



POLICY BRIEF 1/2025

Hate Speech Regulation in Africa: Overview and Current Issues

Hate speech is a form of harmful expression that undermines democratic values and social cohesion.¹ It disproportionately affects minorities, who are the primary targets of incitement to hatred and discrimination.² It is also employed to spread misinformation and cross-border disinformation campaigns.³

While there is no unified definition of hate speech under international human rights law (IHRL),⁴ states across the globe adopt different legal frameworks to regulate it. These approaches are shaped by the tension between limiting hate speech and protecting the right to freedom of expression. The rise of dominant digital platform companies has deepened these dilemmas, disrupting traditional models of internet governance. Africa's *digital divide*⁵ represents an additional challenge in the region.

This policy brief corresponds to points three and four of *the Strategy for Promoting Freedom of Expression in Norwegian Foreign and Development Policy* and to the *Norwegian Guidelines for Freedom of Religion or Belief*. It highlights key regulatory gaps and challenges at the African regional level, provides an overview of South Africa's current regulatory framework on hate speech, and outlines relevant geopolitical developments.

Key points

- Hate speech regulation can advance the protection of human rights for vulnerable populations; however, in the context of authoritarian regimes and weak democracies, these same instruments may be misused to suppress political opposition and curtail freedom of expression.
- The absence of effective content moderation and the control of online information by dominant digital platform companies pose significant challenges for combating hate speech.
- There are emerging regional frameworks aiming to guide African states in defining the responsibilities of digital platform companies for addressing harmful online content.
- South Africa's legal hate speech framework is multilayered, combining constitutional, civil, and criminal provisions. The latest legislative development is the *Hate Crimes and Hate Speech Bill* signed in 2024, by which hate speech is codified as a criminal offense.
- Legal discussions on the criminalization of hate speech in South Africa center on its potential impacts on freedom of expression and Freedom of Religion or Belief (FoRB).

visitors:
Villa Grande
Huk aveny 56
Oslo, Norway

postal address:
PO Box 1168 Blindern
NO-0318 Oslo
Norway

telephone: 22 84 21 00
facsimile: 22 84 21 45
www.hlsenteret.no

Authors: Karoll Andrea Ciceri & Vibeke Moe

e-mail: andrea.ciceri@hlsenteret.no
vibeke.moe@hlsenteret.no

Context

Online hate leads to real-life violence.⁶ It fuels polarization manifesting in forms of group hostility such as xenophobia, antisemitism, Islamophobia and gender-based violence (GBV).

Internationally, state obligations to prohibit hate speech are established in *Article 20* of the *International Covenant on Civil and Political Rights* (ICCPR) and *Article 4* of the *International Convention on the Elimination of All Forms of Racial Discrimination* (ICERD).

Key documents provide guidance for states and stakeholders. The *Rabat Plan of Action* (2012), developed by the *Office of the UN High Commissioner for Human Rights* (OHCHR), sets out a six-part threshold to determine when an expression constitutes incitement to hatred and can be criminalized.⁷ Additionally, the *UN Strategy and Plan of Action on Hate Speech* (2019) aims to coordinate the UN system's response to the societal impact of hate speech and to address its root causes.⁸

At the same time, as social media has become a primary channel for disseminating hate speech, misinformation, and disinformation, different legal frameworks have been adopted to address these information-related harms. In Europe, the *Digital Services Act* (2022) imposes due diligence obligations on online platforms to foster safer digital environments, including mechanisms for removing illegal content such as hate speech. By contrast, in Africa, while instruments such as the *Malabo Convention* (2014) address issues of cybersecurity and data protection, there is currently no binding regional instrument defining the responsibilities of social media platforms for moderating hate speech.

Moreover, across the region, responses to hate speech vary based on the national **political regime** in place. Authoritarian states tend to rely on **technological approaches**, such as internet shutdowns, to moderate harmful content online, whereas democratic states are more likely to adopt **legal approaches**.⁹ Nevertheless, the *African Commission on Human and Peoples' Rights* (ACHPR) remains cautious about measures that explicitly criminalize hate speech, as such legislation could impose excessive restrictions on freedom of expression and be misused to suppress political dissent.

About this document

The content presented in this policy brief draws on information collected from *Key Informant Interviews* (KIIs). The participants are representatives of organisations located in South Africa and across the African continent. Interviewees include participants from the **South African Human Rights Commission (SAHRC)**, the **Special Rapporteur on Freedom of Expression of the African Union**, as well as representatives from Civil Society Organizations (CSOs) with expertise in hate speech legislation and advocacy. The insights from these interviews form the basis of the report's background and are supplemented by relevant literature in the field.

