UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

SOCIAL MEDIA'S DOUBLE-EDGED SWORD: FREEDOM OF EXPRESSION VERSUS HATE SPEECH

by

Masego Entle Palesa Mphuthi

(18097465)

Submitted in fulfilment of the requirements for the degree
Master of Law (International Law) through Coursework

LLM

In the Faculty of Law,

University of Pretoria

July 2025

Supervisor: Prof Sylvia Papadopoulos

© University of Pretoria

EPIGRAPH

"The right to freedom of expression is not absolute; the state may impose restrictions on expression in certain circumstances, such as when the expression is likely to incite violence or hatred, or when it is considered to be defamatory or in breach of privacy."

United Nations Human Rights Committee, General Comment No. 34, Article 19: Freedom of opinion and expression.

© University of Pretoria

DEDICATION

To my spiritual guide, The Alpha and Omega, The Most High God, for blessing me with patience, wisdom, and the desire to always improve. Lord, I am grateful for your guidance, love, and support throughout my academic journey.

To my parents, and my sister who provided me with encouragement, courage, emotional and financial support. Your love and guidance have meant the world to me. I am so grateful for everything you have done for me.

To my family at large, for their understanding, help and indulgence in my studies.

© University of Pretoria

ACKNOWLEDGEMENTS

To my mentor and supervisor Professor Papadopoulos, your belief in me and your constant pushing have helped me to reach my full potential. I am so grateful for your guidance and patience.

To my family, who have always been there for me. Your understanding and support have meant the world to me.

And finally, to myself for having faith. Thank you to myself, for not giving up, even when things were tough, and the road seemed as if it was never ending. For pushing through the challenges and for never giving up on my dreams.

© University of Pretoria

ACRONYMS AND ABBREVIATIONS

ACHPR	African Commission on Human and Peoples' Rights
AI	Artificial Intelligence
AU	African Union
CERD	Convention on the Elimination of All Forms of Racial Discrimination
CEDAW	Convention on the Elimination of All Forms of Discrimination Against Women
CIPESA	Centre for Information, Development, Education and Advocacy
ECTA	Electronic Communications and Transactions Act
EU	European Union
4IR	Fourth Industrial Revolution
ICCPR	International Covenant on Civil and Political Rights
PEPUDA	Prevention of Unfair Discrimination Act
PIPL	Personal Information Protection Law
SAHRC	South African Human Rights Commission
SCA	Supreme Court of Appeal
UDHR	Universal Declaration of Human Rights
UN	United Nations
UNESCO	United Nations Education, Scientific and Cultural Organisation

© University of Pretoria

DECLARATION

I, Masego Entle Palesa Mphuthi, declare 'Social Medias' Double-edged Sword: Freedom of Expression versus Hate Speech,' which is hereby submitted for the award of Master's in Coursework in Private Law is my original work. It has not been previously submitted for the award of a degree at this or any other tertiary institution. Where works of other people are used, comprehensive references have been provided.

© University of Pretoria

ABSTRACT

Social media platforms have become an imperative part of modern life, allowing people to convey not only their thoughts but also exercise their right to freedom of expression. International human rights laws espouse the right to freedom of thought and expression for everyone. However, many countries, under certain circumstances, have laws that restrict or outrightly ban social media expression, particularly of a type that is deemed to incite hatred and violence.

Just as all eyes are meant to see, speech is meant to be heard and words are meant to be written, but it must also be prudently balanced with the right to equality and dignity. The need to balance the recurrent overlap between freedom of expression and what is perceived as hate speech creates a crucial question; where do we establish a boundary between legitimate expression and harmful speech? This analysis focuses on the legal frameworks, policies, and approaches to strike a balance between protecting freedom of expression and addressing potential social harms (being hate speech) on these platforms.

Keywords: Hate Speech, Freedom of Expression, Social Media

© University of Pretoria

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION **11**

1.1 Background 11

1.2 Problem Statement 14

1.3 Research Questions 14
1.3.1. Main Research Question 14
1.3.2 Sub-Questions 14

1.4 Motivation for Study 15

1.5 Limitations of Study 15

1.6 Research Approach and Methodology 16

1.7 Chapter Overview 16

1.9 Summation 17

CHAPTER 2: CONCEPTUAL FRAMEWORK **18**

2.1 Introduction 18

2.2 Freedom of Expression 18
2.1.1 Freedom of Expression on Social Media 19

2.3 Definition and Conceptualisation of Hate Speech 21

2.3.1 Overview 21

2.3.2 Hate Speech Definition by Social Media Platforms 24

2.4 Social Media Regulation in Global Context 25

2.5 Summation 28

CHAPTER 3: HATE SPEECH AND ITS IMPLICATIONS **29**

3.1 To Regulate or Not to Regulate 29

3.2 Types of Arguments for Regulating Hate Speech on Social Media 29

3.3 Impact of Social Media on Hate Speech Propagation 31

3.4 Challenges in Monitoring Hate Speech Effectively 33

3.5 Tension Between Freedom of Expression and Hate Speech 34

© University of Pretoria

3.6 Summation 35

CHAPTER 4: LEGAL AND REGULATORY FRAMEWORK IN SOUTH AFRICA 36

4.1 Introduction 36

4.2 Overview of South Africa's Social Media Landscape 37

4.3 Historical Context of Hate Speech In South Africa 38

4.4 The South African Legal Framework 40
4.4.1 The Constitutional Provisions and Hate Speech Legislation 40
4.4.2 The Promotion of Equality and Prevention of Unfair Discrimination Act 42
4.4.3 Prevention and Combating of Hate Crime and Hate Speech Bill [B9 of 2018] 46
4. 4.4 The Films and Publications Act 65 of 1996 49
4.4.5 The Films and Publications Amendment Act 11 of 2019 50
4.4.6 The Cybercrime Act 19 of 2020 52

4.5 Security Concerns 55

4.6 Case Studies Illustrating the Implementation and Impact of These Laws 57
4.6.1 Overview 57
4.6.2 AfriForum NPC & Another v Pienaar 2017 (1) SA 388 (WCC) 58
4.6.3 The Infamous Sparrow and Momberg cases 58
4.6.4 Islamic Unity Convention v Independent Broadcasting Authority and AfriForum NPC v Nelson Mandela Foundation Trust and Others 60
4.6.5 Cape South African Jewish Board of Deputies vs. Thomas Leyland Torr and Amaning and Others v Ackerman cases 61

4.7 Summation 63

CHAPTER 5: SUMMARY AND CONCLUSION 64

5.1 Introduction 64

5.2 Synopsis of chapters 64

5.3 Recommendations 66
5.3.1 Overview 66
5.3.2 Strengthening Legal Frameworks and Enforcement 66
5.3.3 Collaboration Among Governments, Civil Society, and Social Media Platforms 70

5.4 Conclusion 71

BIBLIOGRAPHY 72

Articles 72

© University of Pretoria

Books 78

Legislation and International Instruments 79

Case Law 80

Thesis and Dissertations 81

Other 82

© University of Pretoria

CHAPTER 1: INTRODUCTION

1.1 Background

In this chapter there is a brief introduction of the topic that lays the foundation for the rest of this essay. Freedom of expression is not absolute; therefore, the aim and objective of this chapter is to comment about the problem statement and provide an overview of other chapters.

The rise of social media has transformed the way people communicate and express themselves.[1] In addition, the anonymity that social media provides can embolden people to become more outspoken and confident in sharing their opinion.[2] However, this comes with challenges, particularly in the form of hate speech.[3] This research aims to explore the issue of freedom of expression on social media, with a specific focus on the problem of the regulation of hate speech in South Africa. International human rights instruments such as the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), and the African Charter on Human and Peoples' Rights will be referenced to provide a comprehensive analysis of the protections afforded within freedom of expression.

Hate speech is a phenomenon whose existence predates the internet and social media, but these new technologies have made it easier to spread and amplify hate speech and have created new ways to attack and harm individuals and groups.[4] Hate speech on social media is a legal issue that has been the subject of much debate around the world.[5] Countries struggle to balance the right to freedom of expression, which is essential for a democracy, with the need to protect people and minority groups from harm caused by hate speech.[6]

---

1 JA Tucker, Y Theocharis, ME Roberts & P Barbera “From Liberation to Turmoil: Social Media and Democracy” (2017) 28 Journal of Democracy 46.

2 B Di Fátima “Hate Speech on Social Media: A Global Approach” (2023) Communication Library vol 7 available at https://store.lexisnexis.co.za/products/hate-speech-on-social-media-a-global-approach-skuZASKUPG00 (accessed 9/10/2023).

3 *Ibid at 2.*

4 N Alkiviadou “Hate Speech on Social Media Networks: Towards a Regulatory Framework?” (2019) 28(1) Information & Communications Technology Law 19.

5 RA Wilson & MK “Land, Hate Speech on Social Media: Content Moderation in Context” (2021) 52 Connecticut Law Review 1029.

6 United Nations "Hate Speech Versus Freedom of Speech" (2023) available at https://www.un.org/en/hate-speech/understanding-hate-speech/hate-speech-versus-freedom-of-speech (accessed 31/10/2023).

© University of Pretoria

In South Africa, hate speech on social media has become a major concern as it can have a significant impact on society.[7] Freedom of expression is a fundamental but not absolute human right in South Africa. Section 16 of the Constitution of South Africa, 1996 (hereafter the Constitution) guarantees the right to freedom of expression, which includes the freedom to receive or impart information or ideas, freedom of artistic creativity, and academic freedom. However, this right does not extend to promoting war, inciting violence, or advocating for hatred based on race, ethnicity, gender, or religion, if such advocacy incites harm.[8]

This can be seen in a 2019 case where the South African Human Rights Commission (SAHRC) ordered Facebook to remove several hate speech posts. Included was one by Velaphi Khumalo, who was found guilty of hate speech by the Equality Court after posting racial slurs against White people in a Facebook tirade.[9] South Africa has been making significant strides in updating its laws to address hate speech on social media platforms like Twitter, WhatsApp and others.[10]

The regulation of social media is the process of governing the content and behaviour of social media platforms and users, which can be undertaken by governments, social media platforms themselves, or a combination of both.[11] The South African government has taken steps to regulate social media content in the name of national security and public safety.[12] This regulation can include, but not limited to, elements such as the restriction of harmful or illegal content, cyberbullying, online grooming and

---

7 K Harrisberg & N Bhalla "As Online Hate Speech Hits Africa, Social Media Firms Told to Act" (7/08/2022) IOL available at https://www.iol.co.za/news/africa/as-online-hate-speech-hits-africa-social-media-firms-told-to-act-3b809769-6b06-533d-a2cb-d834cdfe2440 (accessed 17/10/ 2023).

8 Section 16(2) Constitution.

9 SAHRC, 'Court Finds Government Official Velaphi Khumalo Guilty of Hate Speech' (5 October 2018), available at https://www.thesouthafrican.com/news/velaphi-khumalo-hate-speech-guilty/ (accessed 17/10/2023).

10 Staff Writer "New Laws to Make Hate Speech A Crime In South Africa – Including on Twitter, Whatsapp and Other Social Media" (15/03/2023) Business Tech available at https://businesstech.co.za/news/trending/672725/new-laws-to-make-hate-speech-a-crime-in-south-africa-including-on-twitter-whatsapp-and-other-social-media/ accessed (17/10/2023).

11 M Yar "A Failure to Regulate? The Demands and Dilemmas of Tackling Illegal Content and Behaviour on social media" (2018) 1(1) International Journal of Cybersecurity Intelligence & Cybercrime 5.

12 N Nyathi & K Rajuili "South Africa and Kenya's Legislative Measures to Prevent Hate Speech" (21/07/2017) ACCORD, available at https://www.accord.org.za/conflict-trends/south-africa-kenyas-legislative-measures-prevent-hate-speech/ (accessed 17/10/2023).

© University of Pretoria

harassment.[13] The prevention of the spread of propaganda, the planning of terrorist attacks, the leaking of sensitive information, and the spread of misinformation are also included in this regulatory framework.

Although section 16 of the Constitution guarantees the right to freedom of expression, it also allows for this right to be restricted under specific circumstances, where national security, public order and the rights of others might be unduly threatened. Additionally, the Constitution provides for the limitation of rights in particular conditions. This is outlined in section 36 of the Constitution which delineates the process to ensure proportionate and justified restriction of human rights.

According to section 36, rights contained in the Bill of Rights may be limited only by a law of general application, and the limitation must be reasonable and justifiable in an open and democratic society based on human dignity, equality and freedom.[14] The relevant factors to be considered include the nature of the right, the importance of the purpose of the limitation, the extent and nature of the limitation, the relationship between the limitation and its purpose, and whether there are less restrictive means to achieve the purpose.[15]

The government has also established social media policy guidelines for official government communication purposes and personal use of social media by government employees.[16] These guidelines aim to improve government transparency, participation and interaction with the public.[17] These measures indicate that the South African government is actively working to regulate social media content for national security and public safety as well as protecting privacy, enforcing intellectual property rights and promoting fairness and transparency.[18]

---

13 eSafety Commissioner Australia "What Is Illegal and Restricted Online Content?" (2023) available at https://www.esafety.gov.au/report/what-is-illegal-restricted-content (accessed 18/10/2023).

14 Constitution of the Republic of South Africa, 1996 (hereafter the Constitution), section 36.

15 eSafety Commissioner Australia "What Is Illegal and Restricted Online Content?" (2023) available at https://www.esafety.gov.au/report/what-is-illegal-restricted-content (accessed 18/10/2023).

16 GCIS "Social Media Policy Guidelines" (2011) Government Communication and Information System available at https://www.gcis.gov.za/sites/default/files/docs/resourcecentre/guidelines/social_media_guidelines_final_20_april2011.pdf (accessed 17/10/2023).

17 *Ibid 16*.

18 GCIS "Social Media Policy Guidelines *op cit*.

© University of Pretoria

1.2 Problem Statement

The research problem that this study will address is the complex issue of how South Africa, a country with a long history of social divisions expressed along racial, ethnic, tribal and religious lines, balances the right to freedom of expression on social media with the need to regulate hate speech.

The spread of hate speech on social media has been associated with an increase in violence and discrimination against marginalised groups, eroding the principles of democracy and human rights.[19] There are newly developed legal frameworks and regulations in place to address these issues, like the Prevention and Combating of Hate Crimes and Hate Speech Bill[20] and the Cybercrimes Act.[21] Their effectiveness in balancing the protection of freedom of expression with the need to combat hate speech is still evolving. The goal is to identify areas where South Africa's regulatory approach falls short, and to propose improvements that will better safeguard freedom of expression while addressing the inherent challenges of the digital era.

1.3 Research Questions

1.3.1. Main Research Question

The main research question is: To what extent are the legal frameworks and regulations in South Africa effective in striking a balance between the right to freedom of expression and the need to protect against hate speech?

1.3.2 Sub-Questions

The following related sub-questions help to refine the study for this essay:

How is hate speech defined and regulated in South Africa?

What measures can be taken to combat hate speech on social media in South Africa considering international human rights instruments such as the UDHR, ACHPR and the ICCPR?

---

19 UN News “Independent UN Expert Says, ‘Tsunami of Hate’ Targeting Minorities Must Be Tackled” (15/03/2021) available at https://news.un.org/en/story/2021/03/1087412 accessed 18/10/2023).

20 Prevention and Combating of Hate Crimes and Hate Speech Bill’ (B9-2018).

21 Cybercrimes Act 19 of 2020.

© University of Pretoria

1.4 Motivation for Study

The exponential growth in usage and adoption of social media platforms such as WhatsApp, Instagram, X (formerly Twitter), and TikTok by society at large induced this study. The prevalent positive or negative encounters on social media platforms by the public, move one to wonder about the laws governing these platforms. Each platform has its own terms and conditions which serve as a guideline for users. Consequently, individuals may use these applications to express themselves as they please within the confines of their rules. However, the rules set by these apps may be subject to varying interpretations and be inconsistent with the country's constitution.

This research aims to contribute to the understanding of the challenges and potential solutions in balancing the freedom of expression and addressing harmful content on social media platforms in South Africa. As indicated, international human rights instruments will be examined in comparison to provide valuable insights into the specific challenges of disinformation and hate speech in South Africa. The findings and recommendations of this research will be of interest to policymakers, social media platforms and civil society organisations seeking to promote a safe and inclusive online environment while upholding the principles of freedom of expression.

There also seems to be declining support for media freedom in Africa and the potential reasons for this include several factors, such as concerns about bias, hate speech, the rise of authoritarian governments, the increasing use of social media to spread disinformation, and the growing distrust of the media among the public.[22] These are aspects that are noted in this essay.

1.5 Limitations of Study

The research is constrained by the availability and quality of data on hate speech on social media platforms in South Africa.

Individual biases inherent in society on the basis of race, ethnicity and other demographic elements may present difficulty in conducting this research as all these factors have a direct influence on how and what people may perceive as harmful content leading to or constituting hate speech on social media platforms.

---

22 J Conroy-Krutz “The Squeeze on African Media Freedom” (16/10/2023) Journal of Democracy available at https://www.saflii.org/za/cases/ZASCA/2019/167.pdf (accessed 16/10/2023).

© University of Pretoria

1.6 Research Approach and Methodology

The researcher has conducted an in-depth desk study and reviewed relevant literature that has been published on topics related to the research topic. The research is primarily qualitative, meaning that the researcher focuses on interpreting and understanding existing literature to draw conclusions. Finally, a legal assessment is conducted to examine the implications of hate speech on social media in South Africa, considering international human rights instruments such as the UDHR and the ICCPR. This entails analysing relevant legislation, case law, and international human rights instruments.

1.7 Chapter Overview

This essay consists of five chapters.

Chapter one introduces the topic and provides background information that lays the foundation for the rest of the research.

Chapter two explores the intricate relationship between freedom of expression and hate speech. It first defines these concepts through international frameworks such as the ICCPR and the UDHR, while also considering the African context and the relevant principles set forth in the ACHPR and the African Charter on Democracy, Elections and Governance.

This is followed by the examination of social media as a platform for exercising freedom of expression, discussing both its positive aspects as an enabler of democracy, citizen journalism, and the exchange of ideas, as well as its potential for misuse. The tension between freedom of expression and hate speech emphasises the potential harm that hate speech can cause to individuals and groups. In essence, this chapter offers a nuanced exploration of a complex issue, acknowledging both the power and perils of the online world in shaping how we express ourselves and interact with each other.

Chapter four examines South African laws and regulations that deal with freedom of expression and hate speech, such as the Promotion of Equality and Prevention of Unfair Discrimination Act, Film and Publications Act, Film and Publications Amendment Act, Cybercrime Act, and the Hate Speech Bill. It compares these laws

© University of Pretoria

and regulations to international principles on hate speech and assesses how effective they are in addressing hate speech on social media.

The concluding chapter of the essay outlines the study's key findings and conclusions. A proposal of recommendations is discussed and provided in this chapter.

1.9 Summation

As we proceed to the next chapter, it is crucial to recognise that the initial defence employed by those accused of social media-based hate speech is the assertion of their right to freedom of expression, encapsulated by the phrase, ‘I can say what I want.’ However, this defence is not absolute, as it is important to examine the legal definitions and boundaries of where freedom of expression and hate speech originate to understand and discern what constitutes freedom of expression in juxtaposition to what signifies hate speech.

