultural performance. *Communications Monographs*, 50(2), 126-147.
- Kudlick, C. J. (2003). Disability history: Why we need another “other”. *The American Historical Review*, 108(3), 763-793. .
- Kumar, R. (2011). DEVELOPMENT COMMUNICATION: A PURPOSIVE COMMUNICATION WITH SOCIAL CONSCIENCE-AN INDIAN PERSPECTIVE. *Global Media Journal: Indian Edition*, 2(2).
- Kumar.A. (2022). Fact-Checking Methodology and its Transparency: What Indian Fact-Checking Websites Have to Say? *Journalism Practice*, doi: 10.1080/17512786.2022.2098520.
- L.Shaw, M. E. (1993). *The Evolution of Agenda-Setting Research: Twenty-Five Years in the Marketplace of Ideas*.
- Lunenburg, F. C. (1991). *Educational administration: Concepts and practices*. Wadsworth. Inc. USA.
- Madrid-Morales, D. &. (2021). Disruptions and transformations of digital media in Africa: An. *Journal of African Media Studies*.
- Mafeza, F. (2016). Preventing genocide by fighting against hate speech.
- Matsuda, M. J. (1993). *Public Response to Racist Speech: Considering the Victim's Story*.
- McCombs, M. (2005). *A Look at Agenda-setting: past, present and future*.
- Miller, I. (2020). 'Research Guides: Journalism: Fact-Checking Sites'. .
- MOGHADDAMA, S. A.-N. (2022). Exploring Advanced EFL Learners' Awareness of Communication Strategies.
- Morley, D. &. (2002). Media audiences: Communication and context: Ethnographic perspectives on the media audience. In *A handbook of qualitative methodologies for mass communication research* (pp. 149-162). Routledge.

- Muluken A.I, Mulatu A. and Biset A. (2021). Social media hate speech in the walk of Ethiopian politicalreform: analysis of hate speech prevalence, severity, and natures.
- Muluken Asegidew Chekol, M. A. (2023). *Social media hate speech in the walk of Ethiopian political reform: analysis of hate speech prevalence, severity, and natures, Information, Communication & Society.*
- Naganna Chetty, S. A. (2018). *Hate speech review in the context of online social networks.*
- Nations, U. (2019). United Nations Strategy and plan of action on hate speech.
- Nations, U. (2020). United Nations Strategy and plan of action on hate speech.
- Nigussie, H. (2024). Echoes of Violence:Analyzing Hate Speech in the Ethiopian Broadcast Media.
- O'keefe, D. J. (2008). Elaboration likelihood model.
- Posetti, J. (2019). A short guide to the history of 'fake news' and disinformation.
- Quandt, T. (2018). Dark Participation.
- Rahmi, D. K. (2024). Empathy and Hate Speech in Social Media: The Case of Indonesia. *Dr. Kus Hanna Rahmi.*
- Rimal, R. N. (2009). Why health communication is important in public health. *Why health communication is important in public health.*
- Rimal, R. N. (2009). Why health communication is important in public health. Bulletin of the World Health Organization, 87(4), 247- 247a.
- Rimer, B. K. (2006). Advancing tailored health communication: A persuasion and message effects perspective. Journal of Communication, 56, 184-201.
- S. Pukallus, \*. a. (2024). *Combating Hate Speech on Social Media: Applying Targeted Regulation, Developing Civil-Communicative Skills andUtilising Local Evidence-Based Anti-Hate Speech Interventions.*
- S. Schäfer, I. R. (2023). *Can We Counteract Hate?Effects of Online Hate Speech and Counter Speech on the Perception of Social Groups.*
- Seglow1, J. (2016). *Hate Speech, Dignity and Self-Respect.*
- Shukla, S. (2020). *Concept of Population And Sample.*
- Skjerdal, T. &. (2021). The ethnification of the Ethiopian media. *The ethnification of the Ethiopian media.*
- Stevens, B. (2011). Structural barriers to sexual autonomy for disabled people. Hum. Rts., 38, 14.
- Suaahara, A. F. (2016). An FGD conducted with front line workers of Suaahara (Good. *HERD.*