Hate Speech in the Digital Space

Dominant digital platform companies have compliance obligations under frameworks addressing illegal online content, including hate speech. However, their limited self-regulation presents implementation challenges. This underscores the importance of adopting a **multi-stakeholder co-regulatory approach** to internet governance.¹⁰ Core related issues in Africa are:

Access to information: The absence of information-sharing by platforms presents a significant limitation in the enforcement of hate speech regulations in Africa. National competent authorities and interested parties can require access to platform-held data to substantiate legal claims for hate speech. However, **jurisdictional constraints**, combined with the **lack of local data storage** and **on-the-ground presence of digital platform companies**, present significant obstacles.

Moderation in local languages: As algorithmic moderation based on *large language models* (LLMs) is primarily trained on dominant languages, social media platforms fail to effectively moderate hate speech when expressed in local languages. This gap enables hate speech and inflammatory content to bypass automated detection.¹¹ For example, in South Africa, *Meta*, *TikTok*, *YouTube*, and *X* have allowed the publication of violent content in *Zulu*, *Xhosa*, and *Afrikaans* targeting vulnerable groups, including GBV against women journalists.¹²

Additionally, when requested to disclose the number of moderators proficient in these languages or to provide evidence of how effectively their AI systems detect harmful content, these companies have failed to respond, indicating a lack of policy transparency.

Misalignment between social media policies and hate speech regulation: Law enforcement of hate speech prohibitions can be undermined by platform-specific policies that apply different thresholds for content removal. For instance, in South Africa, the gratuitous display of the apartheid-era flag has been declared to constitute hate speech since 2019. Yet in 2025, *Meta's Oversight Board* ruled that two Facebook posts showing the flag during the 2024 elections did not meet its removal criteria.¹³ Furthermore, Meta appears to grant greater protection to statements made by political figures, although political speech does not enjoy special legal protection under national law.¹⁴

Emerging Regional Frameworks

In response to the growing influence of dominant digital platform companies in spreading harmful content, the ACHPR has tasked the *African Union's Special Rapporteur on Freedom of Expression* with developing different guidelines to assist member states in the regulation of major digital platforms:

Resolution 620 of 2024: Calls for the creation of standards for accountable data collection including online information.

Resolution 630 of 2025: Mandates the creation of state guidelines for information integrity and independent fact-checking online.

Resolution 631 of 2025: Calls for the creation of a focused regional policy framework for access to information as public interest content.

The three guidelines are expected to be launched between late 2025 and early 2026. Resolution 630 represents a significant advancement in the regional response to hate speech, since information integrity is essential for effective content moderation.

Building on these initiatives, the task ahead for African states will be to consider the adoption of new domestic legislation that addresses information challenges in the digital era while safeguarding freedom of expression.

National Regulation of Hate Speech in South Africa

South Africa is a constitutional democracy with strong protections for political rights and civil liberties;¹⁵ nonetheless, the country's legacy of apartheid continues to create social divisions, contributing to **political polarization**.¹⁶ These divisions are deepened by groups such as *Operation Dudula*, a vigilante movement which uses xenophobic rhetoric and online mobilization to incite violence, as well as by the dissemination of hate narratives, including the claim of an alleged ongoing genocide against white people ("white genocide").¹⁷ Within this context, the state maintains a multilayered national framework governing offline and online hate speech, grounded in its constitution:

Constitutional Provisions: Article 16(1) of the South African Constitution establishes the fundamental right to *freedom of expression*. However, this right is not absolute and is interpreted in conjunction with the substantive rights of *dignity* (Section 10) and *equality* (Section 9).¹⁸ This balance is reflected in Article 16 (2), which explicitly excludes hate speech from constitutional protection, understood as "advocacy of hatred that is based on (1) race, (2) ethnicity, (3) gender or (4) religion, and that constitutes **incitement to cause harm**".¹⁹

Civil Law Mechanisms: *The Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000* (PEPUDA) is the primary civil law instrument regulating hate speech in South Africa. It provides a framework to combat *unfair discrimination*, *hate speech*, and *harassment*, while also expanding the four constitutionally prohibited grounds for hate speech to include among others, sexual orientation, disability, belief.²⁰ Since 2023, following a legislative amendment, for an expression to constitute hate speech under PEPUDA, it needs to be: "(1) harmful or to incite harm, and (b) to promote or propagate hatred".²¹

Civil remedies include public apologies or financial compensation for physical or emotional harm and in severe instances, the referral of a case to the *National Prosecuting Authority* for criminal proceedings.