© University of Pretoria

CHAPTER 2: CONCEPTUAL FRAMEWORK

2.1 Introduction

The first chapter sets out the research proposal and questions that should be answered at the end of this essay. This chapter focuses on the conceptual framework of the freedom of expression and hate speech, taking into account their regulatory parameters on social media on a global scale as well as how they are viewed by different global entities.

2.2 Freedom of Expression

Freedom in law refers to the state of being free to enjoy social, political, and economic rights and privileges.[23] It can also be defined as the ability to think, speak, or act without hindrance, interference, or restraint from a dictatorial government. This means that freedom is the ability of an individual to enjoy their fundamental human rights without unjustified or unreasonable limitations by the state or any other entity with the power to impose limitations.[24]

Freedom of expression is a basic human right that is protected by international law. This is because the principles of equality, dignity and personal growth require that people be free to express themselves without fear of punishment or discrimination.[25] Any form of human expression, particularly in speech, communicates the essence of their humanity and the need for this revelatory mode of communication to be respected and protected. Article 19 of the UDHR provides that: “Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.”

Article 19 of the ICCPR expands on several of the provisions of the UDHR in a comparable manner.

---

23 JC Kanamugire & OB Faku “Employees' Right to Freedom of Expression Through Social Media in South Africa” (2022) 11 *Perspectives in Law and Public Administration* 394.

24 *Ibid 23.*

25 Limpitlaw, J, *“Media Law Handbook for Southern Africa”* (2021) Communication Library 6.

© University of Pretoria

2.1.1 Freedom of Expression on Social Media

Technology is neutral, but those who use it are not.[26]

-

The international human rights laws we have traversed indicate that the right to freedom of expression means more than just being able to say what you want. These rights also include the ability of being able to express yourself through a variety of media, such as print, broadcast and online. Therefore, this connects the right of individuals to the medium through which they can express themselves and share information.

For the purposes of this essay, the *Merriam-Webster Dictionary* defines social media as a form of electronic communication that allows users to create online communities to share information, ideas, personal messages, and other content such as photographs and videos.[27] This definition includes sites such as Facebook, Twitter, Instagram, YouTube, blogs and forums.

Social media has played a significant role in enabling the freedom of expression, citizen journalism, and the exchange of ideas which has inadvertently had an influence on how people communicate. Social media has energised the voice of those who were previously marginalised or silenced and has also enabled the sharing of views and opinions across vast communities, the organising of protests at a multi-regional if not on a global scale, and the formation of groups for discussing and assembling agendas in total exercising their freedom of speech.[28] This can be seen in the hashtag and slogan campaign where in 2017, the Hollywood actress Alyssa Milano started the #MeToo movement on Twitter to encourage women to share their experiences of

---

26 Watts, J (2012b) Is technology neutral? Helwys Society Forum. Available at https://www.helwyssocietyforum.com/is-technology-neutral/ (accessed 13/12/2024).

27 Merriam-Webster, 'Social Media' (Merriam-Webster, n.d.), available at https://www.merriam-webster.com/dictionary/social%20media (accessed 16/09/2023).

28 C Dzoguji "The Great Thing About Social Media Was How It Gave a Voice to Voiceless People" (2019) NUHA Foundation available at https://nuhafoundation.org/home/blog/bloggingentries/2019/youth/the-great-thing-about-social-media-was-how-it-gave-a-voice-to-voiceless-people-jon-ronson-2/ (accessed 25/10/2023).

© University of Pretoria

sexual harassment and assault.[29] The movement quickly went viral and sparked a global conversation about sexual violence.[30] Another relevant example is how the Black Lives Matter movement has used social media to raise awareness of police brutality and racial injustice.[31]

Citizen journalism enabled by social media, by allowing people to report on events not covered by traditional media outlets.[32] This has increased the diversity of voices in the media and has helped to hold those in power accountable.[33] The Arab Spring protests of 2011 and 2012 in North Africa, led by young people against corrupt governments, are examples of how social media, the internet and blogs can be used to organise and mobilise people in ways that were not possible before.[34] In the past, young people were often barred from official state media, but social media gave them a platform to make their voices heard.

The exchange of ideas and information facilitated by social media platforms on a global scale has enabled people to connect with others who share similar interests to further discuss and debate these similarities for their shared benefit.[35] However, it is important to note that this argument is not unilaterally focused, but has two sides because as much as social media can be a powerful tool for free speech, it can also be used to spread hate speech, misinformation, and harassment.

---

29 C Ferrière "#metoo and Twitter: The Feminist Movement on Social Media" (2022) Humans R social media, Winter 2022 Open Textbook Edition) available at https://opentextbooks.library.arizona.edu/hrsmwinter2022/chapter/special-topics-me-too/ (accessed 25/10/2023).

30 *Ibid 29.*

31 Black Lives Matter, 'About' (Black Lives Matter, n.d.), available at https://blacklivesmatter.com/about/ (accessed 25/10/2023).

32 WI O'Byrne "Educate, Empower, Advocate: Amplifying Marginalized Voices in a Digital Society" (2019) CITE Journal, Volume 19, Issue 4 available at https://citejournal.org/volume-19/issue-4-19/english-language-arts/educate-empower-advocate-amplifying-marginalized-voices-in-a-digital-society/ (accessed 25/10/2023).

33 Legal Desire "Role of Social Media and Freedom of Speech and Expression" (2021) available at https://legaldesire.com/role-of-social-media-and-freedom-of-speech-and-expression/ (accessed 18/09/2023).

34 W Gumede "Rise in Censorship of The Internet and Social Media in Africa" (2016) Journal of African Media Studies vol 8, 413-421.

35 *Ibid 33.*

© University of Pretoria

2.3 Definition and Conceptualisation of Hate Speech

2.3.1 Overview

Defining hate speech is a challenge, as it has no universally accepted definition[36] which makes it difficult to legislate against hate speech - as laws must be clear and specific. There is no consensus on what constitutes hate speech, and different countries have different laws and regulations governing it.[37]

Several factors influence the complex activity of defining hate speech. First, hate speech can be expressed in many ways which include works, images, and symbols.[38] Second, hate speech can be perceived as a language that attacks or threatens a person or group based on their religion, ethnicity, gender, sexual orientation or race.[39] Third, a form of discrimination that seeks to marginalise and dehumanise its targets can be viewed as hate speech.[40] Examples include using a racial slur against someone; making threats of violence against a group of people, spreading false information about a minority group and denying the existence of a minority group.[41]

The 2020 UN Strategy and Plan of Action on Hate Speech identifies three types of lawful and unlawful expression. The most serious type is advocacy of hatred against a person or group based on their race, religion, or nationality, which is likely to cause discrimination, hostility, or violence.[42] The United Nations Human Rights Council (UNHRC) and General Assembly have affirmed that the freedoms people enjoy offline

---

36 B Di Fátima "Hate Speech on Social Media: A Global Approach" (2023) Communication Library vol 7 available at https://store.lexisnexis.co.za/products/hate-speech-on-social-media-a-global-approach-skuZASKUPG00</a (accessed 9/10/2023) preface 11.

37 Alkiviadou, N, *Global Handbook on Hate Speech Laws* (The Future of Free Speech, 20 November 2020), available at https://futurefreespeech.com/global-handbook-on-hate-speech-laws/ (accessed 25/10/ 2023) 3.

38 *Ibid at page 281 and 282.*

39 Alkiviadou, N, *Global Handbook on Hate Speech Laws op cit at 37.*

40 Speech, "Hate Speech and Social Media: Preventing Atrocities and Protecting Human Rights Online" (2020) *Global Centre for the Responsibility to Protect* available at https://www.globalr2p.org/publications/hate-speech-and-social-media-preventing-atrocities-and-protecting-human-rights-online/ (accessed 6/09/2023).

41 Geldenhuys, J, 'Hate Speech and Racist Slurs in the South African Context' (2020) in *SciELO SA*] (accessed 28/04/2024).

42 United Nations Office on Genocide Prevention and the Responsibility to Protect, *Strategy and Plan of Action on Hate Speech: Detailed Guidance on Implementation for United Nations Field Presences* (United Nations, 2020), accessed 27 September 2023. <URL un strategy and plan of action on hate speech ON United Nations Digital Library digitallibrary.un.org> 16.

© University of Pretoria

also apply online meaning that the right to freedom of expression applies to all forms of communication, including the internet.[43]

The International Covenant on Civil and Political Rights (ICCPR), which guarantees the right to freedom of expression, also allows for restrictions on this right 'for respect of the rights or reputations of others' or "for the protection of national security or public order, or of public health or morals." Article 20(2) of the ICCPR specifically prohibits advocacy of national, racial, or religious hatred that constitutes incitement to discrimination, hostility, or violence.[44]

Article 4(a) of the Convention on the Elimination of All Forms of Racial Discrimination (CERD), prohibits 'all propaganda and all organisations which are based on ideas or theories of superiority of one race or group of persons of one colour or ethnic origin, or which attempt to justify or promote racial discrimination in any way.'[45] The International Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW), prohibits "all forms of discrimination against women," including "incitement to hatred against women."[46]

Hate speech can be subtle and indirect, making it difficult to identify.[47] The concept of hate speech raises many difficult questions such as what does the 'hate' in hate speech refer to, can hate speech be directed at dominant groups, or is it targeted at

43 P Martins & V Ferrari "Human Rights Online at The Human Rights Council 47th Session" (2021) Association for Progressive Communications available at https://www.apc.org/en/news/human-rights-online-human-rights-council-47th-session (accessed 25/10/2023).

44 Office of the High Commissioner for Human Rights "Towards an Interpretation of Article 20 Of The ICCPR: Thresholds for The Prohibition of Incitement to Hatred" (2010) available at https://www.ohchr.org/Documents/Issues/Expression/ICCPR/Santiago/TobyMendel.pdf (accessed 4/09/2023).

45 United Nations, *International Convention on the Elimination of All Forms of Racial Discrimination* (adopted and opened for signature at New York on 21 December 1965, entered into force 4 January 1969), available at https://legal.un.org/avl/ha/cerd/cerd.html (accessed 25/10/2023).

46 United Nations, *Convention on the Elimination of All Forms of Discrimination against Women* (adopted by the General Assembly 18 December 1979, entered into force 3 September 1981), available at https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-elimination-all-forms-discrimination-against-women (accessed 25/10/2023).

47 N Ocampo, E Sviridova, E Cabrio & S Villata "An In-depth Analysis of Implicit and Subtle Hate Speech Messages" (2023) Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics) available at https://aclanthology.org/2023.eacl-main.147 (accessed 25/10/2023).

© University of Pretoria

oppressed or marginalised communities, what is the harm or harms of hate speech, and what can or should be done to counteract hate speech.[48]

Hate speech can be protected by freedom of speech laws, which guarantee the right to express unpopular or offensive opinions. This can be seen in the UDHR, proclaimed in 1948, which is intended to ensure that all human beings have the right to freedom of speech and belief, and freedom from fear. [49]

The preamble of the UDHR states that these rights are "the foundation of justice and peace in the world" and that they should be protected by the rule of law. The UDHR also states that the right to freedom of expression is "the foundation of a free and democratic society" and that it is essential for the "full development of the personality of each human being."[50] The preamble of the UDHR goes on to say that "the nations of the world agree" that the right to freedom of expression is "the core of a democracy and individual freedom."

In other words, the UDHR recognises that the right to freedom of expression is essential for the protection of all other human rights. It is a fundamental right that allows people to participate in the political process, to hold their governments accountable, and to express their opinions freely. The UDHR, which guarantees the right to freedom of expression, but also states that this right "carries with it duties and responsibilities" and "may therefore be subject to certain restrictions."[51]

The African Union (AU) recognises that freedom of expression is an important right, but it must be exercised responsibly and with respect for the rights of others.[52] The AU has adopted the African Charter on Human and Peoples' Rights, which guarantees the right to freedom of expression, but also allows for certain limitations to this right, such as to protect the rights or reputations of others, national security, public order, or

---

48 S Benesch “Dangerous Speech” in C Strippel, S Paasch-Colberg, M Emmer & J Trebbe (eds) *Challenges and Perspectives of Hate Speech Research* (2023) Digital Communication Research available at https://doi.org/10.48541/dcr.v12.11 (accessed 11/09/2023).

49 Universal Declaration of Human Rights (UDHR), adopted 10 December 1948, UN GA Res 217A (III), https://www.un.org/en/about-us/universal-declaration-of-human-rights (28/04/2014)

50 *Ibid 49.*

51 Universal Declaration of Human Rights (UDHR), adopted 10 December *op cit.*

52 African Union, *African Charter on Human and Peoples' Rights* (adopted 27 June 1981, entered into force 21 October 1986), available at https://au.int/sites/default/files/treaties/36390-treaty-0011_-_african_charter_on_human_and_peoples_rights_e.pdf (accessed [28/04/2024]).

© University of Pretoria

public health or morals.[53] Some scholars have taken a human rights approach to the issue of free speech on social media. They argue that implementation of international freedom of expression standards is the best way to assess the boundaries of acceptable speech, as they transcend national laws.[54]

2.3.2 Hate Speech Definition by Social Media Platforms

Social media sites often define hate speech in their own terms of service.[55] Facebook's Community Standards prohibit hate speech, which is defined as a direct attack on people based on their protected characteristics, such as race, ethnicity, religion, sexual orientation and gender.[56] Facebook also prohibits content that promotes violence or dehumanises individuals or groups based on their protected characteristics.[57] Educational, documentary, scientific and artistic content that includes hate speech must provide context in the video itself, not just in the title or description.[58]

YouTube does not allow hate speech, which is content that promotes violence or hatred against people based on their age, caste, disability, ethnicity, gender identity and expression, nationality, race, immigration status, religion, sex/gender, sexual orientation, or veteran status.[59]

According to Twitter's terms of service, it is prohibited to post content that attacks or threatens people based on their race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religion, age, disability or disease.[60] This includes

---

53 Office of the High Commissioner for Human Rights "Training Manual on Human Rights Monitoring (2001) available at https://www.ohchr.org/sites/default/files/Reporting-ICCPR-Training-Guide.pdf (accessed 4/09/2023).

54 Z Takhshid "Regulating Social Media in the Global South" (2021) 24 *Vanderbilt Journal of Entertainment & Technology Law*, Volume 24, Issue 1(2022). U Denver Legal Studies Research Paper No. 21-18. Available *at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3836986*

55 N Alkiviadou "Hate Speech on Social Media Networks: Towards a Regulatory Framework?" (2019) 28(1) Information & Communications Technology Law 19.

56 Facebook "Hate Speech" (n.d.) available at https://transparency.meta.com/policies/community-standards/hate-speech.(accessed 10/10/2023).

57 *Ibid 56*.

58 Facebook "Hate Speech *op cit.*

59 Google "Hate Speech Policy" (n.d.) available at https://support.google.com/contributionpolicy/answer/11412392?hl=en (accessed 10/10/2023).

60 Twitter "Hateful Conduct Policy" (n.d.) available at https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy (accessed 10/10/2023).

© University of Pretoria

content that promotes violence against or threatens other people based on these categories.[61]

TikTok does not allow hate speech, which is defined as content that attacks, threatens, incites violence against, or dehumanises people based on their race, ethnicity, national origin, religion, caste, sexual orientation, sex, gender, gender identity, serious disease, disability or immigration status.[62]

2.4 Social Media Regulation in Global Context

The multidimensional nature of social media platforms has made it difficult for lawmakers and policymakers to agree on the best way to regulate them.[63] Social media law is not a single, unified set of rules. Instead, it is a mixture of laws from different areas, including constitutional law, employment law, criminal law, administrative law, intellectual property law, tort law, and even the rules of discovery.[64] These laws have been adapted to apply to the widespread and relatively new use of online and mobile platforms for sharing and creating content.[65]

The key issues that are being debated in the global context of social media regulation includes content moderation, transparency, accountability, and competition. According to the International Telecommunication Union (ITU), the estimated number of internet users worldwide as of 2022 was 5.3 billion, representing approximately 66% of the global population.[66]

On the surface, the fact that content reaches a wide audience is not necessarily a problem. However, the type of content that is being shared is often the real issue.[67] As a result, any regulation of social media platforms would need to balance the interests of different stakeholders, such as users, businesses, and governments. In recent

---

61 *Ibid 60.*

62 TikTok "Community Guidelines'" (n.d.) available at https://www.tiktok.com/safety/community-guidelines (accessed 10/10/2023).

63 Z Takhshid "Regulating Social Media in the Global South" (2021) 24 *Vanderbilt Journal of Entertainment & Technology Law* 1.

64 CE Hart "Social Media Law: Significant Developments" (2016) 72 Business Law 235.

65 *Ibid 64*.

66 See Statista.com "Number of Internet Users Worldwide 2022" (2023) available at https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/ (accessed 27/09/2023).

67 I Rutenberg & A Sugow "Regulation of the Social Media in Electoral Democracies: A Case of Kenya" (2020) 7 SOAS University of London Law Journal 301.

© University of Pretoria

years, governments around the world have put increasing pressure on providers of internet communication services to monitor and control speech and privacy online.[68] The pace of technological development is at times faster than the pace of legal change. As a result, existing laws that govern offline conduct are often applied to illegal online behaviour.[69]

The difficulty of governments regulating social media has been attributed to the internet's dual nature as both a global commons and a privately owned and operated space.[70] Social media platforms have been criticised for their poor efforts to moderate content in many instances around the world.[71] For example, they have been accused of failing to protect users from foreign influence during elections in the United States and France, and of not adequately moderating hate speech in Ethiopia that incited violent ethnic protests.[72]

Tech companies and social media platforms have responded to online hate speech in a haphazard and inconsistent manner.[73] However, it is nearly impossible to effectively monitor social media and protect against incitement on an *ad hoc* basis. The sheer number of platforms, groups, and online spaces on social media makes the task even more difficult.[74] Social media companies are not legally required under international law to suppress hate speech and incitement to violence on their platforms.[75] However,

68 H Bloch-Wehba "Global Platform Governance: Private Power in the Shadow of the State" (2019) 72 Southern Methodist University Law Review 27.

69 Rutenberg & Sugow (2020) 301.

70 *Ibid 69*.

71 L Garbe, LM Selvik & P Lemaire "How African Countries Respond to Fake News and Hate Speech" (2023) 26(1) Information, Communication & Society 86.

72 J Samples "Why the Government Should Not Regulate Content Moderation of Social Media' (Cato Institute 9 April 2019) available at https://www.cato.org/policy-analysis/why-government-should-not-regulate-content-moderation-social-media (accessed 25/10/2023).

73 RA Wilson & MK Land "Hate Speech on Social Media: Content Moderation in Context" (2021) 52 Connecticut Law Review 1029 available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3690616 (accessed 25/10/2023).

74 Speech "Hate Speech and Social Media: Preventing Atrocities and Protecting Human Rights Online" (2020) *Global Centre for the Responsibility to Protect* available at https://www.globalr2p.org/publications/hate-speech-and-social-media-preventing-atrocities-and-protecting-human-rights-online/ (accessed 6/09/2023).

75 J Samples "International Law and "Hate Speech" (July 2020) available at https://www.cato.org/blog/international-law-hate-speech-online (accessed 25/10/2023)

© University of Pretoria

the UN Guiding Principles on Business and Human Rights (UNGPs)[76] require social media companies to respect human rights.