- Szafrańska, M. (2024). War Fact-Checking: Strategies and Methodology for Debunking Disinformation on the War in Ukraine.
- Tatum, B. D. (2001). Defining racism: Can we talk. Race, class, and gender in the United States: An integrated study, 100-107.
- Tekoniemi, S. K. (2022). - Tekoniemi, S., Kotilainen, S., Fact-checking as digital media literacy in higher education. Seminar.net, . - Tekoniemi, S., Kotilainen, S., Fact-checking as digital media literacy in higher education. Seminar.net, .
- Tornberg, A. &. (2016). *Muslims in social networks discourse: Combining topic modeling and critical discourse analysis. Discourse, Context and Networks, 13, 132–142, <http://dx.doi.org/10.1016/j.dcm.2016.04.003>.*
- Waltman, M. S. (2011). The communication of hate. New York, NY: Peter Lang Publishers.
- Waltman, M. S. (2017). Understanding hate speech. In Oxford encyclopedia of communication. Oxford, England: Oxford University Press. doi:10.1093/acrefore/9780190228613.013.422.
- Wardle, C. &. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe report.
- Wardle, C. (2019). Understanding Information disorder.
- Williams, S. S. (2021). Countering Online Misinformation in the First-Year Composition. *Countering Online Misinformation in the First-Year Composition*.
- Wodak, R. &. (1999). *Discourse and racism: European perspectives. Annual Review of Anthropology, 28(1), 175-199.*
- Workneh, T. W. (2019). *Ethiopia's hate speech predicament: Seeking antidotes beyond a legislative response.*
- Workneh, T. W. (2021). *Social media, protest, & outrage communication in Ethiopia: Toward fractured publics.*
- Worldbank. (2006). Global Information & Communication Technologies Department. (2006). 2006 Information and Communications for Development: Global Trends and Policies. World Bank Publications.
- Wubetu Demilie and Ayodeji Salau. (2022). Detection of fake news and hate speech for Ethiopian languages: a systematic review of the approaches.
- Yong, C. (2011). Does Freedom of Speech Include Hate Speech?

## Appendices

### Appendix I: Interview questions

Guidelines for In-depth Interview Questions and Focus Group Discussion

Full Name: \_\_\_\_\_

Level of Education: \_\_\_\_\_

Position in the Organization: \_\_\_\_\_

Contact Details – Phone: \_\_\_\_\_

E-mail: \_\_\_\_\_

Date of Interview: \_\_\_\_\_

**Dear sir/ Madam,**

I am Birshewa Seyoum, MA student in Multimedia journalism at Addis Ababa University. As part of my degree requirements, I am conducting a research study titled " Effective Communication Strategies for Raising Awareness of Hate Speech and Fact-Checking: A Case Study of the Ethiopian Media Authority". I sincerely appreciate your willingness to participate in an interview for my research. To ensure accuracy and efficiency, I kindly request permission to record the interview. Please be assured that any information provided will be used solely for academic purposes.

1. How do you define hate speech and fact checking?
2. How do you address the skepticism or distrust people might have toward automated fact checking system?
3. What strategies do you think are most effective in engaging different audience (e.g youth, old generation) in understanding the harmful effect of hate speech?

4. What communication channels (e.g social media, TV, public events) have you found most effective for raising awareness about hate speech? And how can traditional media and social media be leverage differently in combating hate speech?
5. What obstacles have you faced in your efforts to spread the word about hate speech and its repercussions?
6. How do you measure the effectiveness of your communication effort to combat hate speech?
7. How do you think the general public perceives the importance of fact-checking? What role do you believe fact checking plays in combating the spread of hate speech?
8. How can technology (e.g. AI) be used to improve fact checking and what a challenge does this present?
9. How do you address the skepticism or distrust people might have toward automated fact checking system?
10. What metrics or feedback mechanisms do you use to evaluate the success of your fact checking efforts and how do you ensure that the fact checking process itself is transparent and trustworthy to the public?
11. How do you perceive the role of the Ethiopian Media Authority in addressing hate speech?
12. What challenges does the Ethiopian Media Authority face when trying to implement effective communication strategies against hate speech?
13. How should Ethiopian Media Authority balance regulating hate speech with protecting freedom of expression?