Criminal Law Mechanisms: Under *common law*, through court-development rules, a person may be charged with *crimen injuria* if their conduct intentionally violates another person's dignity or privacy. Prosecuted cases under this figure have included antisemitic, Islamophobic and racist expressions.²²

The most recent legislative development, however, is the **Prevention and Combating of Hate Crimes and Hate Speech Act**. Signed into law in 2024, the Act was adopted by South Africa to comply with its international human rights obligations on hate speech prohibition. It also responds to voluntary pledges made during the third cycle of the **Universal Periodic Review**, with **Norway** among the recommending states.²³

The new law explicitly codifies expressions of hate speech as a criminal offense punishable by a fine *and/or* imprisonment of up to 5 years. Importantly, the provision also includes an exception clause protecting *bona fide* religious, artistic, academic, and public interest expressions.²⁴ In terms of its status, while signed, the bill has **not yet become operational**, and to date no cases have been prosecuted under this framework.

Current Issues

Some CSOs in South Africa have welcomed the adoption of the new Act's hate crime provisions as an important step toward strengthening national protection for vulnerable groups, including migrants, refugees, and members of the LGBTQIA+ community.²⁵ However, others raise concerns about the Bill's hate speech provisions and their potential implications for freedom of expression. These concerns focus on two main issues:

1. Criminalization of hate speech vs implementation of existing frameworks: Before the adoption of the new legislation, PEPUDA already addressed incitement to harm, and *crimen injuria* covered expressions that intentionally violate human dignity. This has prompted debates on whether the criminalization of hate speech represents a **proportionate restriction** on the right to freedom of expression, given that less severe measures are available.²⁶ From this perspective, critics argue that government priority should not be given to creating new regulation, but to strengthening the implementation of the existing framework under PEPUDA.

Complementarily, the current framework is viewed as largely reactive, with insufficient preventive educational measures, particularly for **media and information literacy (MIL)** initiatives, to address root causes before harmful speech escalates to incitement. The **digital divide** further stresses this need, as individuals can lack the critical digital skills needed for responsible

engagement on social media platforms.

2. Freedom of Religion and Belief: The new law adds “social detriment” as a dimension of **harm** in hate-speech offences, defined as harm that “undermines the social cohesion of South Africa”.²⁷ During parliamentary debates, religious CSOs expressed that this definition could invite subjective judicial interpretations, potentially criminalizing legitimate religious expression and restricting their FoRB.²⁸ The extent to which this will take place will be clear once the law is implemented.

Policy Recommendations

- Prioritize funding initiatives in Africa designed to foster media and information literacy (MIL) as a preventive measure to combat hate speech. While these programs should be accessible online, they need also to be available in offline formats mindful of the *digital divide* in the continent.
- Support civil society initiatives that monitor and document hateful expressions, disaggregated by language and vulnerable groups, to combat hate speech.
- Facilitate national and regional multistakeholder platforms bringing together civil society, government representatives, and dominant digital platform companies to discuss hate speech regulation and develop co-regulatory models. These should promote hiring local language moderators and establishing national offices for digital platform companies.
- Support regional initiatives led by the *ACHPR* through *Resolutions 620, 630, and 631* to advance human rights-based approaches to digital governance in Africa.
- Recommend to South Africa, in the context of the 5th Universal Periodic Review (UPR) cycle, to ensure the effective implementation of PEPUDA and the Prevention and Combating of Hate Crimes and Hate Speech Act, as well as the adoption of a national law regulating dominant digital platform companies if no such provisions exist at the time.