This means that social media companies should take steps to identify and address the human rights impact of their activities, including the spread of hate speech and incitement to violence.[77] Governments can also pass legislation in their territories to require social media companies to follow human rights rules when it comes to hate speech.[78] Unfortunately, many social media companies do not invest enough resources in filtering or blocking hate speech. Instead, they rely on a combination of artificial intelligence (AI), user reporting, and 'content moderators' to monitor content and decide whether it is appropriate or not.[79] Content moderators are often not familiar with the specific political and historical context of situations where hate speech is prevalent.[80]

The challenges of combating hate speech are further compounded by the fact that content on social media is often posted in local languages.[81] However, most major tech companies are based in the United States,[82] and their content moderators are typically English-speaking. This means that they may not understand the danger and

---

76 United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations 'Protect, Respect and Remedy' Framework* (United Nations, 2012), available at https://www.ohchr.org/en/business-and-human-rights (accessed 25/10/2023).

77 B Paulos, SÇelik "The Challenges of Regulating Hate Speech on Social Media in Light of the Theory of Freedom of Expression" (2021) Süleyman Demirel Üniversitesi Hukuk Fakültesi Dergisi https://dergipark.org.tr/en/pub/sduhfd/issue/63393/922588 (accessed 25/10/2023) 110,113,121 - 123.

78 *Ibid 77.*

79 BK. Wiederhold, 'Femtech: Digital Help for Women's Health Care Across the Life Span' (Cyberpsychology, Behaviour, and Social Networking Vol. 24, No. 11, 10 November 2021) https://www.liebertpub.com/doi/full/10.1089/cyber.2021.29230.editorial (accessed 25/10/2023); Geoffrey A. Fowler and Chris Alcantara, 'Gatekeepers: These tech firms control what's allowed online' (The Washington Post 24 March 2021) https://www.washingtonpost.com/technology/2021/03/24/online-moderation-tech-stack/ (accessed 25/10/2023).

80 'Hate Speech and Social Media: Preventing Atrocities and Protecting Human Rights Online' (*Global Centre for the Responsibility to Protect*, 21 February 2020) https://www.globalr2p.org/publications/hate-speech-and-social-media-preventing-atrocities-and-protecting-human-rights-online/ (accessed 6/09/2023).

81 Kovács, György, Alonso, Pedro, and Saini, Rajkumar, 'Challenges of Hate Speech Detection in Social Media' (2021) 13 SN Computer Science 1 [online] 13 February. Available. (accessed 25/10/2023) 95.

82 Disfold Data (2023), *Top 1000 Largest US Technology Companies 2024* [online] Available at: https://disfold.com/united-states/sector/technology/companies/ (accessed 25/10/2023).

© University of Pretoria

harmfulness of certain content, especially if it is in another language or is culturally specific.[83]

2.5 Summation

This chapter has defined freedom of expression and hate speech by identifying the legal landscape for these concepts in addition to spotlighting the complex issue of regulation. The next chapter acknowledges the existence of both proponents and opponents of regulating these concepts and will explore these contrasting viewpoints. It will also draw on how hate speech as a concept can fit into freedom of expression as well as identifying the challenges that are posed by the current legal provisions on hate speech.

83 Global Centre for the Responsibility to Protect. 'Hate Speech and Social Media: Preventing Atrocities and Protecting Human Rights Online' (Speech delivered by Dr Simon Adams on 16 February 2020) (2020) Global Centre for the Responsibility to Protect available at: https://www.globalr2p.org/publications/hate-speech-and-social-media-preventing-atrocities-and-protecting-human-rights-online/ (accessed 06/09/2023)

© University of Pretoria

CHAPTER 3: HATE SPEECH AND ITS IMPLICATIONS

3.1 To Regulate or Not to Regulate

Social media hate speech is more complex and varied, it spreads rapidly and can have real-world consequences beyond its borders,[84]resulting in the dissemination of surface level opinions that lack an in-depth analysis on a wide range of issues.[85] It is ubiquitous, interactive, and has different media channels, available 24/7 to a much larger audience.[86] The works of Papadopoulos and Snail demonstrate the inherent difficulty in identifying and prosecuting hate speech due to challenges presented by finitely tracing the diverse online sources.[8788] On social media, haters can be anonymous and find support from others with the same aggressive mindset. Most recently, social media has been weaponised to target the 'other' and incite violence.[89]

3.2 Types of Arguments for Regulating Hate Speech on Social Media

There are three main types of arguments for and against regulating hate speech on social media which include liberal arguments (post-structuralist), progressivist arguments and multiculturalism debates,[90] however, they are not the only arguments that can be made, and there are other perspectives on this issue. Arguments for and against regulating hate speech on social media are not mutually exclusive.[91] The core of the debate is finding a balance between protecting free expression and regulating hate speech. These arguments represent different viewpoints on the issue of regulating hate speech on social media.

---

84 Global Handbook on Hate Speech Laws' (The Future of Free Speech, November 2020) https://futurefreespeech.com/global-handbook-on-hate-speech-laws/ (accessed 25/10/2023)

85 S Papadopoulos and Snail S Cyberlaw @ SA IV: The Law of the Internet in South Africa9 4TH edn, Van Schaik Publishers 2022) 11.1

86 B Di Fátima, 'Hate Speech on Social Media: A Global Approach' (2023) Communication Library 7.

87 S Papadopoulos and Snail S Cyberlaw *op cit* 10.1

88 Ibid 87.

89 Hate Speech and social media: Preventing Atrocities and Protecting Human Rights Online' (Global Centre for the Responsibility to Protect, February 2020) https://www.globalr2p.org/publications/hate-speech-and-social-media-preventing-atrocities-and-protecting-human-rights-online/ (accessed 06/09/2023).

90 B Paulos, 'The Challenges of Regulating Hate Speech on Social Media in Light of the Theory of Freedom of Expression' (2021) 11 *Süleyman Demirel Üniversitesi Hukuk Fakültesi Dergisi* 97.

91 Hate Speech and social media *op cit.*

© University of Pretoria

The central issue under consideration is striking the appropriate balance between safeguarding the fundamental right to freedom of speech and implementing the necessary regulations on hateful, insightful, and discriminatory rhetoric that is disseminated through social media platforms. The complexity of this discourse is evidenced by the different perspectives and interests presented.[92]

Liberal arguments against regulating hate speech are based on the principle of free speech.[93] They argue that free speech is essential for a democratic society, and that any restrictions on free speech should be narrowly tailored to achieve a compelling government interest.[94] Liberals also argue that it is difficult to define hate speech without censoring legitimate speech.[95]

Progressivist arguments for regulating hate speech are based on the principle of equality and social justice.[96] They hold that hate speech can cause serious harm to individuals and minority groups, and that regulating hate speech can help to protect these groups from harm.[97] Progressivists also argue that hate speech can perpetuate harmful stereotypes and prejudices, and that regulating hate speech can help to promote equality and social justice by challenging these stereotypes and prejudices.[98]

Multiculturalist debates on hate speech are concerned with the tension between the right to free speech and the right of minority groups to be free from discrimination and violence.[99] Multiculturalists argue that the harm caused by hate speech is often disproportionately borne by minority groups, and that regulating hate speech is

---

92 V Guillén-Nieto *Hate Speech: Linguistic Perspectives* (De Gruyter Inc 2023) 123 https://ebookcentral-proquest-com.uplib.idm.oclc.org/lib/pretoria-ebooks/detail.action?docID=7177394 accessed 10 July 2025.

93 A Reid, 'Does Regulating Hate Speech Undermine Democratic Legitimacy? A Cautious "No"' (*Res Publica*, 6 May 2019) https://link.springer.com/article/10.1007/s11158-019-09431-6 (accessed 25/10/2023).

94 J Maclure, 'The Regulation of Hateful and Hurtful Speech: Liberalism's Uncomfortable Predicament' (2017) 63 McGill Law Journal 133 [online] McGill University https://lawjournal.mcgill.ca/article/the-regulation-of-hateful-and-hurtful-speech-liberalisms-uncomfortable-predicament/ (accessed 25/10/2023).

95 *Ibid*.

96 C Demaske, 'Hate Speech' (The Free Speech Center, last updated 30 July 2023) [online] Middle Tennessee State University https://firstamendment.mtsu.edu/article/hate-speech/ (accessed 30/10/2023).

97 *Ibid* and An Altman, 'Liberalism and Campus Hate Speech: A Philosophical Examination' (1993) 103 (2) Ethics 344 [online] JSTOR (accessed 25/10/2023).

98 A Reid, *op cit*.

99 D Chong, 'Free Speech and Multiculturalism in and out of the Academy' (2006) 27(1) Political Psychology 1 [online] JSTOR (accessed 25/10/2023).

© University of Pretoria

necessary to protect these groups.[100] However, multiculturalists also recognise the importance of free speech, and they argue that any restrictions on free speech should be carefully tailored to avoid stifling legitimate debate.[101]

3.3 Impact of Social Media on Hate Speech Propagation

Hate speech on social media can have a global impact, crossing borders and cultures. A message posted in one country can be seen by users all over the world. The propagation of hate speech has taken many forms throughout history as media platforms have evolved. The Nazis used virulently antisemitic newspapers like 'Der Stürmer' to help incite the German people into active persecution of Jews.[102] *Der Sturmer* publisher Julius Streicher was later found guilty of inciting genocide during the Nuremburg Tribunal and sentenced to death.[103]

Five decades later, during the 1994 genocide in Rwanda, the RTLM radio station played a key role promoting hate speech regarding the Tutsi,[104] inciting the genocide, and even directing the killers with the names of potential victims who had not yet been murdered.[105] Hate speech often relies on distorted or false information about targeted groups, and social media can help this misinformation spread quickly, making it difficult to stop.

For example, misinformation about COVID-19 has led to vaccine hesitancy.[106] Political bias has also been a very relatable topic when it concerns the negative aspects of social media as companies that own these platforms use them to promote certain

---

100 *Ibid*.

101 D Chong *op cit*.

102 Der Stürmer | German newspaper | Britannica" (Encyclopaedia Britannica, 2023) https://www.britannica.com/topic/Der-Sturmer (accessed 28/10/2023) 'Der Stürmer Antisemitic Newspaper' (Montreal Holocaust Museum) available at https://museeholocauste.ca/en/objects/der-sturmer-antisemitic-newspaper/ (accessed 30/11/2023).

103 D Milo and P Stein a Practical Guide to Media Law (2ed, Lexis Nexis, 2013) 164.

104 Author (n.d.), 'Rwanda and RTLM Radio Media Effects - United States Holocaust Memorial Museum' (United States Holocaust Memorial Museum), available at [https://www.ushmm.org/m/pdfs/20100423-atrauss-rtlm-radio-hate.pdf] (accessed 25/10/2023 – 05/02/2024).

105 Author "Propaganda vs. Education: A Case Study of Hate Radio in Rwanda" (Harvard Evidence for Policy Design, 2023) available at https://epod.cid.harvard.edu/sites/default/files/2018-02/propaganda_vs._education_a_case_study_of_hate_radio_in_rwanda.pdf (accessed 25 - 30/11/2023).

106 Mejía "Misinformation and Vaccine Hesitancy" (Julian Samora Research Institute, Michigan State University, Fall (2021) available at https://jsri.msu.edu/publications/nexo/vol-xxv/no-1-fall-2021/misinformation-and-vaccine-hesitancy (accessed 28/11/ 2023).

© University of Pretoria

viewpoints and suppress others.[107] This can harm democracy by making it difficult for people to access diverse information and viewpoints.

Social media can also be used to organise campaigns of harassment, doxing and other forms of online abuse against specific individuals or groups which directly impacts individuals' self-esteem and mental wellbeing, with many reporting high rates of anxiety.[108] These platforms have been noted to harass and target women, LGBTQI+ people, and people of colour.[109]

Extremist groups and ideologies can use social media to recruit, radicalise, and mobilise followers, often by exploiting the platforms' algorithms to gain visibility which can magnify discord and echo changes in the political climate.[110] These platforms also tend to show people content that aligns with their existing beliefs, which can make it difficult for people to see other perspectives. An example of this can be seen in the United States, where white supremacists who have committed recent attacks have connected with racist communities online and used social media to promote violence.[111]

Prosecutors say that the Charleston church shooter, who killed nine black clergy and worshippers in June 2015, became radicalised online by white supremacist ideology.[112] Additionally, users may feel less detached from their online actions, which can lead to a perception that there are fewer consequences for engaging in hate speech.[113] Online anonymity enables individuals to engage in toxic behaviour without

---

107 Law and Internet Foundation "The Negative Impact of Social Media on Politics" Law and Internet Foundation (2022) available at https://www.netlaw.bg/en/a/the-negative-impact-of-social-media-on-politics (accessed 02/11/2023).

108 Jones, S, 'Internet Matters: What Is the Real-World Impact of Online Hate Speech on Young People?' 2024 available at https://www.internetmatters.org/hub/question/what-is-the-real-world-impact-of-online-hate-speech-on-young-people/ (accessed 26/09/2023).

109 Boston "Unfriendly Content: The Rise of Hate Speech on Social Media" (accessed 26/09/2023).

110 Council on Foreign Relations "Hate Speech on Social Media: Global Comparisons" (accessed 26/09/2023).

111 Bjork-James "Fuelled by Virtually Unrestricted Social Media Access, White Nationalism Is on The Rise and Attracting Violent Young White Men" (The Conversation, 02/08/2022) available at https://theconversation.com/fueled-by-virtually-unrestricted-social-media-access-white-nationalism-is-on-the-rise-and-attracting-violent-young-white-men-186896 (accessed 25/10/2023).

112 Council on Foreign Relations "Hate Speech on Social Media: Global Comparisons" (accessed 26/09/2023).

113 Clean Up The Internet (2023), 'Academic Research About Anonymity, Online Disinhibition, Trolling And Abuse', available at https://www.cleanuptheinternet.org.uk/post/some-useful-scholarly-articles-about-online-disinhibition-anonymity-and-online-harms (accessed 25/10/2023).

© University of Pretoria

fear of being identified.[114] While online anonymity can protect whistle-blowers, political dissidents and targets seeking safety from their abusers and harassers, it can also enable online hate, harassment and misinformation.[115] There is also the 'priority' by these platforms to include controversial content, including hate speech, as it keeps people engaged.[116]

3.4 Challenges in Monitoring Hate Speech Effectively

There are privacy and security concerns associated with monitoring hate speech online.[117] It is important to balance the need to protect users from harm with the right to privacy. Because hate speech can be expressed in a variety of forms, including text, images, videos, and memes[118] makes it difficult to detect automatically, as there is also a vast amount of content generated and shared online, which increases the difficulty to monitor it for hate speech.[119] The constant evolution of social media platforms and new trends emerging all the time,[120] make it difficult to keep up with the latest tactics used to disguise hate speech as legitimate speech, making the two forms of speech indistinguishable.

Hate speech can be used to express political opinions or social commentary which complicates determining whether speech is hateful or simply offensive or controversial. Examples of hate speech include statements made by political figures that categorise immigrants as a threat to national security. Such statements could be

114 Ferrer "The Dangerous Shield Of Online Anonymity And Dealing With Toxic Behaviour Online" (The Ionian) 24/02/2023) available at https://theionian.org/11417/opinion/the-dangerous-shield-of-online-anonymity-and-dealing-with-toxic-behavior-online/ (accessed 25/10/2023).

115 eSafety Commissioner (2023), 'Anonymity And Identity Shielding', available at [https://www.esafety.gov.au/industry/tech-trends-and-challenges/anonymity] (accessed 25/10/2023).

116 MediaSmarts (2023), 'Impact of Online Hate', available at [https://mediasmarts.ca/online-hate/impact-online-hate] (accessed 25/10/2023)

117 International Association of Privacy Professionals (2023), 'In Curtailing Hate Speech Online, Will Privacy Sometimes Have to Take a Back?', available at [https://iapp.org/news/a/in-curtailing-hate-speech-online-will-privacy-sometimes-have-to-take-a-back/] (accessed 26/10/2023).

118 Fátima *Hate Speech on social media: A Global Approach* (Communication Library vol 7) (2023) available at https://store.lexisnexis.co.za/products/hate-speech-on-social-media-a-global-approach-skuZASKUPG00 (accessed 09/10/2023) 76,144 and 160.

119 Council on Foreign Relations (2023), 'Hate Speech on Social Media: Global Comparisons', available at [https://www.cfr.org/backgrounder/hate-speech-social-media-global-comparisons] (accessed 26/10/2023).

120 Maryville University, (2020), 'The Evolution of Social Media: How Did It Begin, and Where Could It Go Next?', available at [https://online.maryville.edu/blog/evolution-social-media/] (accessed 26/10/2023).

© University of Pretoria

considered as hate speech if they are intended to incite hatred or violence against immigrants. Similarly, a comedian making a joke about a specific racial or religious group could be considered hate speech if the intention is to dehumanise or belittle that group. However, it could also be considered legitimate comedy if it is simply intended to be funny and does not promote hatred or violence. There are complex legal and ethical questions surrounding hate speech monitoring. Different countries have different laws and cultural norms regarding hate speech.[121]

3.5 Tension Between Freedom of Expression and Hate Speech

It is difficult to define hate speech in a way that is both comprehensive and respectful of freedom of expression.[122] Some countries have extremely strict laws against hate speech, and in some cases, these laws are too strict. For example, the German government passed a law to punish online platforms for not removing illegal hate speech[123] and this law allows the government to fine platforms up to 50 million euros if they do not remove illegal content within 24 hours of being notified about it.[124]

Hate speech can be subtle and implicit, and it can vary depending on the cultural context.[125] The tension between freedom of expression and hate speech emerges when the right to free expression clashes with the boundaries set to contain this right from transgressing into hate speech. As much as it is important to protect the right of individuals to engage in robust public debate, even if their views are controversial, hate speech can have a negative impact by creating a hostile environment and potentially leading to violence or discrimination against marginalised groups.[126]

---

121 The Future of Free Speech (2020), *Global Handbook on Hate Speech Laws*, available at https://futurefreespeech.com/global-handbook-on-hate-speech-laws/ (accessed 25 - 30/10/2023).

122 Research Outreach, 'Hate Speech Regulation on Social Media: An Intractable Contemporary Challenge' (accessed 26/09/2023).

123 B Paulos, 'The Challenges of Regulating Hate Speech on Social Media in Light of the Theory of Freedom of Expression' (2021) 11 *Süleyman Demirel Üniversitesi Hukuk Fakültesi Dergisi* 97.

124 J Braun, 'Germany's balancing act: Fighting online hate while protecting free speech' (POLITICO, 1 October 2020) [online] [[invalid URL removed]) (accessed 26/10/ 2023).

125 Hate Speech, a FAQ' (DW Akademie, 9 March 2016) [online] https://akademie.dw.com/en/hate-speech-a-faq/a-19103744 [invalid URL removed] (accessed 26 -30/10/2023).

126 United Nations, 'International Human Rights Law' [online] United Nations Office on Genocide Prevention and the Responsibility to Protect available at https://www.un.org/en/hate-speech/united-nations-and-hate-speech/international-human-rights-law (accessed 26/10/2023).