## Appendix II

### Amharic version of appendix

1. የጥላቻ ንግግረ እና ሃቅ ማጣራት እንዴት ይገለጻል?
2. ሰዎች በራስ አገዝ የሃቅ ማጣራት ስርዓት ላይ ሊኖራቸው የሚችለውን ጥርጣሬ ወይም አለመተማመን እንዴት መፍታት ይቻላል?
3. የተለያዩ ታዳሚዎችን (ለምሳሌ ወጣቶች፣ ጎልማሳ፣ በእድሜ ገፋ ያሉ ስዎች) የጥላቻ ንግግርን ጎጂ ውጤት ለመረዳት ምን አይነት ስልቶች ውጤታማ ናቸው ተብለው ይታመናሉ?
4. የጥላቻ ንግግር ግንዛቤን ለማሳደግ የትኞቹን የመገናኛ ብዙሃን ማስተላለፊያ መንገዶች (ለምሳሌ ማህበራዊ ሚዲያ፣ ቲቪ፣ ህዝባዊ ዝግጅቶች) ውጤታማ ናቸው? እና ባህላዊ ሚዲያ እና ማህበራዊ ሚዲያዎች የጥላቻ ንግግርን ለመዋጋት በተለያዩ መንገድ እንዴት ሊጠቀሙበት ይችላሉ?
5. የጥላቻ ንግግሮችንና መዘዞቹን ለማስፋፋት በምታደርገው ጥረት ምን እንቅፋት አገጥሟል?
6. የጥላቻ ንግግርን ለመዋጋት የምታደርጉትን የግንኙነት ጥረት ውጤታማነት እንዴት ይለካል?
7. ህዝቡ ሃቅ የማጣራት አስፈላጊነት እንዴት ይገነዘቡዋል? የጥላቻ ንግግር መስፋፋትን ለመዋጋት ሃቅ ማጣራት ምን ሚና አለው ብለው ያምናሉ?
8. ቴክኖሎጂ (ለምሳሌ AI) የሃቅ ማጣራት ፍተሻን ለማሻሻል እንዴት ጥቅም ላይ ሊውል ይችላል እንዲሁም ምን ተግዳሮቶችን ሊያመጣ ይችላል?
9. ሰዎች በራስ አገዝ የሃቅ ማጣራት ፍተሻ ስርዓት ላይ ሊኖራቸው የሚችለውን ጥርጣሬ ወይም አለመተማመን እንዴት መፍታት ይቻላል?
10. የሃቅ ማጣሪያ ጥረቶችን ስኬት ለመገምገም ምን አይነት መለኪያዎችን ወይም የግብረመልስ ዘዴዎችን ይጠቀማሉ እና የፍተሻ ሂደቱ እራሱ ግልፅ እና ለህዝብ ታማኝ መሆኑን እንዴት ማረጋገጥ ይቻላል?

11. የኢትዮጵያ መገናኛ ብዙሃን ባለስልጣን የጥላቻ ንግግርን ለመፍታት ያለውን ሚና እንዴት ይመለከቱታል?

12. የኢትዮጵያ መገናኛ ብዙሃን ባለስልጣን የጥላቻ ንግግርን ለመከላከል ውጤታማ የግንኙነት ስትራቴጂዎችን ለመተግበር ሲሞክር ምን ተግዳሮቶች አጋጥመዋል?

13. የኢትዮጵያ መገናኛ ብዙሃን ባለስልጣን የጥላቻ ንግግርን ከመቆጣጠር እና ሃሳብን በነፃነት ከመግለፅ ጋር እንዴት ማጣጣም አለበት?