Notes

- 1 Norwegian Ministry of Foreign Affairs. 2021. “Press Statement on Countering Hate Speech.”
- 2 United Nations Human Rights Council. 2021. “Report of the Special Rapporteur on minority issues, Fernand de Varennes, (A/HRC/46/57).”
- 3 NATO Strategic Communications Centre of excellence. 2025. “Russian Information Operations Outside the West Information Environment.”
- 4 United Nations. 2020. “United Nations Strategy and Plan of Action on Hate Speech.”
- 5 Norwegian Ministry of Foreign Affairs. 2021. “Freedom of Expression, Strategy for promoting freedom of expression in Norwegian foreign and development policy.” p. 23
- 6 Rousset, Thierry; Maluleke, Gavaza; and Mendelsohn, Adam. 2022. “The Dynamics of Racism, Antisemitism and Xenophobia on Social Media in South Africa. Johannesburg: *Konrad-Adenauer-Stiftung*.”
- 7 United Nations General Assembly. 2013. “Annual report of the United Nations High Commissioner for Human Rights.”
- 8 United Nations. 2020. “United Nations Strategy and Plan of Action on Hate Speech.”
- 9 Garbe, Lisa, Lisa-Marie Selvik, and Pauline Lemaire. 2021. “How African Countries Respond to Fake News and Hate Speech.” In *Information, Communication & Society* 26 (1): 86–103.
- 10 Stockman, Daniela. 2023. “Tech companies and the Public interest: the role of the state in governing social media platforms.” In *Information, Communication*
- 11 Rousset, Thierry; Maluleke, Gavaza; and Mendelsohn, Adam. 2022. “The Dynamics of Racism, Antisemitism and Xenophobia on Social Media in South Africa. Johannesburg: *Konrad-Adenauer-Stiftung*.”
- 12 Legal Resources Centre. 2023. “Facebook, X/Twitter, YouTube and TikTok approve violent misogynistic hate speech adverts for publication in South Africa.”
- 13 Oversight Board. 2025. “Posts Displaying South Africa’s Apartheid-Era Flag.”
- 14 Fisher, Sarah, Beatriz Kira, Kiran Arabaghata Basavaraj, and Jeffrey Howard. 2024. “Should Politicians Be Exempt from Fact Checking?” *Journal of Online Trust & Safety* 2 (2).
- 15 Freedom House. 2025. “Freedom in the World 2025.”
- 16 Adonis, Cyril Kenneth. 2025. “‘The Crown without the Jewels’: South Africa’s Flawed Transition and Its Implications for Social Justice and Equity in Post-Apartheid South Africa.” *Memory Studies*.
- 17 Rousset, Thierry; Maluleke, Gavaza; and Mendelsohn, Adam. 2022. “The Dynamics of Racism, Antisemitism and Xenophobia on Social Media in South Africa. Johannesburg: *Konrad-Adenauer-Stiftung*.”
- 18 Oukouomi Noutchie, Suare Clovis. 2025. “Between Expression and Equality: Constitutional Challenges in Adjudicating Racially Charged Political Speech in South Africa.” In *International Journal of Research in Business and Social Science* 14 (5): 507–14.
- 19 South African Constitutional Assembly. 1996. “The Constitution of the Republic of South Africa.”
- 20 Parliament of the Republic of South Africa. 2000. “Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000.”
- 21 Parliament of the Republic of South Africa. 2000. “Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000.”
- 22 South African Human Rights Commission. 2023. “Social Media Charter.”
- 23 UPR Info. “Database.”
- 24 Parliament of the Republic of South Africa. 2018. “Prevention and Combating of Hate Crimes and Hate Speech Bill.”
- 25 Hate Crimes Working Group. 2023. “Media Statement: Passage of the Prevention & Combating of Hate Crime & Hate Speech Bill.”
- 26 Venter, Roxan. 2025. “Dignity, Freedom of Expression and the Battle over Hate Speech: A Case Study in Post Apartheid South Africa.” In *Probing Human Dignity*, Springer.
- 27 Parliament of the Republic of South Africa. 2018. “Prevention and Combating of Hate Crimes and Hate Speech Bill.”
- 28 Parliamentary Monitoring Group. 2018. “Summary of Submissions to Select Committee on Security and Justice: Prevention and Combating of Hate Crimes and Hate Speech Bill.”

The authors

Karoll Andrea Ciceri is a research assistant at the Norwegian Center for Holocaust and Minority Studies. She has a background in legal studies and holds a Master’s degree in Human Rights and Multiculturalism from the University of South-Eastern Norway.

Vibeke Moe is a senior researcher at the Norwegian Center for Holocaust and Minority Studies. She holds a PhD in Religious Studies from the University of Oslo. Moe’s research explores various forms of prejudice and hate, with particular emphasis on religious minorities and the social and societal consequences of hate speech.

The Norwegian Center for Holocaust and **Minority Studies** is a research, education and documentation center in Oslo focusing on the Holocaust, other genocides and the situation of minorities in contemporary societies.



HL-SENTERET
The Norwegian Center for Holocaust and Minority Studies