© University of Pretoria

There is no easy answer to the question of how to balance freedom of expression and hate speech. Different societies have different approaches, and the debate is ongoing. Some countries have laws that make it illegal to say or write things that are hateful or discriminatory,[127] whilst others have laws that protect freedom of speech, even if it means that people can say or write things that are hateful or discriminatory.[128]

3.6 Summation

Social media platforms have become a den for hate speech because anonymity and interactivity of these platforms enable individuals to engage in hate speech with relative impunity, often finding support from individuals who share their ideology. As mentioned above this chapter has explored the contrasting views of those who advocate for and against the regulation of these concepts. It has highlighted how hate speech as a concept fit into freedom of expression, as well as identifying the challenges that are posed by the current legal provisions on hate speech.

127 Hate Speech, a FAQ' (DW Akademie, 9 March 2016) available at https://akademie.dw.com/en/hate-speech-a-faq/a-19103744 (accessed 26/10/2023).

128 *Ibid.*

© University of Pretoria

CHAPTER 4: LEGAL AND REGULATORY FRAMEWORK IN SOUTH AFRICA

4.1 Introduction

The preceding three chapters examine the freedom of expression and hate speech, exploring how they are regulated globally on social media platforms whilst outlining the pros and cons of each concept. Considering South Africa's apartheid history and incorporation of international law, this chapter gives a perspective on moderating the freedom of expression and hate speech while identifying legislation and relevant case laws to better understand how these concepts are decided.

Freedom of expression is a constitutional right protected by section 16 of the South African Constitution.[129] This right includes the freedom to express ideas that may offend or disturb others, as well as the right to receive information and unrestricted ideas.[130] The importance of the right to freedom of expression was clearly stated in the case of South African National Defence Union v Minister of Defence and Another,[131] where the court held that the freedom of expression is essential to democracy. This case acknowledged and applied three foundational bases which are typically associated with the right to freedom: its role in supporting the functioning of a proper of democracy, the contribution it has in the pursuit of truth and lastly the important place it has in the promotion of individual autonomy and self- fulfilment.[132]

The court was also of the opinion that freedom of expression is valuable for many reasons. In *S v Khumalo*[133] it was confirmed that freedom of expression is a fundamental right in our democracy, but it is not the most important right. It must be considered in the context of other rights enshrined in the Constitution, such as human dignity, freedom, and equality. The rights in Chapter 2 of the Constitution are all equally important as there is no hierarchy of rights. Consequently, rights, irrespective of whether they be freedom of expression or dignity are not absolute.[134] The assumption

---

129 S16(1)(a)-(d) the Constitution.

130 South African Human Rights Commission (SAHRC) "Hate Speech Information Sheet" available at Hate Speech Information Sheet- print ready-.pdf (accessed 26/09/2023).

131 *South African National Defence Union v Minister of Defence and Another* 1999 (4) SA 469 CC.

132 D Milo and P Stein a Practical Guide to Media Law (2ed, Lexis Nexis, 2013) 4.

133 *S v Khumalo* [1973] 3 SA 697 (A).

134 S Papadopoulos and Snail S Cyberlaw @ SA IV: The Law of the Internet in South Africa9 4TH edn, Van Schaik Publishers 2022 11.3.1

© University of Pretoria

cannot be made that the right to dignity will always override freedom of expression, or vice versa.[135]

The Constitution's historical, social and political context are important as its foundational values and transformative vision show us how to interpret or understand the constitutionally guaranteed rights.[136] When evaluation of whether a particular provision of media law infringes on the right of freedom of expression, the analysis involves three distinct stages.[137] The first distinction raised is whether the expression in question is excluded from constitutional protection under section 16(2) where categories of expression such as propaganda for war, incitement of imminent violence and etc exist. If the expression does not fall under these exclusions, then it is considered protected under section 16(1).[138]

The second stage assesses whether the common law or statutory provision in question limits the protection of freedom of expression. If such a limitation is found, the assessment is then conducted in terms of section 36 of the Constitution.[139] It is important to note that a limitation clause exists in section 36 of the Constitution which allows for restrictions on these guaranteed rights on the grounds that the restriction is reasonable, proportionate and based on the values of human dignity, equality and freedom.[140] Hate speech, which undermines the dignity of others and causes harm, is not protected by the Constitution. This is as it creates a conflict between the right to freedom of expression and the rights to human dignity and equality for which does not align with what free and fair societies strive.

4.2 Overview of South Africa's Social Media Landscape

South Africa's social media landscape is growing rapidly, attested to by the fact that in January 2023, there were 25.8 million social media users in the country. This was the

135 S Papadopoulos and Snail S Cyberlaw @ SA IV: The Law of the Internet in South Africa9 4TH edn, Van Schaik Publishers 2022 11.3.1

136 Botha, J *Hate Speech as a Limitation to Freedom of Expression* (PhD, Nelson Mandela Metropolitan University 2016) 21.

137 D Milo and P Stein a Practical Guide to Media Law (2ed, Lexis Nexis, 2013) 1.

138 *Ibid 137*.

139 D Milo and P Stein a Practical Guide to Media Law (2ed, Lexis Nexis, 2013) 2.

140 United Nations, 'International Human Rights Law' [online] United Nations Office on Genocide Prevention and the Responsibility to Protect available at https://www.un.org/en/hate-speech/united-nations-and-hate-speech/international-human-rights-law (accessed 26/10/2023).

© University of Pretoria

equivalent of 42.9% of the total population.[141] Facebook is the most popular platform, with a market share of 56.7%. TikTok, Instagram, Twitter, and LinkedIn are also gaining in popularity.[142] The number of active social media users in South Africa has grown exponentially in the last eight years.[143] In 2015, there were only 9.4 million social media users in the country.[144] This growth is being driven by several factors, including the increasing availability of smartphones and the growing accessibility of mobile data.[145] South Africans use social media for a variety of purposes, including entertainment, news and communication.[146]

4.3 Historical Context of Hate Speech In South Africa

The Republic of South Africa, established in 1961, was governed by a White minority who implemented an apartheid policy that violated the fundamental human rights of the majority.[147] This violation of human rights took many forms, including repressive laws, arbitrary arrests, detention without trial and harassment by the police.[148] In addition, the government had a tight grip on the media, denying citizens the right to full access to information.[149] Under the Publications Act 42 of 1974, they censored information disseminated through the media, which could be considered a form of indoctrination.

These laws include the Native Administration Act 30 of 1927 which criminalised "promoting feelings of hostility" between "natives" and Europeans and the Riotous Assemblies Amendment Act 19 of 1930, which was set to control public gatherings

---

141 DataReportal (2023), *Digital 2023 South Africa*, available at https://datareportal.com/reports/digital-2023-south-africa (accessed 12/09/2023).

142 Business Live (2023), 'Facebook Still Dominates South Africa's Social Media Landscape', available at https://www.businesslive.co.za/redzone/news-insights/2023-07-04-facebook-still-dominates-south-africas-social-media-landscape/ (accessed 12/09/2023).

143 Statista "Social Media in South Africa" (2023) available at https://www.statista.com/topics/9923/social-media-in-south-africa/ (accessed 12/09/2023).

144 Statista "Number Of Social Media Users In South Africa From 2017 to 2025" (2021) available at https://www.statista.com/statistics/1306799/number-of-social-media-users-south-africa/ (accessed 11/10/2023).

145 *Ibid.*

146 Fátima *Hate Speech on social media: A Global Approach* (Communication Library vol 7) (2023) available at https://store.lexisnexis.co.za/products/hate-speech-on-social-media-a-global-approach-skuZASKUPG00 (accessed 09/10/2023) 76,144 and 160.

147 United States Department of State, Office of the Historian (n.d.), 'South Africa - Countries - Office of the Historian', available at [https://history.state.gov/countries/south-africa] (accessed 27/10/2023).

148 Van Vollenhoven, W. J. (n.d.), *Learners' Understanding of Their Right to Freedom of Expression In South Africa* (PhD thesis, University of Pretoria) 2.

149 *Ibid.*

© University of Pretoria

and effectively prohibit publications that could have potentially incited hostility.[150] These laws effectively silenced and stifled any free expression of the Black population.

The 1994 elections in South Africa were a watershed moment in the country's history as they marked the end of apartheid and the beginning of a new era of democracy, unity, and diversity.[151] South Africa enacted a Constitution which includes a Bill of Rights as the supreme law of the country, and the Constitutional Court has the power to declare any law or action that violates the Bill of Rights invalid.[152]

Before the Constitution was enacted, there was systematic discrimination, oppression, and unjustified limitations on human rights.[153] The right to freedom of expression was one of the rights that was most restricted by the state and the experiences of victims of hate crimes in South Africa are shaped by the country's history of discrimination.[154] Racism and the legacy of colonialism and apartheid continue to have a significant impact on South African society, economy, and public and private discourse.[155] These legacies are evident in the high levels of poverty and inequality that persist in the country, as well as in the ongoing discrimination that people of colour face in many areas of life.[156]

The post-apartheid South African government was founded on the principles of safety and security, human dignity, equality, and social justice.[157] This is why hate speech in the country is defined as speech that advocates hatred against a person or group based on their race, ethnicity, gender, or religion, and that is likely to incite harm. This

150 Segregationist Legislation Timeline 1930-1939 | South African History Online; Riotous Assemblies Amendment Act No 19 - The O'Malley Archives

151 SA History Online “The South African General Elections: 1994” (date) (South African History Online) available at https://www.sahistory.org.za/article/south-african-general-elections-1994 (accessed 27/10/2023).

152 Department of Justice and Constitutional Development, ‘The SA Constitution Bill of Rights Chapter 2, Section 7-39’ (Department of Justice and Constitutional Development) https://www.justice.gov.za/constitution/chp02.html (accessed 27/10/2023)

153 Kanamugire & Faku “Employees' Right to Freedom of Expression Through Social Media in South Africa” (2022) 11 Persp L Pub Admin 394.

154 Marais, Nel & Govender “Emotional Consequences of Hate Incidents: Experiences of a South African Cohort” (2019) 36(2) South African Journal of Psychology 1.

155 The Editors of Encyclopaedia Britannica, ‘Apartheid | South Africa, Definition, Facts, Beginning, & End | Britannica’ (Encyclopaedia Britannica, 28 Sep 2023) https://www.britannica.com/topic/apartheid (accessed 27 October 2023)

156 Reddy, ES, “The Struggle against Apartheid: Lessons for Today’s World” (United Nations) https://www.un.org/en/chronicle/article/struggle-against-apartheid-lessons-todays-world (accessed 27/10/2023).

157 The Constitution of the Republic of South Africa

© University of Pretoria

definition is based on the Constitution of the Republic of South Africa, which not only guarantees freedom of expression but also allows for restrictions on this right in certain circumstances when it is necessary to protect the rights of others.

Despite its progressive legislative framework, the country still struggles with daily incidences of injustice and violent crime.[158] The difficulty in finding an appropriate balance between the prohibition of hate speech and freedom of expression lies in the fact that free speech is a fundamental right that is essential for a healthy democracy.

4.4 The South African Legal Framework

4.4.1 The Constitutional Provisions and Hate Speech Legislation

As mentioned above it is important to note that a limitation clause exists in section 36 of the Constitution which allows for restrictions on these guaranteed rights on the grounds that the restriction is reasonable, proportionate, and based on the values of human dignity, equality, and freedom.[159] This view is further supported by Marais who mentions that, South Africa law, fulfils a constitutional mandate to enact legislation that prohibits unfair discrimination by both the state and private individuals.[160]

The Bill of Rights is the cornerstone of democracy in South Africa.[161] It protects and upholds various rights and democratic values, such as human dignity, equality, and privacy. The Constitution states that everyone has a fundamental right to dignity and the right to have their dignity respected and protected. This can be seen in the case of Dawood & another v Minister of Home Affairs & others; Thomas & another v Minister of Home Affairs & others[162] which said:

> *"The value of dignity on our Constitution framework cannot..., be doubted: The Constitution asserts dignity to contradict our past in which human dignity for black South Africans was routinely and cruelly denied. It asserts it too to inform the future, to invest in our democracy respect for the intrinsic*

158 Marais, Nel & Govender "Emotional Consequences of Hate Incidents: Experiences of a South African Cohort" (2019) 36(2) *South African Journal of Psychology* 1.
159 *Ibid.*
160 M Marais, 'Hate Speech in the Equality Act Following the Constitutional Court Judgement in Qwelane v SAHRC (2023) 26 PER/PELJ 12.
161 The Bill of Rights Chapter 2, ss 7-39 the Constitution of South Africa 1996.
162 *Dawood and Another v Minister of Home Affairs and Others; Shalabi and Another v Minister of Home Affairs and Others; Thomas and Another v Minister of Home Affairs and Others* [2000] ZACC 8 (7 June 2000) para 35.

© University of Pretoria

worth of human beings. Human dignity therefore informs constitutional adjudication and the interpretation at a range of levels."

Hate speech is increasingly becoming the main cause of the violation of the right to dignity in South Africa.[163]. The overarching purpose of hate speech laws is to safeguard and protect human dignity in an egalitarian framework that is note determined by personal honour and self-worth.[164] The laws aim for inclusivity through participation in ordinary social life with respect and acceptance whilst detracting from exclusion and marginalisation on the basis of being a member of a minority group.[165]

In South Africa, hate speech is often directed towards marginalised groups such as LGBTQI+ individuals, immigrants, and people of colour.[166] Examples of hate speech on social media in South Africa include racist, homophobic, and xenophobic comments as well as incitement to violence. Users abuse social media to spread racist and xenophobic comments, especially towards immigrants from other African countries.[167] Others have also used social media platforms to spread homophobic comments, especially towards LGBTQI+ individuals.[168] Posts have used social media platforms to incite violence against certain groups, especially during times of political unrest.[169]

As mentioned above, tension between freedom of expression and hate speech is a complex issue. While freedom of expression is a fundamental human right, hate speech can cause harm to individuals and groups. South Africa's constitution allows it to recognise international law, such as the ICCPR, Malabo Convention and UDHR.[170] This is why South Africa is one of the most international law-friendly countries in the

163 SAnews.gov.za, 'Hate Crimes Policy to Balance Right to Dignity, Free Speech' (18/09/2013) available at https://www.sanews.gov.za/south-africa/hate-crimes-policy-balance-right-dignity-free-speech (accessed 27/10/2023).

164 D Milo and P Stein a Practical Guide to Media Law (2ed, Lexis Nexis, 2013) 164.

165 *Ibid*.

166 UNESCO "International Standard Classification of Education (ISCED)" (2011) available at https://unesdoc.unesco.org/ark:/48223/pf0000379177 (accessed 12/09/2023).

167 Marais, Nel & Govender "Emotional Consequences of Hate Incidents: Experiences of a South African Cohort" (2019) 36(2) *South African Journal of Psychology* 1.

168 *Ibid*.

169 UNESCO "International Standard Classification of Education *op cit*.

170 Human Rights Watch (1998), Incomplete 'Appendix A: South Africa's Obligations Under International And Domestic Law', available at [https://www.hrw.org/legacy/reports98/sareport/App1a.htm] (accessed 30/10/2023).

© University of Pretoria

world.[171] Therefore, the constitution also recognises the right to freedom of expression and the need for limitations on this right in certain circumstances.[172]

The ICCPR's Article 20 and the UDHR's limitations clause allow for permissible restrictions on freedom of expression in cases of hate speech that incites violence or discrimination.[173] Article 29(3) of the Malabo Convention requires states parties to pass laws that make it a crime to spread racist or xenophobic ideas or theories through computer systems, to attack people through computer systems because of their race, colour, descent, national or ethnic origin, religion, or political opinion, or to deny, approve or justify acts of genocide or crimes against humanity through computer systems.[174]

4.4.2 The Promotion of Equality and Prevention of Unfair Discrimination Act

The Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000 (hereafter PEPUDA) was enacted to ensure that South Africa complies with its international obligations.[175] This Act, which can be found in section 2, lists its goals as wanting to protect equality and stop unfair discrimination and hate speech.[176] It is important to understand its objectives in order to find ways to fix problems which could be created by the law.[177] This Act prohibits hate speech that promotes or propagates hatred because of race, gender, sex, pregnancy, marital status, ethnic or social origin, colour, sexual orientation, age, disability, religion, conscience, belief, culture, language and birth.[178] Its promulgation is in accordance with section 9(4) of the Constitution making it constitutionally mandated.[179]

---

171 Tladi, D "Interpretation and International Law in South African Courts: The Supreme Court of Appeal and the Al Bashir Saga" (2016) 16(2) African Human Rights Law Journal 310 - 338

172 Mendel, T. (2000), 'ICCPR Article 19: Freedom of Expression' (Office of the United Nations High Commissioner for Human Rights, Santiago), available at https://www.ohchr.org/Documents/Issues/Expression/ICCPR/Santiago/TobyMendel.pdf (accessed 12/09/2023).

173 Mendel T *op cit.*

174 African Union Convention on Cyber Security and Personal Data Protection' (African Union, 31 October 2023) available at https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection (accessed 31/10/2023).

175 Section 3 PEPUDA; Section 233 of the Constitution.

176 Section 2 PEPUDA; cf Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 15

177 *Ibid.*

178 Section 7 PEPUDA.

179 D Milo and P Stein a Practical Guide to Media Law (2ed, Lexis Nexis, 2013) 166.

© University of Pretoria

This Act which is regarded as a legislative milestone has been seen to have many interpretative and Constitutional concerns.[180] The focus of this essay will mainly be on section 7(a) which regulates hate speech as a form of unfair discrimination, section 10 which is the primary hate speech regulatory piece and section 12 which has a special rule related to section 10 also stops people from broadcasting or publishing anything that unfairly discriminates.[181]

Bearing in mind the country's history of racial segregation it is agreeable to identify that section 7(a) of PEPUDA prohibits the promotion of racial superiority of any person or the incitement of racial violence as a form of unfair discrimination.[182] Section 7(a) states that:

neither the state nor any person may unfairly discriminate against another person on the ground of race, including –

a) the dissemination of any propaganda or idea, which propounds the racial superiority or inferiority of any person, including incitement to, or participation in, any form of racial violence.[183]

Similarly, Article 4(a) of the ICERD requires countries to criminalise the spread of ideas based on racial superiority or hatred, as well as incitement to racial discrimination. However, the placement of this provision, especially where racial discrimination is the focus, has led to an ambiguous interpretation as to whether it limits freedom of expression as a form of hate speech or as unfair discrimination.[184]

This is demonstrated in *South African Human Rights Commission v Khumalo*[185] where the court grappled with the application of this section in regard to a controversial utterance made by Mr Khumalo.[186] With the court having to balance protection of individuals and harmful speech with one's constitutional right to freedom of expression,

180 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 403 - 516
181 *Ibid.*
182 Botha, J. (2016*), op cit.*
183 Section 7 PEPUDA
184 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 403 - 516
185 *South African Human Rights Commission v Khumalo* 2019 1 SA 289 (GJ)
186 Geldenhuys, J & Kelly-Louw, M 'Jurisdictional and Procedural Technicalities in Hate Speech Cases: South African Human Rights Commission v Khumalo' (2020) 23 ER (Potchefstroom) 1.

© University of Pretoria

the ambiguity created by the wording made it difficult to ascertain whether the utterance constituted hate speech or unfair discrimination.

Section 12 of the Act also prohibits speech as a form of unfair discrimination and provides that:

No person may –

a) disseminate or broadcast any information.

b) publish or display any advertisement or notice that could be construed or reasonably be understood to demonstrate a clear intention to unfairly discriminate against any person.

It is clear to see that Section 12 that strives to uphold the dignity of all individuals, and which prohibits discriminatory speech, aims at preventing the dissemination of information that could be interpreted as unfair discrimination. This is not without its challenges as the section's broad wording is open to interpretation which allows for inconsistent application of the law. This could in turn lead to suppression of legitimate speech in the digital age. Because of its wide range of speech forms that have been mentioned the section has the possibility of being considered an unjustified limitation on freedom of expression, as per Section 36 of the Constitution.[187]

As mentioned above, uncertainty surrounds Section 7(a) and 12 as they do not clearly specify whether they address hate speech itself or speech constituting unfair discrimination. Section 10(1) mentions “Subject to the proviso in section 12, no person may publish, propagate, advocate or communicate words based on one or more of the prohibited grounds, against any person, that could reasonably be construed to demonstrate a clear intention to be hurtful; be harmful or to incite harm; promote or propagate hatred.”[188]

Section 10(1) is seen to only target the forms of expression that are motivated by malicious intent engaged in with malicious intent.[189] An example made by Marais of the good-faith display of a painting in an art gallery depicting the brutal mistreatment

187 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 403 - 516

188 Section 10(1) substituted by s 20 of Act 15 of 2023 with effect from 3 April 2024.

189 M Marais, ‘Hate Speech in the Equality Act Following the Constitutional Court Judgement in Qwelane v SAHRC (2023) 26 PER/PELJ 15.

© University of Pretoria

of slaves would not amount to hate speech, even if it would cause psychological distress or discomfort in others.[190] However, positioning that same painting on someone's doorstep in a racially charged context could be seen as one's intention to incite hatred and harm.[191] If the same painting was delivered by mistake, it would not constitute discrimination as the mistake made wasn't based on race.[192]

Looking at the example above has demonstrated how the interpretive and constitutional concerns over this section can be easily challenged in the Courts. This can be seen in November 2019, when the Supreme Court of Appeal (SCA) ruled that section 10(1) of the Equality Act, which prohibits 'hate speech,' was unconstitutional.[193] Section 10 prohibits any speech that is 'hurtful' or 'degrades' a person or group based on their race, ethnicity, gender, religion or sexual orientation. The SCA found that section 10 was unconstitutional because it was too broad, vague, did not provide adequate safeguards against abuse and could be used to censor legitimate speech.

This view does not leave section 12 unscathed as its own broad language issues made it challenging to determine the precise boundaries of permissible speech.[194] The section did not require the State to prove that the speech was likely to cause harm, and it did not allow for any exceptions, such as for artistic or academic speech. The ruling made it clear that the government cannot censor speech simply because it is offensive or hurtful and made it clear that the government must provide adequate safeguards against abuse of hate speech laws.[195]

Despite its challenges this does not mean that this legislation is without merit, PEPUDA has provided for the establishment of the Equality Court, which has the power to hear cases related to hate speech and discrimination.[196] The court allows a range of remedies - these include but are not limited to an interim order, a declaratory order, an order for payment of damages to the aggrieved party for financial losses, impairment of dignity or pain and suffering. This is inclusive of emotional distress and

190 *Ibid.*
191 M Marais, *op cit.*
192 M Marais, *op cit.*
193 *Qwelane v South African Human Rights Commission* 2020 3 BCLR 334 (SCA).
194 *Ibid.*
195 Qwelane v South African Human Rights Commission *op cit.*
196 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression page 403 - 516.

© University of Pretoria

an order that an unconditional apology be made just to name a few.[197] It is important to note that the court can refer the matter to the National Prosecuting Authority to be charged as a crime in terms of the common law or relevant legislation.[198]

Taking all this into consideration it can be seen that PEPUDA was founded to address the inequalities of the past and promote reconciliation in society by providing victims of hate speech and unfair discrimination with remedies and forums within which to pursue complaints. This is commendable. However, this reminds one of the idioms 'the devil is in the detail.' This Act set the standard with its principles, establishing legal precedents and definitions which have allowed for more legislation to emerge, namely, the Hate Speech Bill and Cybercrimes Act which in their own right attempt to strive for greater clarity and specificity. By drawing on the criticisms faced by PEPUDA, these newer laws can be redrafted or in other instances amended to avoid similar ambiguities and ensure effective enforcement without confusion on application.

The question that now remains is whether PEPUDA is a good or bad legislation? This author believes while this legislation is not without its flaws and has been subject to criticism, its positive impact outweighs its shortcomings, therefore making it effective. This is evidenced by cases like those of Sparrow and Momberg. Its practical application and alignment with its original intent are discussed further below.

4.4.3 Prevention and Combating of Hate Crime and Hate Speech Bill [B9 of 2018]

Both PEPUDA and the Prevention and Combating Hate Crime and Hate Speech Bill have the similar aim of combating discrimination and hate speech. It is important to note that they use varying methods to achieve this. PEPUDA, which covers a wide range of discriminatory behaviour beyond hate speech, has resorted to dealing with hate speech and discrimination as civil cases by employing civil remedies such as damages and apologies. The Bill, on the other hand, which focuses specifically on hate crimes and hate speech, deals with these issues by treating them as criminal offences. It therefore imposes criminal penalties such as imprisonment and fines.

South Africa is committed to criminalising hate speech due to several high-profile cases and its commitment to the 1956 International Convention on the Elimination of

197 *Ibid.*
198 D Milo and P Stein a Practical Guide to Media Law (2ed, Lexis Nexis, 2013) 166

© University of Pretoria

All Forms of Racial Discrimination.[199] In September 2022, the South African National Assembly introduced the Prevention and Combating of Hate Crime and Hate Speech Bill.[200] This new law would create two new offences - hate crime and hate speech.[201]

Hate crime is defined as any crime that is motivated by hatred or prejudice against a person's race, religion, sexual orientation, gender identity, disability or other intrinsic characteristics.[202] Hate speech is defined as any public statement that is intended to be harmful or incite harm against a person or group of people based on their race, religion, gender, sexual orientation or other protected characteristics. The Bill is still under consideration by the National Assembly,[203] but if it is passed into law, it would be a significant step forward in South Africa's fight against hate crime and hate speech.

This step however raises questions such as whether any of the existing laws can be used to regulate extreme hate speech directed at target groups and which causes violence or severe harm.[204] It is crucial to consider existing laws such as *crimen iniuria*, defamation, public assault and incitement to punish hate speech.[205]

This Bill which is in essence an extension or narrower approach of PEPUDA has its foundation in introducing crimen injuria in order to regulate hate speech at a criminal level, especially where there is emphasis on derogatory terms used.[206] *Crimen iniuria,* defined as the unlawful, intentional and serious violation of dignity and privacy of another, has been used[207] for many cases, some of which are discussed below. These are cases where the complainants have been successful in suing others for abusive or insulting language. With this considered the question of whether this bill is

---

199 Burchell "Balancing Equality of Respect with Freedom of Expression: The *Actio Iniuriarum* and Hate Speech" (2019) 2019 Acta Juridica 203.

200 Prevention and Combating of Hate Crimes and Hate Speech Bill' (Parliamentary Monitoring Group 2023) https://pmg.org.za/bill/779/ (accessed 30/10/2023).

201 Singo, D, 'The Hate Crimes and Hate Speech Bill is a step closer to becoming law' (*Werksmans e-Bulletin,* 20 March 2023) (accessed 30/10//2023. Available at: https://www.werksmans.com/legal-updates-and-opinions/the-hate-crimes-and-hate-speech-bill-is-a-step-further-to-becoming-law/ (accessed 30/10/2023).

202 Marais A, Nel JA & Govender, R, 'Emotional Consequences of Hate Incidents: Experiences of a South African Cohort" (2019) 36(2) *South African Journal of Psychology* 1.

203 Prevention and Combating of Hate Crimes and Hate Speech Bill B9B-2018' (Parliamentary Monitoring Group 2023) available at https://static.pmg.org.za/Hate_Speech_Bill_B9B-202316065.pdf (accessed 30/10/2023).

204 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 403 - 516

205 *Ibid.*

206 Botha, J. *op cit 403 - 516.*

207 Botha, J. *op cit 403 - 516.*

© University of Pretoria

redundant in the face of the existing legal mechanisms already provided by PEPUDA and *crimen iniuria* remains to be answered.[208]

Crimen iniuria has been used to prosecute cases such as those of Sparrow and Momberg which will be discussed in more detail below. The United Nations human rights office of the high commissioner recommended there should be a distinction between three types of expression that may be considered hate speech.[209] These include "expression that constitutes a criminal offence, expression that is not criminally punishable but may justify a civil suit or administrative sanctions and expression that does not give rise to criminal sanctions but still raises concerns in terms of tolerance.[210]

In the Rabat Plan it is stipulated that states must establish a rigorous standard for criminalising hate speech, focusing especially on the severity of the offence.[211] In line with the ICCPR criminal liability for hate speech requires there to be an intent to advocate hatred and incite harm, excluding mere recklessness or negligence.[212]

The publication of the Bill has ignited significant debate and anxiety within South African society as it criminalises hate speech instead of offering rehabilitative measures.[213] The harsh prison penalties associated with the Bill may lead to self-censorship, both among journalists and the public, even when speech does not violate its provisions.[214]As it stands the bill has used a blanket approach and criminalised all forms of hate speech. It attracts a penal sanction and the Bill's mentioned goals can already be achieved through existing laws and constitutional values on freedom of expression.[215] Whilst the law is stricter it remains to be seen if it is an ineffectual piece of legislation or game changing.

---

208 Bohler, N, Pienaar, G, Houston, G, Barolsky V & Majozi, N, (2017) 'Criminalising Racism in Tumultuous Times'. HSRC Policy Brief, March. Available at: https://hsrc.ac.za/uploads/pageContent/8382/02_HSRC%20POLICY%20BRIEF%20Racism%20policy.pdf

209 United Nations, Office of the High Commissioner for Human Rights, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression* UN Doc A/HRC/48/54, para 32 (2021).

210 Griffin, R. Why the hate speech bill should be withdrawn. Helen Suzman Foundation. Available at: https://hsf.org.za/publications/hsf-briefs/why-the-hate-speech-bill-should-be-withdrawn.

211 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 403 - 516.

212 *Ibid*.

213 Democratic Alliance, *Prevention and Combating of Hate Crimes and Hate Speech Bill: Public Submission of the Democratic Alliance* (May 2023).

214 *Ibid*.

215 Griffin, R. Why the hate speech bill should be withdrawn. Helen Suzman Foundation. Available at: https://hsf.org.za/publications/hsf-briefs/why-the-hate-speech-bill-should-be-withdrawn.

© University of Pretoria

One of the posited arguments in the hate speech debate is that no law, no matter the strictness, can entirely eliminate racism, the attempt to seek a legal remedy to deter racism and racist expression is futile.[216] Louw in support of this view indicated that banning right to a fundamental freedom of expression will enable this hatred to germinate exponentially underground.[217]

However, the author is in agreement with Gutto who challenges this view. He mentions that the purpose of laws against the propagation of hate speech is not to eradicate racist or sexist beliefs altogether, but rather to express society's collective condemnation of such views and to discourage their public expression.[218] Gutto emphasizes that the goal of such legislation is to limit and discourage—not completely eliminate—the belief or its outward promotion.[219] He draws a parallel to laws against intentional homicide and theft: while these crimes still occur, the existence of laws prohibiting them plays a crucial role in reducing their incidence.[220] Similarly, laws against hate speech serve to curb the public spread of harmful ideologies, even if they cannot fully erase the beliefs behind them.[221]

4. 4.4 The Films and Publications Act 65 of 1996

The primary legislation in South Africa for regulating online content is the Films and Publications Act 65 of 1996 (hereafter FPA). This law was enacted to replace previous legislation that aimed to censor cultural productions during the apartheid era. The FPA established the Film and Publications Board (FPB) to receive complaints or applications to evaluate the classification of cultural production regarding its suitability for an audience.[222] This Act criminalises categorical types of hate speech.[223] The Act did not specifically include hate speech as a category of prohibited content but it did mention that any publication, film, or game that promotes hatred against a specific group of people and encourages violence will be banned.[224] This legislation covers

---

216 Between Speech and Silence, Hate speech, Pornography, and the New South Africa page Shadrack B. Gutto 122
217 *Ibid at page 19*
218 Between Speech and Silence, Hate speech, Pornography, and the New South Africa page Shadrack B. Gutto 122
219 *Ibid.*
220 Shadrack B. Gutto op cit 122
221 *Ibid.*
222 Freedom House (2022),*South Africa: Freedom in the World 2022 Country Report*, available at [https://freedomhouse.org/country/south-africa/freedom-world/2022] (accessed 04/09/2023).
223 D Milo and P Stein a Practical Guide to Media Law (2ed, Lexis Nexis, 2013) 168.
224 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 546 - 555.

© University of Pretoria

traditional media outlets such as films, videos, publications and games with the aim of regulating the distribution and classification of these materials to ensure that they were suitable for the different age groups and did not promote harmful or offensive content.[225] An offence can be committed when an individual who knowingly fails to fall in line with the classification, broadcasts, distributes or publicly exhibits such material without prior classification.[226] The offence will be punishable by either a fine or imprisonment or both.[227]

The FPA does not cater to the advancements in the technological field regarding the digital content creation and distribution in the new age while the Films and Productions Amendment Act aims to bridge the gap between the two, which will be discussed below.[228]

4.4.5 The Films and Publications Amendment Act 11 of 2019

In 2015 the Minister of Communications proposed amendments to the Act, primarily to regulate online content, including hate speech.[229] The proposed key changes included increased fines, a definition of hate speech , a new prohibition under section (18H) which proposes to prohibit propaganda for war, incitement to violence and hate speech and internet service provider liability in certain cases.[230]This expansion was necessary due to the increasing prevalence of online content and the need to regulate its distribution.[231]

On 3 November 2019, the South African President signed into law an Amendment Act to the FPA, known as the 'Internet Censorship Bill' by opponents.[232] The new version of the Act shifts the intermediary liability completely by imposing new duties and

225 *Ibid*.
226 *Botha J op cit 546 - 555.*
227 *Ibid.*
228 Strachan, W & Ramoabi, N 2022, 'Amendments to the Films and Publications Act, 1996 are now in force' (ENSafrica, 1 March) accessed 11 December 2024,https://www.ensafrica.com/news/detail/5375/amendments-to-the-films-and-publications-act-.
229 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 456 - 555
230 *Ibid.*
231 Baker McKenzie (2022), 'South Africa: Films and Publications Amendment Act comes into force', InsightPlus, available at [https://insightplus.bakermckenzie.com/bm/technology-media-telecommunications_1/south-africa-films-and-publications-amendment-act-comes-into-force] (accessed 23/11/2024).
232 *Ibid.*

© University of Pretoria

obligations to ISPs, which become obliged to monitor illicit, abusive and harmful content, such as child exploitation and abuse imagery, war propaganda, incitement to violence, hate speech, and more.[233]

The FPAA has aimed to update/modify the definition of child pornography to align with the Constitutional Court's ruling in the case of *De Reuck v DPP* and also align the definition of sexual conduct with the Criminal Law (Sexual Offences and Related Matters) ACT 32 of 2007.[234][235] Social media has also been included and mentions various online technology tools and forms of electronic communication via the internet, such as websites for social networking and micro-blogging through which users create online communities to share information, ideas, personal messages and other content.[236]

The Act has defined hate speech as the following:

> Hate speech has been added to mean any speech, gesture, conduct, writing, display or publication which is prohibited in terms of section 16(2) of the Constitution, which propagates, advocates or communicates words against any person or identifiable group, which words could reasonably be construed to demonstrate a clear intention to be harmful, to incite harm and promote or propagate hatred against the said person or identifiable group;[237]

This definition requires all three elements to be present: a clear intention to harm, the potential to incite harm, and the promotion of hatred.[238] Additionally, the definition specifies that harm can include emotional, psychological, or moral distress.[239] The

---

233 Belli, Curzi and Britto "Online Content Regulation in the BRICS Countries: A Cybersecurity Approach to Responsible Social Media Platforms" (date?) (*SSRN*, 2 May 2023) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4424913 (accessed 04/09/2023).

234 The Daily Maverick (2020), 'Film and Publications Amendment Act: Protecting, Not Censoring Our Citizens in the Digital Age?', by Barnaschone Attorneys, available at https://www.dailymaverick.co.za/article/2020-08-27-film-and-publications-amendment-act-protecting-not-censoring-our-citizens-in-the-digital-age/ (accessed 23/11/2024).

235 Parliamentary Monitoring Group (PMG). Committee Meeting Reports [website]. Available at: https://pmg.org.za/committee-meetings/ (Accessed: 2024/11/24)

236 *Ibid.*

237 Parliamentary Monitoring Group (PMG) *op cit.*

238 Forsyth, RSA (n.d.), Internet Censorship: A Look at the Films and Publications Amendment Acts' Criminalisation of Hate Speech, FORSA, https://journals.sagepub.com/page/ioc/collections/reading-lists/apartheid-era-south-africa, (accessed 23/11/2024)

239 FOR SA (Freedom of Religion South Africa) (n.d.), 'Internet censorship? A look at the Films and Publications Amendment Act's criminalisation of hate speech', available at

© University of Pretoria

definition of hate speech in the FPAA is similar to that in PEPUDA as discussed above, focusing on speech that incites harm or promotes hatred against a specific group.[240] However, in contrast to PEPUDA and similarly to the Hate Speech Bill the FPAA amendments criminalise the distribution of hate speech, potentially leading to imprisonment or fines. With this new definition emanating from the Constitutional definition, it is important to note that the scope of prohibited speech has been broadened. An objective *mens rea* test has been introduced which, however, is in conflict with the subjective test of the Constitution.[241]

Sections 18H and 24G of the Act, criminalise the distribution of content that promotes war, incites immediate violence or advocates hate speech which is inclusive of online distribution through platforms like social media.[242] With this Act creating new criminal offences there has been concerns about the potential overreach and the impact on freedom of expression.[243] These inconsistencies and overly-broad definitions require careful examination in order to ensure a balanced approach to the regulation of hate speech and the protection of fundamental rights.[244]

4.4.6 The Cybercrime Act 19 of 2020

The South African Cybercrimes Act 19 of 2020, passed in May 2021, is a law that aims to combat and prosecute cybercrime and address harmful speech in the online environment, including incitement of violence and other harm.[245] It includes several new provisions, such as offences related to the spread of malware, the use of social engineering techniques and the abuse of online anonymity.[246] It does this by creating new offences related to cybercrime, such as unauthorised access to computer

[https://www.forsa.org.za/articles/internet-censorship-a-look-at-the-films-and-publications-amendment-acts-criminalisation-of-hate-speech] (accessed 23/11/2024).

240 *Ibid.*

241 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 546 - 555.

242 FOR SA (Freedom of Religion South Africa) (n.d.), 'Internet censorship? A look at the Films and Publications Amendment Act's criminalisation of hate speech', available at [https://www.forsa.org.za/articles/internet-censorship-a-look-at-the-films-and-publications-amendment-acts-criminalisation-of-hate-speech] (accessed 23/11/2024).

243 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 546 - 555.

244 *Ibid.*

245 Control Risks 2023, 'How the South African Cybercrimes Act 19 of 2020 Will Affect Individuals And Businesses', available at [https://www.controlrisks.com/our-thinking/insights/how-the-south-african-cybercrimes-act-19-of-2022-will-affect-individuals-and-businesses] (accessed 31/10/2023)

246 *Ibid*.

© University of Pretoria

systems,[247]data theft, malware and ransomware,[248] phishing and identity theft,[249] hate speech and incitement to violence[250] and cyber harassment.[251]

The Cybercrimes Act allows South African court's jurisdiction over various cybercrimes even if they were committed outside South Africa[252] under specific circumstances, and it regulates the powers of law enforcement to investigate cybercrimes.[253] These circumstances include cases where the accused was arrested within the country; the accused is a South African citizen or resident; if the crime was committed on a South African vessel, ship, offshore installation or aircraft; the crime was committed outside South Africa against a South African citizen or resident, a South African company or organisation or a South African government facility.[254]

The Cybercrimes Act is the first standalone law in South Africa that addresses both the procedures for investigating and prosecuting cybercrimes, as well as the specific criminal offences related to cybercrime.[255] This is supported by authors such as Sizwe Snail ka Mtuze and Melody Musoni [256], who mention that legislations such as this one has disturbed the course of making this country a safe haven for cybercriminals.

However, this authors' view is that if the current measures put into place to avoid cybercrimes are sufficient then there should be no need for such strong and severe laws to avert potential disaster.[257] In other words prevention is better than cure and as referenced by Watney " if proactive security measures are adequate, the need for a

247 S2 and 3 of the Cybercrimes Act ; MBA Incorporated (2024), 'Understanding the Cybercrimes Act, 19 of 2020, and Pig Butchering Schemes in South Africa', available at [https://mbaincorporated.co.za/thought-leadership-articles/posts/understanding-the-cybercrimes-act-19-of-2020-and-pig-butchering-schemes-in-south-africa/] (accessed 23/11/2024).

248 Section 5 and 8 of the Cybercrimes Act

249 Section 9 of the Cybercrimes Act

250 Section 14 and 15 of the Cybercrimes Act

251 Section 16 of the Cybercrimes Act

252 Chapter 3 of the Cybercrimes Act

253 *Ibid*.

254 Chapter 3 of the Cybercrimes Act

255 F Cronje, S Snail, An Olivier, J Jordaan, *Cybercrime in South Africa: Law and Practice* (LexisNexis Butterworths, 2022).

256 Sizwe Snail ka Mtuze and Melody Musoni, 'An Overview of Cybercrime Law in South Africa' (2023) 4 *International Cybersecurity Law Review* 299 https://doi.org/10.1365/s43439-023-00089-8.

257 S Papadopoulos and Snail S Cyberlaw @ SA IV: The Law of the Internet in South Africa9 4TH edn, Van Schaik Publishers 2022 13.2

© University of Pretoria

drastic and reactive system e.g. criminal law to be invoked might fall away completely or at least be diminished considerably."[258]

This Act is linked to the Hate Speech Bill insofar as the Bill has defined communications to include electronic communications. This can be seen in section one which mentions:[259] "In this Act, unless the context indicates otherwise— ''communication'' includes any— (a) display; (b) written, illustrated, visual or other descriptive matter; (c) oral statement; (d) representation or reference; or (e) an electronic communication, and ''communicates'' and ''communicated'' have a corresponding meaning;"

The Cybercrimes Act therefore correlates to the Prevention and Combating of Hate Crime and Hate Speech Bill under section 4 which mentions the offence of hate speech and how it is communicated. It even goes a step further by including incitement to damage of property as an offence whilst the Bill focuses primarily on incitement to create hatred against an individual.[260]

The Cybercrimes Act also amends several other South African laws, such as the Criminal Procedure Act, the South African Police Service Act and the Electronic Communications and Transactions Act just to name a few.[261]

The acts above vary in regard to their primary focus/purpose. PEPUDA, which has a broad scope of protection, focuses on equality, discrimination and offers civil remedies. The Hate Crimes and Hate Speech Act has introduced the criminalization of hate crimes and hate speech with specific focus on harmful intent The Films and Publications Act deals with content regulation, the regulation of traditional and online media as well as the classification of content and hate speech. Lastly the Cybercrimes Act criminalises cybercrimes, including online harassment and hate speech.

It is important to note the significant overlap and intersectionality of mechanisms exhibited by the mentioned acts in addressing hate speech. An example is the Hate Crimes and Hate Speech Act and the Cybercrimes Act that specifically criminalise

---

258 *Ibid.*
259 Section 1 of the Prevention of and Combating of Hate Crime and Hate Speech Bill [B9 of 2018]
260 *Ibid.*
261 Cybercrimes Act 19 of 2020 (South Africa) [legislation online] accessed 24 November 2024https://www.gov.za/sites/default/files/gcis_document/202011/20201127cybercrimes_act_19_of_2020.pdf

© University of Pretoria

hate speech. The FPAA and the Cybercrimes Act both regulate online content, inclusive of hate speech, which is disseminated through social media and other online platforms.

These laws complement each other in addressing the key issues, which are discrimination, hate speech and online harm. It is important to note that while these laws are in effect, the Hate Speech Bill is a welcome addition more specific to combating hate speech in the digital age with additional protections and remedies. The Bill provides a definition for hate speech, outlines penalties and authorises law enforcement to take affirmative action against hate speech offenders. The existing legislative framework as discussed has facilitated the formation of a bill that is more comprehensive in targeting the societal issue that is hate speech.

4.5 Security Concerns

The nature of the internet makes it difficult for South African public authorities to regulate speech on social media.[262] The underlying technology is constantly evolving and is often owned and controlled by foreign companies. Additionally, online expression can easily cross international borders. As a result, the power to regulate speech on social media may lie with a variety of private and public bodies, both in South Africa and in other countries. These bodies include internet service providers, other internet intermediaries and the social media platforms themselves.[263]

Numerous internet-related technologies significantly impact individual privacy.[264] This is because the design of computer systems is to amass, process and store significant collections of data.[265] The spread of connected devices such as digital cameras, cell phones, geo-identification tools and GPS– further underscores the critical importance of robust data protection measures.[266]

The internet presents a realized fear of loss of data and erosion of personal privacy to its users. The primary concern identified by internet users is a fear and lack of trust

262 Sive & Price "Regulating Expression on Social Media" (2019) *SALJ* 51.
263 *Ibid.*
264 S Papadopoulos and Snail S Cyberlaw @ SA IV: The Law of the Internet in South Africa9 4TH edn, Van Schaik Publishers 2022) 1.2.6
265 *Ibid.*
266 S Papadopoulos and Snail S *op cit* 1.2.6

© University of Pretoria

stemming from the potential loss of personal privacy.[267] Security breaches and vulnerabilities in firmware which exposes personal data have left internet users exposed to financial fraud and identity theft.[268] The South African government passed the Protection of Personal Information Act 4 of 2013 (hereafter POPIA). Prior to POPIA the author notes that the legislation and common law in place was not adequate, nor did it provide sufficient protection compared to its international counterparts.[269] Swales argues that this is because the users could not exercise full control of how their data was being accessed or utilised. The available remedies at the time were not up to standard in dealing with access and management of personal data by third parties.[270]

The POPIA created an Information Regulator to ensure data privacy.[271] POPIA has eight essential principles that organisations must follow, including accountability, processing limitation, purpose specification, further processing limitation, information quality, openness, security safeguards and data subject participation.[272] The Act came into full effect in 2020, and non-compliance can result in significant penalties.[273]

Personal information is defined as "information relating to an identifiable, living, natural person, and where applicable, an identifiable existing juristic person." The information about a person includes but is not limited to name, age, race, national or ethnic origin, religion, beliefs, identity number, education and medical history just to name a few examples.[274] This Act has been implemented to address issues related to social media usage as it has imposed strict rules on how organisations collect, store and process personal information including that gathered from social media.[275]

POPIA plays a crucial role as it protects personal information and mitigates the risks associated with the misuse of personal information which is what hate speech often

---

267 S Papadopoulos and Snail S Cyberlaw @ SA IV: The Law of the Internet in South Africa9 4TH edn, Van Schaik Publishers 2022) 10.1

268 *Ibid.*

269 Swales L "The protection of Personal Information Act 34 of 2013 in the Context of Health Research: Enabler of Privacy Rights or Roadblock? PER/PERJ 2022(25) 10.

270 *Ibid.*

271 Section 36 of POPIA.

272 Section 17 of POPIA; Tripwire, *Summary of South Africa's Protection of Personal Information Act (PoPIA)*(31 October 2023) https://www.tripwire.com/state-of-security/summary-south-africas-protection-personal-information-act-popia (accessed 24/11/2024).

273 Section 85 of POPIA.

274 CookieYes (2023), 'POPIA South Africa', available at [https://www.cookieyes.com/blog/popia-south-africa/] (accessed 23/11/ 2024).

275 Meyer Attorneys (2024), 'Social Media and the Law in South Africa: Navigating the Digital Landscape', available at https://meyerattorneys.co.za/2024/04/05/social-media-and-the-law-in-south-africa-navigating-the-digital-landscape/ (accessed 23/11/2024).

© University of Pretoria

thrives on such as race, gender, or other sensitive details etc. POPIA holds social media platforms accountable when personal information is misused especially when it is used to facilitate hate speech. This plays a part when cyberbullying is a factor as it also sometimes relies on personal information including private conversations or images.

4.6 Case Studies Illustrating the Implementation and Impact of These Laws

4.6.1 Overview

The law protects people from being subjected to derogatory or contemptuous words, even if those words are not directed at a third party.[276] This means that someone can be sued or prosecuted for hate speech, even if they only say it to the victim directly. In addition to words, other conduct can also be considered hate speech if it impairs a person's dignity.[277] When it comes to social media related disputes, there is often a conflict between freedom of expression and dignity.[278] It is the authors view that courts are therefore required to strike an appropriate balance between upholding freedom of expression and protecting the inherent value of human dignity.[279]

This is further supported by the works of Snail and Swales who mention that this must work in tandem with the analysis of the case details, the parties involved , the surrounding context and the competing rights in question which all will summarily assist in determining which right takes precedence given the situation.[280] As mentioned above the factors to consider are intention to harm, and whether the speech targets a specific group or individual based on protected characteristics such as race or religion. There are several cases of hate speech on social media that have been successfully prosecuted - or at the very least addressed - by our judicial system, discussed below.

---

276 Burchell “Balancing Equality of Respect with Freedom of Expression: The *Actio Iniuriarum* and Hate Speech” (2019) *Acta Juridica* 203.

277 *Ibid.*

278 S Papadopoulos and Snail S Cyberlaw @ SA IV: The Law of the Internet in South Africa9 4TH edn, Van Schaik Publishers 2022) 11.1

279 *Ibid.*

280 S Papadopoulos and Snail *op cit* 11.3.1

© University of Pretoria

4.6.2 AfriForum NPC & Another v Pienaar 2017 (1) SA 388 (WCC)

In AfriForum v Pienaar the right to freedom of expression and freedom of the press and other media, as expressed in the section 16(1) of the Constitution, was considered in so far as whether press and media includes social media such as Facebook and Twitter.[281] AfriForum, by entering the public and politicised campus, intentionally sought to engage in political debate.[282] The courts have acknowledged that such actions, under the constitutional right to freedom of expression, which includes social media, can have legal ramifications.[283] The court considered whether social media should be considered as part of the 'press and media' for purposes of constitutional protection.[284] "The court found that the postings, while strong and provocative, were fair comment on a matter of public interest."[285] The considerations taken into account by the court included the context of the event, the constitutional right of freedom of expression by looking at the judgement in Democratic Alliance v African National Congress[286], where the Constitutional Court categorised the internet alongside traditional media like public meetings, radio, television and newspapers as a space for robust public debate,"[287] as well as the political nature of the issue and came to a conclusion that the posts made were a legitimate exercise of this right.[288] In the end the Western Cape High Court expanded press freedom to social media.[289]

4.6.3 The Infamous Sparrow and Momberg cases

As mentioned above, using abusive or insulting language is grounds for a criminal and civil offence. A relevant, infamous case is that of Ms Penny Sparrow who sparked anger and rage when she referred to black Africans as monkeys in a Facebook post following the New Year's Eve celebrations in 2016.[290] The Equality Court convicted

---

281 *AfriForum NPC & Another v Pienaar* 2017 **(**1) SA 388 (WCC) par 55.

282 *Ibid.*

283 *AfriForum NPC & Another v Pienaar op cit* para 55.

284 Davis T, 'Social media content moderation and limitations on freedom of expression: the role of the state' (2022)

285 *Ibid.*

286 *Democratic Alliance v African National Congress and Another 2015 (2) SA 232 (CC)*

287 *AfriForum NPC & Another v Pienaar* 2017 **(**1) SA 388 (WCC) par 58.

288 *Ibid par 55.*

289 Davis T, 'Social media content moderation and limitations on freedom of expression: the role of the state' (2022)

290 Singh, K 2019, 'Penny Sparrow, whose racist post sparked fury, has died', News24 (25 July) accessed 11 December 2024,https://www.news24.com/news24/breaking-penny-sparrow-has-died-20190725.

© University of Pretoria

her of hate speech as defined in Section 10 of the PEPUDA.[291] The Umzinto Equality Court found that Sparrow's words "convey the message both explicitly and implicitly to the reader that black people are not worthy of being described as human beings - implicit in this is that they have lowered or sub-human intelligence" and within six months of her comments ordered her to pay R150 000 to the Adelaide and Oliver Tambo Foundation, with the presiding officer warning Sparrow against further hate speech. [292] At the end of her Equality Court hearing, the NPA considered charging her. Subsequently, in a criminal court she pleaded guilty to *crimen injuria* and entered into a plea agreement where she was fined R5 000 or 12-months imprisonment by the Scottburgh Magistrates court, a further two-year suspended sentence and the reading of an apology to those she had offended[293] This goes hand in hand with the Hate Speech Bill that pushes for criminalisation of hate speech.

The Momberg matter provided an unequivocal legal precedent that the use of the racial epithet 'kaffir' is considered hate speech under section 10 of PEPUDA and crimen injuria. This all began when South African Police Service officials attended to Momberg, following a smash-and-grab incident where she repeated the reference multiple times.[294] Her defence which relied on the suffering from 'sane automatism' due to the trauma she experienced in the smash-and-grab led the Randburg Magistrates Court to convict her on all four charges of *crimen injuria*. She was sentenced to an effective three-year imprisonment, a year of which was suspended on condition that she was not convicted of *crimen injuria* during the period.[295]Following the Equality Court's June 2017 order for a R100 000 fine, a public apology like Sparrow's mentioned above, and community service, Momberg served four months of her sentence. She was then released on R2 000 bail in August 2018, pending the appeal of her conviction. Despite the order, the obligations set for Momberg have not been fulfilled.

---

291 Maart R, (2020) 'Black Bodies on South African Beaches: Lus en smaak jou lekkerding 1' (2020) 27 Alternation.

292 Marais ME, 'A constitutional perspective on the Sparrow judgements' (2017) 42 Journal for Juridical Science 25 - 64.

293 *Ibid.*

294 Momberg v S (A206/2018) [2019] ZAGPJHC 183; 2019 (2) SACR 505 (GJ) (28 June 2019); SAHRC, 'Racist Vicki Momberg hands herself over to police' (2019) SAHRChttps://www.sahrc.org.za/index.php/sahrc-media/news/item/2206-racist-vicki-momberg-hands-herself-over-to-police (accessed 23/11/2024).

295 *Ibid para 1.*

© University of Pretoria

This brings us to the point raised above whether PEPUDA is an effective piece of legislation. As evidenced by these cases this law is still used to drive home a point that even with its ambiguous nature it is a deterrent and continues to be a point of reference to successfully prosecute those who do not comply with it. While PEPUDA provided the legislation that made the above prosecutions actionable it was criticised for not being able to deal with the rapid evolution of online communication, and the prevalence of hate speech on these platforms.

The introduction of new legislation discussed above has allowed society and legislation to move with the times better to meet the demands of the present. This is evidenced by Momberg's video recording and Sparrow's post which were circulated many times on not just virtual platforms but television and other forms of media. The FPAA and Cybercrimes Act, as evidenced above, are relevant to these scenarios as they provide the legal framework that speaks specifically and most importantly, narrowly to the aspects of hate speech on online communications, television, and the regulation regarding distribution of these matters.

4.6.4 Islamic Unity Convention v Independent Broadcasting Authority and AfriForum NPC v Nelson Mandela Foundation Trust and Others

The Constitutional Court's landmark ruling on hate speech in the Islamic Unity Convention v Independent Broadcasting Authority established the boundaries of freedom of expression in South Africa. The Court determined that most forms of expression are protected under the Constitution, but limitations must adhere to strict constitutional requirements. However, certain types of speech, including propaganda for war, incitement of violence and hate speech, are excluded from constitutional protection due to their potential to harm human dignity and undermine the principles of equality and non-discrimination. The Court emphasised that these excluded categories are clearly defined and restrictive. Laws regulating speech within these boundaries do not infringe on the right to freedom of expression.[296]

AfriForum NPC v Nelson Mandela Foundation Trust and Others also known as 'the old RSA flag case' highlighted the destructiveness of the old South African flag, a notable symbol of apartheid and a brutal and unjust oppressive regime. The key

296 Botha *Hate Speech as a Limitation to Freedom of Expression* (PhD, Nelson Mandela Metropolitan University 2016).

© University of Pretoria

question asked was whether the gratuitous display of this flag constitutes hate speech within the meaning of section 10(1) of the Equality Act. The high court had found that displaying the flag was a symbol of racial segregation and White supremacy and that it constituted hate speech, unfair discrimination, and harassment.[297]This ruling was met with opposition from AfriForum as it challenged the high court's order, arguing that the court lacked the authority to issue such an order and that public displays of the old flag were protected by constitutional rights. However, AfriForum failed to provide sufficient evidence to support these claims.[298]

When one thinks of hate speech it is usually alluded to what one says or writes. In this case it was proved that symbols matter as well. Context matters when hate speech is the discussion and how it is uttered or displayed plays a big part. This was demonstrated in this case where the historical, social, and symbolic significance of the flag was essential in determining whether its display constituted hate speech. The court in this matter ruled that the flag represented White supremacy and the oppression of Black people and created a hostile and intimidating environment for victims of apartheid. The public display of the flag, which was based on the grounds of race, was prohibited.[299] The recognition of the potential for speech to cause harm was considered, especially when it targeted marginalised groups or perpetuated stereotypes. The court further held that the public display of the flag constituted hate speech, unfair discrimination, and harassment, as defined in Sections 10(1), 7 and 11 of the Equality Act. It interpreted section 10(1) expansively to encompass the expression of ideas through conduct, such as the display of the flag.[300]

4.6.5 Cape South African Jewish Board of Deputies vs. Thomas Leyland Torr and Amaning and Others v Ackerman cases

Cape South African Jewish Board of Deputies vs Torr received judgement on the 9th of May 2025 where the courts had taken PEPUDA, specifically section 10(1) of this key legislation in terms of equality and section 16 of the Constitution in order to effectively balance the freedom of expression and hate speech which have been

297 *AfriForum NPC v Nelson Mandela Foundation Trust and Others (371/2020) [2023] ZASCA 58; 2023 (4) SA 1 (SCA); [2023] 3 All SA 1 (SCA) (21 April 2023).*

298 *Ibid.*

299 *AfriForum NPC v Nelson Mandela Foundation Trust and Others op cit.*

300 *AfriForum NPC v Nelson Mandela Foundation Trust and Others op cit.*

© University of Pretoria

extensively discussed above.[301] This is proof that although PEPUDA has had its issues, it is still presently able to provide legal justification to judgements made. The inception of this case was when Thomas Leyland Torr posted 22 videos during the 2023 – 2024 period on Tik Tok, Instagram and YouTube which had a clear intention to incite harm and promote hatred towards those who are of Jewish descent, religion, culture etc.[302] The court which found Torr guilty of hate speech had explicitly rejected his defence of protected speech after his content was described as “nasty, vicious and appalling" and finding a clear "intent to fuel antisemitism and cause harm."[303]

Hate in Private? Although, the Hate Speech Bill would have been a lovely addition in the Amaning[304] matter, this case demonstrated the broadening for the understanding of “public communication” for hate speech which is as it seems inclusive of private online messages. In this instance a businessperson used the infamous k word when communicating with a business associate on various occasions.[305] As discussed above, the communication of hate speech can take many forms, and the court has resolved this matter by applying the existing PEPUDA legislation. In paragraph 116 of the matter, it was mentioned that “there is no requirement inherent in the verb “communicate” that the offensive words be communicated widely.” It was further indicated that it was sufficient for an individual be addressed and that the intention of promoting hate speech might address one person at a time.

The presiding officer then effectively dealt with what constitutes privacy and quoted from Langa DP in Hyundai at paragraph 123, “privacy is a right which becomes more intense the closer it moves to the intimate personal sphere of life of human beings and less intense as it moves away from that core.” Consequently, this quote led a differentiation between degrees of privacy, and in this instance communications by Mr Ackerman and Mr Wellman were not considered to be in the “intimate personal sphere” in relation to section 10(1). Subsequently the businessman, Mr Ackerman was found guilty of hate speech. This case highlights and supports the authors views of the country being able to successfully apply existing legislation when it comes to the balance of freedom of expression and hate speech in the online sphere.

---

301 *Cape South African Jewish Board of Deputies v Torr EC11/2025(9 May 2025)*
302 *Ibid.*
303 *Op cit.*
304 Amaning and Others v Ackerman (EQ3/2023) [2024] ZAEQC 2 (10 May 2024)
305 Ibid at para 19,21,22

© University of Pretoria

4.7 Summation

In this chapter a significant point discussed above tries to strike a balance between protecting freedom of expression and curbing hate speech. They share a common goal of promoting social cohesion and protecting individuals from the harm caused by hate speech. Each law adds a piece to the hate speech puzzle in order to address the constantly evolving technological advances and societal shifts. It is essential to continually assess whether legislation relates to the current environment in order for it to be effective.

South Africa which is party to international laws and agreements such as the UDHR AND ICCPR has demonstrated its commitment to promoting equality and regulating hate speech with our domestic laws such as PEPUDA, the Hate Speech Bill and Cybercrimes Act. While this study seeks to understand the balancing of hate speech and freedom of expression on social media with reference to case laws, judgements, and interpretation of legislation, it is aware that even with consideration of international instruments and local application there are still challenges. Examples are whether this legislation is applied correctly, is too strict and leaves no room for rehabilitation, and also infringes on freedom of expression.

The chapter has revealed that unambiguous wording is essential in effective interpretation of the law, in which hate speech is not simply words. Context matters, be it historical or societal. There is legislation that caters for hate speech experienced within the applicable courts to address these grievances and offer remedies. In addition, it shows how one cannot merely rely on sane automatism to get away with hate speech because in the spotlight of the social media, transgressors will answer to the law.

© University of Pretoria

CHAPTER 5: SUMMARY AND CONCLUSION

5.1 Introduction

In the previous chapter it can be seen that there are various ways one can become an active participant online, just by posting a few words on a small rectangular device, making individuals creators and publishers of their own content.[306] In the past, social media platforms were primarily used for sharing information and connecting with friends and family. However, today the internet and social media have changed and challenge how one exercises one's right to freedom of expression.[307] This author has identified that this includes the rise of citizen journalism which has introduced a new dynamic to how news is shared. In the digital era, these social media platforms have become places where people can spread hate speech, target others and incite violence.[308] Hate speech is no longer confined to the dark corners of the internet; it is now increasingly common on all major social media platforms.[309]

This study has aimed to identify the effectiveness of South African legislation in balancing the right to freedom of expression and the need to protect against hate speech by looking at how hate speech is defined and regulated. International instruments were considered in providing a guideline when it comes to combating hate speech on social media. Finally, consolidation of key findings of the preceding chapters, discussion of the implications of these findings and recommendations for addressing identified issues are addressed in this chapter.

5.2 Synopsis of chapters

The first chapter introduces the study by reviewing existing research, outlining the research goals, questions, and limitations, and explaining the research methods and approach.

306 Davis T, Social Media Content Moderation and Limitations on Freedom of Expression: The Role of the State

307 *Ibid.*

308 Davis T *op cit.*

309 'Hate Speech and Social Media: Preventing Atrocities and Protecting Human Rights Online' (*Global Centre for the Responsibility to Protect*, 21 February 2020) https://www.globalr2p.org/publications/hate-speech-and-social-media-preventing-atrocities-and-protecting-human-rights-online/ (accessed 6/09/2023)

© University of Pretoria

Chapter two identified freedom of expression to be an internationally recognised right which includes the freedom for citizens to express themselves on social media. This facilitates the exchange of ideas and communication, drawing on the fact that it is an internationally recognized right. However, this freedom can be exploited for harmful purposes such as hate speech. International laws such as CERD, UNHRC and ICCPR prohibit advocacy of national, racial, gender or religious hatred that constitutes incitement to discrimination, hostility, or violence.

While there is no universally accepted definition of hate speech, social media platforms have their own preferences and guidelines as to what constitutes hate speech. The global regulation of social media is up for discussion as it remains a complex issue with many platforms facing criticism for their inadequate content moderation efforts. Different countries have varying regulations which further complicates the issue (despite social media platforms' ability to induce citizen journalism).

Chapter three is dedicated to hate speech and its implications particularly, in the context of social media. The debate surrounding the regulation of hate speech is explored, considering arguments for and against such regulation. Three distinct arguments arise: one advocating for unrestricted speech, another that advocates for and prioritises the principles of equality and social justice, and a third which opts for navigating the tension between these rights.

This chapter also examines the role of social media in the hate speech propagation by highlighting the ease with which misinformation and harmful speech can spread through various platforms and throughout history. The challenges of effectively monitoring hate speech were discussed, inclusive of privacy concerns and the difficulty in distinguishing between genuine hate speech and satire or humour. Finally, the tension between freedom of expression and the protection of vulnerable groups from hate speech was explored.

Chapter four was devoted to South Africa's legal system regarding freedom of expression and hate speech on social media. The background of the deep-rooted past in a very controlled system of freedom of expression is introduced which demonstrates why the country has the existing legislation in place in order to learn from the past. Various legislations are introduced such as PEPUDA, Hate Speech Bill, Films and Publications Act (former and amended versions), Cybercrimes Act and POPIA. As

© University of Pretoria

seen these legislations share the same goal of dealing with hate speech, however, as seen above they each contribute to the different aspects such as where hate speech can come from, the remedies available for victims of hate speech, how the perpetrators are punished and how to better protect groups and individuals. Case law is provided to show how the courts approach, uphold, and define these various legislations. In conclusion to this chapter, a comparison of and how international principles and local legislation are intertwined in the management of hate speech and freedom of expression on social media, is presented.

5.3 Recommendations

5.3.1 Overview

Hate speech is not just about words, it is about power. It is about using language to silence and intimidate those who are different. Strategies for balancing freedom of expression and addressing hate speech on social media in South Africa as recommended by the Report of the UNHCHR are explored. Expert workshops on the prohibition of incitement to national, racial, or religious hatred include education and media literacy, proactive moderation by social media platforms, and transparent and consistent enforcement of regulations are covered.[310]

5.3.2 Strengthening Legal Frameworks and Enforcement

As referenced above the ambiguous nature of section 10 and 12 of PEPUDA is open to exploration. Therefore, this author supports the proposed amendments of these sections by J Botha. Section 10(1) would now read:

(1) No person may publish, propagate, or advocate any form of expression in public against an identifiable group of persons on one or more of the prohibited grounds that could reasonably be construed to demonstrate a clear intention to promote hatred against such group and which is likely to cause harm.[311]

(2) Defences sub-section (1) does not apply to:

a. The *bona fide* engagement in artistic creativity, academic or scientific enquiry.

---

310 UN High Commissioner for Human Rights, *Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial, or religious hatred* (UN Doc A/HRC/49/34, 2021).

311 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression 502,600 - 602.

© University of Pretoria

b. The *bona fide* engagement in fair media reporting and commentary in the public interest.[312]

The integration of these defenses into this section allows for the enhancement of clarity and avoids potential ambiguities that emanate from the current structure.[313] The Section 12 amendment mentions the prohibition of the dissemination or broadcast of discriminatory symbols and the publication or display of discriminatory advertisements or signs that promote unfair discrimination against a vulnerable group these are based on one or more prohibited grounds. This was a topic of debate in the 'old RSA flag case' with a focus on 'vulnerable groups,' The amendment provides a clear approach to the protection of marginalised individuals and communities from harmful discriminatory expression.[314]

As the previous recommendation focused on addressing hate speech, the following suggestion pertains to the protection of freedom of expression. This approach aims to maintain a balanced perspective, avoiding undue emphasis on one aspect over the other. Concerning the Hate Speech bill, the author recommends a tiered system, much like one that mirrors that of criminal law offences namely murder and culpable homicide.[315] In South African law, there is a distinction between murder and culpable homicide which hinges on the concept of *mens rea*. Murder is the deliberate and intentional killing while culpable homicide involves a lesser degree of intent and is defined as the unlawful and negligent killing of another human being.[316] For example, a driver who causes a fatal accident due to reckless driving by speeding, or driving under the influence of drugs or alcohol may be charged with culpable homicide as their intention was not to kill , but their lack of responsible behaviour led to the tragic

312 *Ibid.*

313 Botha, J *Op. cit. 600 – 602.*

314 Botha, J *Op. cit. 600 – 602.*

315 Schedule 1 substituted by section 24(5) of the *Protection of Constitutional Democracy against Terrorism and related activities Amendment Act*, 2022 (Act No. 23 of 2022), Notice No. 1533, G47803, dated 29 December 2022, effective 4 January 2023 per Proclamation Notice 110, GG47820, dated 4 January 2023.

316 Hagglund, K & Khan, F 2020, 'Reformulating dolus eventualis: Guidance from USA and Germany' South African Crime Quarterly, vol. 69, no. 69a6108,http://dx.doi.org/10.17159/2413-3108/2020/v0n69a6108 accessed [accessed 01/12/2024].

© University of Pretoria

outcome.[317] This demonstrates how the law considers both the mental state of the perpetrator and the act committed.[318]

The recommendation to apply a tiered system to hate speech would allow for various degrees of severity.[319] This recommendation would align with the Rabat Plan that stipulates that there must be an implementation of a high threshold test for a criminal sanction for hate speech and that there should always be a consideration of severity.[320] This threshold test from the ICCPR ,[321] which is similar to section 16(2)(c) of the Constitution, must be applied in order to ensure that the expressed conduct by the perpetrator advocates hatred, on prohibited grounds and incites harm.[322]

A casual, insensitive statement or remark made as a joke, though considered offensive, may not constitute the same level of punishment as a deliberate attack designed to incite violence or hatred. A similar question was raised by Mr Motsamai who highlighted the trend of individuals being prosecuted for singing the 'Kill the Boer, Kill the Farmer' song, often initiated by right-wing groups.[323] Considering the historical and cultural significance of such songs, particularly during periods of oppression, there is a concern that this could lead to widespread imprisonment for cultural expression.[324]

This system recommended by the author would allow for a proportionate response considering numerous factors such as intent, context and impact allowing for a more case-by-case basis instead of the answer simply being jail and a fine. As mentioned above, a system of this nature could facilitate rehabilitative measures, such as drawing on groups like Alcoholics Anonymous, alongside initiatives aimed at curbing hate speech. This view is supported by the likes of Gutto who asserts that there is a need for a form of punishment to the offender or an appropriate remedy to the victim ought

317 *Zulu v S (189/2011) [2012] ZAKZPHC 26 (4 May 2012)*

318 *Ibid.*

319 Prevention and Combating of Hate Crimes and Hate Speech Bill: public hearings day 2, NCOP Security and Justice, 20 September 2023, chaired by Ms S Shaikh (ANC, Limpopo).

320 Botha, J. (2016), *Hate Speech as a Limitation to Freedom of Expression; UN High Commissioner for Human Rights, *Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial, or religious hatred* (UN Doc A/HRC/49/34, 2021)

321 Article 20(2) of the ICCPR

322 Botha, J (2016) op cit

323 Prevention and Combating of Hate Crimes and Hate Speech Bill: public hearings day 2, NCOP Security and Justice, 20 September 2023, chaired by Ms S Shaikh (ANC, Limpopo).

324 *Ibid.*

© University of Pretoria

to consider rehabilitation as well.[325] Such programmes foster respect for others and would ensure that the punishment fits the crime, promotes justice and fairness whilst discouraging harmful speech. During the Second Day of Public Hearings on the Prevention and Combating of Hate Crimes and Hate Speech Bill, it was mentioned that as a country we prioritise restorative justice, as outlined in the Equality Act.

However, we recognise that severe cases of hate speech may necessitate criminalisation. Any such criminalisation should incorporate restorative justice principles.[326]A suggestion by the FW De Klerk Foundation is also welcomed as they proposed that the bill would contract recommendations of the Rabat Plan of Action and ICERD and that statutory criminal measures should only be considered as a measure of last resort in addressing hate speech.[327]

The author's recommendation is consistent with the report's proposal to prioritise teacher training in human rights values and principles, as well as to introduce or enhance intercultural education programs in schools for all age groups.[328]This would be effective with teaching the public on how to better conduct themselves on and off social media and promote safe online behaviour. As is, a fine and prison sentence are possible responses to transgressions.

The legal frameworks discussed in this dissertation need to be reconstituted to eliminate ambiguities and perceived weaknesses which would in turn result in the need to train legal practioners in cybercrimes, the collection of electronic evidence and the rules that govern the admissibility of digital evidence in judicial proceedings.[329] This suggestion is supported by Sizwe Snail ka Mtuze and Melody Musoni who mention that although South Africa does not possess comprehensive skills associated with cybercrime matters, there should be training of police officers in the gathering of

---

325 Between Speech and Silence, Hate speech, Pornography, and the New South Africa page Shadrack B. Gutto 123

326 Prevention and Combating of Hate Crimes and Hate Speech Bill: public hearings day 2, NCOP Security and Justice, 20 September 2023, chaired by Ms S Shaikh (ANC, Limpopo).

327 *Ibid.*

328 UN High Commissioner for Human Rights, *Report of the United Nations High Commissioner forHuman Rights on the expert workshops on the prohibition of incitement to national, racial, or religious hatred* (UN Doc A/HRC/49/34, 2021).

329 Sizwe Snail ka Mtuze and Melody Musoni, 'An Overview of Cybercrime Law in South Africa' (2023) 4 *International Cybersecurity Law Review* 299 https://doi.org/10.1365/s43439-023-00089-8 321.

© University of Pretoria

electronic and forensic evidence as they are mainly responsible for the initial investigations of cybercrimes.[330]

5.3.3 Collaboration Among Governments, Civil Society, and Social Media Platforms

It is this author's opinion that regulating and monitoring hate speech cannot be solved by one country with a unilateral effort. Given the vast global outreach of social media companies such as Facebook and TikTok it would be impractical for these platforms to establish physical offices in every country in which they operate or to have a strong media presence. In theory, this would be necessary to effectively address issues like hate speech and cultural sensitivities that are contextually unique.

This author suggests that there needs to be a collaboration with regional bodies such as SADC, ECOWAS,[331] EEA and others as this model could leverage shared cultural understandings and legal frameworks to develop tailored solutions to address hate speech and other online harms. Example of this can be seen as the globally recognised company Apple was forced to change charger in Europe as the EU approved its overhaul[332] for its devices. This demonstrates the power of regional cooperation in influencing global tech giants. By working together, regional bodies can exert pressure on tech companies to adhere to local laws and cultural norms. Also, a working group, comprising civil society and academic experts, should review current hate speech laws and determine whether criminalising them is necessary.[333] International human rights law, including the International Covenant on Civil and Political Rights (ICCPR) and the European Convention on Human Rights (ECHR), recognises the right to freedom of expression, but also acknowledges limitations to protect public order, morality and the rights of others.[334]

---

330 *Ibid.*

331 United Nations Economic Commission for Africa, *Regional Economic Communities* (United Nations Economic Commission for Africa, 2023) https://au.int/en/recs (accessed 2/12/2023).

332 As reported by Reuters, Apple was forced to change its charger in Europe due to EU regulations.

333 Prevention and Combating of Hate Crimes and Hate Speech Bill: public hearings day 2, NCOP Security and Justice, 20 September 2023, chaired by Ms S Shaikh (ANC, Limpopo).

334 UN High Commissioner for Human Rights, *Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial, or religious hatred* (UN Doc A/HRC/49/34, 2021).

© University of Pretoria

5.4 Conclusion

The question asked in the beginning is to what extent are the legal frameworks and regulations in South Africa effective in striking a balance between freedom of expression and protection against hate speech? While there is no straightforward answer, this author examined the implications for South Africa, considering its specific legal framework, in order to answer the question.

As it stands the current laws have constitutional compatibility issues namely the threshold tests that are inconsistent and not aligned with international standards and the Constitution that require modification. This is also inclusive of others that regulate the advocacy of hatred may be an obstacle and therefore makes them inadequate in tackling this issue.

As seen from the information above, the complexity and nuanced interplay between hate speech and freedom of expression on social media presents a significant challenge to societies worldwide, including South Africa. International principles such as ICCPR, UNDR and CEDAW, have provided valuable guidance on how to strike a delicate balance between these two fundamental rights; however, it is up to the discretion of governments on how they choose to bring the world into their country and vice versa. This, however, needs to go hand-in-hand with collaboration between all parties involved, such as civil society, technology companies and the relevant stakeholders, in order to develop strategies that are effective. These will identify, report and address hate speech online. This author believes that South Africa has effectively balanced freedom of expression and protection against hate speech. Cases like those of Momberg and Sparrow demonstrate how South African courts have used the law to address hate speech, while upholding the principles of free speech. By carefully considering the specific circumstances of each case and applying appropriate legal remedies, South Africa has shown a commitment to both individual liberties and social harmony. The author believes that South Africa has effectively balanced freedom of expression and protection against hate speech.

[19895 words]

© University of Pretoria

BIBLIOGRAPHY

Articles

Alkiviadou N, 'Hate Speech on Social Media Networks: Towards a Regulatory Framework?' (2019) Information & Communications Technology Law DOI: 10.1080/13600834.2018.1494417.

Altman A, 'Liberalism and Campus Hate Speech: A Philosophical Examination' (Ethics) http://www.jstor.org/stable/2381524 accessed 25 October 2023.

Aswad EM, 'The Future of Freedom of Expression Online' (2018-2019) 17 Duke L & Tech Rev 26.

Black Lives Matter, 'About' https://blacklivesmatter.com/about/ accessed 25 October 2023; Howard University School of Law, 'Black Lives Matter Movement' https://library.law.howard.edu/civilrightshistory/BLM accessed 25 October 2023

Benesch S, 'Dangerous Speech' in Strippel C et al. (eds), Challenges and Perspectives of Hate Speech Research (Digital Communication Research) https://doi.org/10.48541/dcr.v12.11 accessed 11 September 2023

Bloch-Wehba, H., 2019. Global Platform Governance: Private Power in the Shadow of the State. SMU L. Rev.

Campbell, C. M. (1987) 'Law, Technology and Developing Countries', Commw L Bull, 13, pp. 1029.

Chong D, 'Free Speech and Multiculturalism in and out of the Academy' (Political Psychology) http://www.jstor.org/stable/3792382 accessed 25 October 2023.

Claude C.E., 'The African Charter and Freedom of Expression in Africa' (1998) 4 University at Buff Human Rights, L Rev 103. Available at: https://digitalcommons.law.buffalo.edu/bhrlr/vol4/iss1/7/

© University of Pretoria

Coe, P., 'The social media paradox: an intersection with freedom of expression and the criminal law' (2015) 24 Information & Communications Technology Law 16, DOI: 10.1080/13600834.2015.1004242 (accessed: 16 May 2023)

Demaske C, 'Hate Speech' (The Free Speech Center) https://firstamendment.mtsu.edu/article/hate-speech/ accessed 25 October 2023

D Moyo and W Chuma, 'The state of South African media: a space to contest democracy' (2020) 26(2) Journal of International Communication 178-196 https://doi.org/10.1080/13216597.2020.1738827 accessed [13 May 2023].

Dzoguji C, "The great thing about social media was how it gave a voice to voiceless people"- Jon Ronson' (NUHA Foundation 9 November 2019) https://nuhafoundation.org/home/blog/bloggingentries/2019/youth/the-great-thing-about-social-media-was-how-it-gave-a-voice-to-voiceless-people-jon-ronson-2/ accessed 25 October 2023.

Ferrière CA, '#metoo and Twitter: The Feminist Movement on Social Media' (Humans R Social Media, Winter Open Textbook Edition 2022) https://opentextbooks.library.arizona.edu/hrsmwinter2022/chapter/special-topics-me-too/ accessed 25 October 2023

Fowler GA and Alcantara C, 'Gatekeepers: These tech firms control what's allowed online' (The Washington Post) https://www.washingtonpost.com/technology/2021/03/24/online-moderation-tech-stack/ accessed 25 October 2023

FOR SA (Freedom of Religion South Africa) (n.d.), 'Internet censorship? A look at the Films and Publications Amendment Act's criminalisation of hate speech', available at [https://www.forsa.org.za/articles/internet-censorship-a-look-at-the-films-and-publications-amendment-acts-criminalisation-of-hate-speech] (accessed 23 November 2024).

© University of Pretoria

Forsyth, RSA (n.d.), Internet Censorship: A Look at the Films and Publications Amendment Act's Criminalisation of Hate Speech, FORSA, https://journals.sagepub.com/page/ioc/collections/reading-lists/apartheid-era-south-africa, (accessed 23/11/2024)

Garbe L et al., 'How African Countries Respond to Fake News and Hate Speech' (2023) Information, Communication & Society DOI: https://doi.org/10.1080/1369118X.2021.199462

Global Centre for the Responsibility to Protect, 'Hate Speech and Social Media: Preventing Atrocities and Protecting Human Rights Online' (21 February 2020) . Available at: https://www.globalr2p.org/publications/hate-speech-and-social-media-preventing-atrocities-and-protecting-human-rights-online/

Government Communication and Information System, 'Social Media Policy Guidelines' (April 2011). Available at: https://www.gcis.gov.za/sites/default/files/docs/resourcecentre/guidelines/social_media_guidelines_final_20_april2011.pdf

Gumede W, 'Rise in censorship of the Internet and social media in Africa' (2016) Journal of African Media Studies vol. 8. Available at: https://www.researchgate.net/publication/312443019_Rise_in_censorship_of_the_Internet_and_social_media_in_Africa_Journal_of_African_Media_Studies_83_pp413-421>

Hart CE, 'Social Media Law: Significant Developments' (2016) Bus Law. Available at: https://www.jstor.org/stable/26419119

Hagglund, K & Khan, F 2020, 'Reformulating *dolus eventualis*: Guidance from USA and Germany' *South African Crime Quarterly*, vol. 69, no. 69a6108, http://dx.doi.org/10.17159/2413-3108/2020/v0n69a6108 accessed [accessed 01/12/2024].

Harrisberg, K and Bhalla, N. 'As online hate speech hits Africa, social media firms told to act' (IOL, 7 August 2022). available at: https://www.iol.co.za/news/africa/as-online-

© University of Pretoria

hate-speech-hits-africa-social-media-firms-told-to-act-3b809769-6b06-533d-a2cb-d834cdfe2440

Joshua A, Tucker at al., 'From Liberation to Turmoil: Social Media and Democracy' (2017) 28 J Democracy 46. Available at: https://journalofdemocracy.org/articles/from-liberation-to-turmoil-social-media-and-democracy/

Kanamugire JC and Faku OB, 'Employees' Right to Freedom of Expression through Social Media in South Africa' (2022) 11 Persp L Pub Admin 394. Available at: https://journals.indexcopernicus.com/api/file/viewByFileId/1824804

K Schwab, 'The Fourth Industrial Revolution' (Essay, Britannica) https://www.britannica.com/topic/The-Fourth-Industrial-Revolution-2119734 accessed [13 May 2023].

Kovács G et al., 'Challenges of Hate Speech Detection in Social Media' (SN Computer Science) https://link.springer.com/article/10.1007/s42979-021-00457-3 accessed 25 October 2023

Luescher, T., Loader, L. and Mugume, T., '# FeesMustFall: An Internet-age student movement in South Africa and the case of the University of the Free State' (2017) 44(2) Politikon 231-245.

Maart R, (2020) 'Black Bodies on South African Beaches: Lus en smaak jou lekkerding 1' 27 Alternation.

Maclure J, 'The Regulation of Hateful and Hurtful Speech: Liberalism's Uncomfortable Predicament' (McGill Law Journal Volume) https://lawjournal.mcgill.ca/article/the-regulation-of-hateful-and-hurtful-speech-liberalisms-uncomfortable-predicament/ accessed 25 October 2023

Marais M 'Hate Speech in the Equality Act Following the Constitutional Court Judgment in Qwelane v SAHRC' (2023) 26 PER / PELJ http://dx.doi.org/10.17159/1727-3781/2023/v26i0a15438 accessed 16 July 2025

© University of Pretoria

Martins P and Ferrari V, 'Human rights online at the Human Rights Council 47th session' (Association for Progressive Communications) https://www.apc.org/en/news/human-rights-online-human-rights-council-47th-session accessed 25 October 2023.

Meyer Attorneys, (2024), 'Social Media and the Law in South Africa: Navigating the Digital Landscape', available at https://meyerattorneys.co.za/2024/04/05/social-media-and-the-law-in-south-africa-navigating-the-digital-landscape/ (accessed 23/11/2024).

Nyathi, N and Rajuili, K, 'South Africa and Kenya's Legislative Measures to Prevent Hate Speech' (ACCORD, 21 July 2017). Available at: https://journals.co.za/doi/10.10520/EJC-95737f6ea

O'Byrne WI, 'Educate, Empower, Advocate: Amplifying Marginalized Voices in a Digital Society' (CITE Journal vol19 issue4) https://citejournal.org/volume-19/issue-4-19/english-language-arts/educate-empower-advocate-amplifying-marginalized-voices-in-a-digital-society/ accessed 25 October 2023.

Ocampo N et al., 'An In-depth Analysis of Implicit and Subtle Hate Speech Messages' (Proceedings of the EACL) https://aclanthology.org/2023.eacl-main.147 accessed 25 October 2023

Reid A, 'Does Regulating Hate Speech Undermine Democratic Legitimacy? A Cautious 'No'' (*Res Publica*) https://link.springer.com/article/10.1007/s11158-019-09431-6 accessed 25 October 2023.

Rutenberg I and Sugow A, 'Regulation of the Social Media in Electoral Democracies: A Case of Kenya' (2020) SOAS Law Journal. Available at: https://cipit.strathmore.edu/wp-content/uploads/2020/06/Journal-7.pdf

Racist Vicki Momberg hands herself over to police' (2019) SAHRC. https://www.sahrc.org.za/index.php/sahrc-media/news/item/2206-racist-vicki-momberg-hands-herself-over-to-police (accessed 23/11/2024).

© University of Pretoria

Samples J, 'Why the Government Should Not Regulate Content Moderation of Social Media' (Cato Institute) https://www.cato.org/policy-analysis/why-government-should-not-regulate-content-moderation-social-media accessed 25 October 2023

Schwab, K, 'The Fourth Industrial Revolution' (Essay, Britannica) https://www.britannica.com/topic/The-Fourth-Industrial-Revolution-2119734 accessed [13 May 2023]

Sizwe Snail ka Mtuze and Melody Musoni, 'An Overview of Cybercrime Law in South Africa' (2023) 4 *International Cybersecurity Law Review* 299 https://doi.org/10.1365/s43439-023-00089-8

South African Human Rights Commission, 'Court finds government official Velaphi Khumalo guilty of hate speech' (5 October 2018).Available at: https://sahrc.org.za/index.php/sahrc-media/news/item/1633-court-finds-government-official-velaphi-khumalo-guilty-of-hate-speech#:~:text=Velaphi%20Khumalo%2C%20who%20posted%20racial%20slurs%20against%20white,in%20the%20South%20Gauteng%20High%20Court%20on%20Friday

Staff Writer, 'New laws to make hate speech a crime in South Africa – including on Twitter, WhatsApp and other social media' (Business Tech, 15 March 2023).Available at: https://businesstech.co.za/news/trending/672725/new-laws-to-make-hate-speech-a-crime-in-south-africa-including-on-twitter-whatsapp-and-other-social-media/

Takhshid Z, 'Regulating Social Media in the Global South' (2021), Vanderbilt Journal of Entertainment & Technology Law, Volume 24, Issue 1 (2022). U Denver Legal Studies Research Paper No. 21-18. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3836986

The Daily Maverick (2020), 'Film and Publications Amendment Act: Protecting, Not Censoring Our Citizens in the Digital Age?', by Barnaschone Attorneys, available at:https://www.dailymaverick.co.za/article/2020-08-27-film-and-publications-amendment-act-protecting-not-censoring-our-citizens-in-the-digital-age/ (accessed 23 November 2024).

© University of Pretoria

The Future of Free Speech, 'Global Handbook on Hate Speech Laws' https://futurefreespeech.com/global-handbook-on-hate-speech-laws/ accessed 25 October 2023

UN News, 'Independent UN expert says, 'tsunami of hate' targeting minorities must be tackled' (15 March 2021).Available at: https://www.ungeneva.org/en/news-media/news/2021/03/66588/independent-un-expert-says-tsunami-hate-targeting-minorities-must-be#:~:text=Social%20media%20has%20too%20often%20been%20used%20with,an%20international%20treaty%20to%20address%20the%20growing%20scourge

Wiederhold BK, 'Femtech: Digital Help for Women's Health Care Across the Life Span' (Cyberpsychology, Behaviour, and Social Networking) https://www.liebertpub.com/doi/full/10.1089/cyber.2021.29230.editorial a ccessed 25 October 2023.

Wilson RA and Land MK, 'Hate Speech on Social Media: Content Moderation in Context' (Connecticut Law Review) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3690616 accessed 25 October 2023.

Books

Branco Di Fátima, *Hate Speech on Social Media: A Global Approach* (2023) Communication Library vol 7.

Limpitlaw J, *Media Law Handbook for Southern Africa*: Volume 1 (2021) Communication Library.

Cronje, F, Snail, S, Olivier A & Jordaan, J, *Cybercrime in South Africa: Law and Practice* (LexisNexis Butterworths, 2022).

Milo, D and Stein, P, A Practical Guide to Media Law (2nd Edition, Lexis Nexis 2014)

S Papadopoulos and Snail S Cyberlaw @ SA IV: The Law of the Internet in South Africa9 4TH edition, Van Schaik Publishers 2022)

© University of Pretoria

Guillén-Nieto V *Hate Speech: Linguistic Perspectives* (De Gruyter Inc 2023) https://ebookcentral-proquest-com.uplib.idm.oclc.org/lib/pretoria-ebooks/detail.action?docID=7177394 accessed 10 July 2025.

Duncan, J (ed), *Between Speech and Silence: Hate Speech, Pornography and the New South Africa* (Freedom of Expression Institute and Institute for Democracy in South Africa 1996)

Legislation and International Instruments

African Charter on Human and Peoples Rights (ACPHR)

Constitution of the Republic of South Africa Act 108 of 1996 (RSA).

South Africa, 'Cybercrimes Act' (Act 19 of 2020).

South African Human Rights Commission, 'Hate Speech Information Sheet' (2023).

South Africa, 'Prevention and Combating of Hate Crimes and Hate Speech Bill' (B9-2018).

Schedule 1 substituted by section 24(5) of the *Protection of Constitutional Democracy against Terrorism and Related Activities Amendment Act*, 2022 (Act No. 23 of 2022), Notice No. 1533, G47803, dated 29 December 2022, effective 4 January 2023 per Proclamation Notice 110, GG47820, dated 4 January 2023

International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171.Available at: https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights

Office of the High Commissioner for Human Rights, 'Towards an interpretation of article 20 of the ICCPR: Thresholds for the prohibition of incitement to hatred' https://www.ohchr.org/Documents/Issues/Expression/ICCPR/Santiago/Toby Mendel.pdf accessed 4 September 2023.

© University of Pretoria

Office of the High Commissioner for Human Rights, 'Training Manual on Human Rights Monitoring' (OHCHR, 2001) https://www.ohchr.org/sites/default/files/Reporting-ICCPR-Training-Guide.pdf accessed 4 September 2023.

United Nations, 'International Convention on the Elimination of All Forms of Racial Discrimination' https://legal.un.org/avl/ha/cerd/cerd.html accessed 25 October 2023.

United Nations, 'Convention on the Elimination of All Forms of Discrimination against Women' https://www.un.org/womenwatch/daw/cedaw/ accessed 25 October 2023

United Nations Office on Genocide Prevention and the Responsibility to Protect, 'United Nations strategy and plan of action on hate speech: Detailed guidance on implementation for United Nations field presences. Available at: https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20PoA%20on%20Hate%20Speech_Guidance%20on%20Addressing%20in%20field.pdf

Universal Declaration of Human Rights (adopted 10 December 1948) UNGA Res 217 A(III) (UDHR)Available at: https://www.ohchr.org/en/universal-declaration-of-human-rights

Case Law

AfriForum NPC & Another v Pienaar 2017 (1) SA 388 (WCC).

AfriForum NPC v Nelson Mandela Foundation Trust and Others (371/2020) [2023] ZASCA 58; 2023 (4) SA 1 (SCA); [2023] 3 All SA 1 (SCA) (21 April 2023).

Amaning and Others v Ackerman (EQ3/2023) [2024] ZAEQC 2

Dawood and Another v Minister of Home Affairs and Others; Shalabi and Another v Minister of Home Affairs and Others; Thomas and Another v Minister of Home Affairs and Others [2000] ZACC 8 (7 June 2000).

Democratic Alliance v African National Congress and Another 2015 (2) SA 232 (CC)

Momberg v S (A206/2018) [2019] ZAGPJHC 183; 2019 (2) SACR 505 (GJ) (28 June 2019).

© University of Pretoria

Sedick and Another v Krisray (Pty) Ltd (2011) 8 BALR 879 (CCMA).

S v Khumalo [1973] 3 SA 697 (A).

South African Human Rights Commission v Khumalo 2019 1 SA 289 (GJ)

South African National Defence Union v Minister of Defence and Another 1999 (4) SA 469 CC

Qwelane v South African Human Rights Commission and Another [2019] ZACC 22.

Zulu v S (189/2011) [2012] ZAKZPHC 26 (4 May 2012).

Cape South African Jewish Board of Deputies v Torr EC11/2025(9 May 2025).

Thesis and Dissertations

Botha, J *Hate Speech as a Limitation to Freedom of Expression* (PhD, Nelson Mandela Metropolitan University 2016).

Davis T, Social Media Content Moderation and Limitations on Freedom of Expression: The Role of the State, (2022). Unpublished dissertation submitted to the Law Faculty, University of the Witwatersrand.

Nyathi, N, The Poverty of Law: A Critical Analysis of Hate Speech Jurisprudence in South Africa (unpublished PhD thesis, University of Pretoria, 2018)

Paulos B and Çelik S, “The Challenges of Regulating Hate Speech on Social Media in Light of the Theory of Freedom of Expression” (Süleyman Demirel Üniversitesi Hukuk Fakültesi Dergisi).

Van Vollenhoven, W. J. (n.d.), *Learners’ Understanding of Their Right to Freedom of Expression in South Africa* (PhD thesis, University of Pretoria).

© University of Pretoria

Other

Disfold Data, 'Largest 1000 US Technology Companies in 2023' https://disfold.com/united-states/sector/technology/companies/ accessed 25 October 2023

Facebook, 'Hate Speech' https://www.facebook.com/communitystandards/hate_speech accessed 10 October 2023; TikTok, 'Community Guidelines' https://www.tiktok.com/community-guidelines?lang=en accessed 10 October 2023

Google 'Hate Speech Policy' (n.d.) available at https://support.google.com/contributionpolicy/answer/11412392?hl=en (accessed 10/10/2023).

Legal Desire, 'Role of Social Media and Freedom of Speech and Expression' (Legal Desire, 2021) https://legaldesire.com/role-of-social-media-and-freedom-of-speech-and-expression/ accessed 18 September 2023

Merriam-Webster, 'Social Media' (Merriam-Webster, 16 September 2023) https://www.merriam-webster.com/dictionary/social media

Montreal Holocaust Museum, 'Der Stürmer Antisemitic Newspaper' https://museeholocauste.ca/en/objects/der-sturmer-antisemitic-newspaper/ accessed 25 October 2023.

N Ocampo, E Sviridova, E Cabrio & S Villata "An In-depth Analysis of Implicit and Subtle Hate Speech Messages" (2023) Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics) available at https://aclanthology.org/2023.eacl-main.147 accessed 25 October 2023.

Prevention and Combating of Hate Crimes and Hate Speech Bill: public hearings day 2, NCOP Security and Justice, 20 September 2023, chaired by Ms S Shaikh (ANC, Limpopo).

P Martins & V Ferrari "Human Rights Online at The Human Rights Council 47th Session" (2021) Association for Progressive Communications available at

© University of Pretoria

https://www.apc.org/en/news/human-rights-online-human-rights-council-47th-session (accessed 25/10/2023).

Statista, 'Number of internet users worldwide 2022' https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/ accessed 27 September 2023.

TikTok 'Community Guidelines' (n.d.) available at https://www.tiktok.com/safety/community-guidelines (accessed 10/10/2023).

Twitter 'Hateful Conduct Policy' (n.d.) available at https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy (accessed 10/10/2023).

© University of Pretoria