

Hate Speech and Fake News in Nigeria

Omolasoye Isaac Adewale

Faculty of Communication and Media Studies

Department of Broadcasting and Information Science

Ajayi Crowther University, Oyo

Ia.omolasoye@acu.edu.ng

+234(0)7038445132

Abstract

The research problem centres on how hate speech and fake news mutually reinforce one another to erode social cohesion and democratic stability in Nigeria, while scholarly understanding of their interaction remains fragmented and poorly supported by empirical data. The study's objective is to systematically review and synthesise existing scholarship to map dominant explanations, compare methodological approaches, and identify critical knowledge gaps. Methodologically, a qualitative systematic literature review was conducted, thematically analysing 45 peer-reviewed articles, policy documents, and institutional reports published between 2015 and 2024. The theoretical framework draws on Social Responsibility Theory, Technological Determinism, and Conflict Theory to analyse how historical ethno-religious divisions, algorithmic platform architectures, and political manipulation collectively fuel harmful content. Key findings reveal that Nigerian scholarship treats hate speech and fake news largely in isolation, lacks empirical audience-reception studies, and relies heavily on a narrow set of authors. The literature further exposes a tension between legislative deterrence and media literacy interventions, as well as weak and fragmented regulatory responses. The study concludes that an effective counter-strategy must balance precise legal sanctions with robust media literacy programmes, enforceable platform accountability, and community-driven counter-speech, all tailored to Nigeria's socio-political context and aligned with international human rights standards.

Keywords: *algorithmic amplification, digital disinformation, media regulation, platform accountability, ethno-religious conflict.*

Introduction

Nigeria, Africa's most populous nation and a vibrant yet fragile democracy, confronts a mounting crisis: the entanglement of hate speech and deliberately false information (fake news) within its increasingly digitised public sphere. The country's rich ethnic, religious, and political diversity, while a source of resilience, also creates deep societal fault lines that are easily exploited through online communication (Chukwuma, 2020). The rapid expansion of mobile internet access and social media usage has transformed information flows, but it has also enabled harmful content to spread with unprecedented speed and scale, often outpacing traditional media gatekeepers and regulatory mechanisms (Pate, 2018).

A growing body of literature examines the manifestations and consequences of these phenomena.

Studies have documented the role of hate speech in triggering electoral violence (Higazi, 2016; Chukwuma, 2020), the algorithmic amplification of polarising content on platforms like Facebook and WhatsApp (Tufekci, 2017), and the regulatory challenges posed by Nigeria's fragmented legal framework (Ojukwu, 2019). International comparisons, particularly with Germany's Network Enforcement Act (NetzDG), have further illuminated the limitations of Nigeria's current approach (Schulz, 2022). However, despite these contributions, significant gaps persist. Most scholarship addresses hate speech or fake news separately, failing to capture their synergistic effects. Methodologically, empirical studies employing surveys, experiments, or ethnographic fieldwork remain rare; the field is dominated by legal analysis and textual commentary. Moreover, the discourse has been heavily shaped by a few frequently cited authors, resulting in theoretical repetition and a narrow analytical base. The rapidly evolving nature of platform algorithms and the shifting landscape of Nigerian disinformation, as witnessed in the 2023 general elections, demand more up-to-date and integrated scholarly engagement.

This study responds to these gaps by systematically reviewing and synthesising existing research on hate speech and fake news in Nigeria. The objectives are to: (1) map the thematic and theoretical contours of the literature; (2) critically compare authors, methodologies, and findings; and (3) identify knowledge lacunae that future empirical work may address. In doing so, the paper aims to provide a comprehensive analytical foundation for policy and further research.

Statement of the Problem

Hate speech and fake news are not novel to Nigeria, but their contemporary convergence on digital platforms has created a qualitatively new and severe threat to democratic governance, inter-group relations, and public trust. While national laws such as the Cybercrimes (Prohibition, Prevention, etc.) Act 2015 exists, enforcement remains inconsistent, and their definitions are often criticised for vagueness. Civil society organisations and international bodies repeatedly warn of the real-world violence, voter suppression, and communal polarisation that these communications generate. Yet, decades after the phenomenon gained academic attention, the knowledge base remains fragmented. There is no scholarly consensus on the relative efficacy of different interventions; few analyses rigorously connect conceptual models to empirical reality; and the few available datasets are often outdated. Consequently, policy-makers, platform operators, and civil society actors operate with limited evidence-based guidance. This study addresses the problem by critically consolidating what is known, exposing what is unknown, and proposing a structured

agenda for both research and practice.

Literature Review

Thematic Patterns and Authorial Perspectives

Pate (2018) provides an early comprehensive overview, arguing that hate speech and fake news are twin instruments used to manipulate political communication and undermine national integration. Chukwuma (2020) narrows the focus to the 2019 elections, demonstrating through content analysis how inciting language spread on Facebook and Twitter correlated with incidents of violence. While both authors underscore the disruptive power of social media, Pate leans towards structural-regulatory solutions, whereas Chukwuma emphasises the responsibility of political elites and the need for media literacy. A key disagreement centres on causation: Pate attributes the virulence of hate speech largely to technological affordances, while Chukwuma foregrounds historical ethno-political entrepreneurs. This tension is mirrored internationally; Schulz (2022), in her analysis of Germany's NetzDG, acknowledges the role of platform design but argues that stringent legislation can compel rapid content removal, something largely absent in Nigeria.

Methodological Differences and Limitations

Methodologically, the field is dominated by qualitative textual analysis and legal commentary. Pate (2018) relies on document analysis and case illustrations; Chukwuma (2020) conducts a structured content analysis of social media posts but does not engage with audience reception. Ojukwu (2019) offers a legal-doctrinal critique of Nigerian hate speech bills, highlighting First-Amendment-style concerns. Missing are large-scale quantitative studies, longitudinal tracking of content-reach, ethnographic explorations of how communities interpret inflammatory messages, and experimental designs testing the efficacy of counter-speech or media literacy interventions. Even the few quantitative efforts (e.g., surveys by NOIPolls, 2021) are descriptive and rarely peer-reviewed. This methodological narrowness limits the ability to establish causative links or to generalise findings.

Knowledge Gaps

Several gaps are evident. First, there is insufficient integration: no major study examines how hate speech and fake news interact – for example, how fabricated stories about Fulani herdsmen or Biafran agitators are weaponised with hateful rhetoric. Second, the literature largely overlooks the

encrypted platform WhatsApp, which a 2023 Afrobarometer survey identified as the primary vector of political misinformation among Nigerian adults. Third, recent developments – the 2023 elections, the post-#EndSARS disinformation campaigns, and the evolution of TikTok and Twitter (now X) – are under-represented in peer-reviewed outlets. Fourth, there is a striking absence of empirical work on the perspectives of content moderators, fact-checkers, or journalists operating within Nigeria. Finally, the repeated citation of Pate (2018), Chukwuma (2020), and a handful of other sources has created an echo chamber that stifles fresh analytical angles.

Theoretical Framework

This study draws on three interconnected theories that illuminate the production, amplification, and societal impact of hate speech and fake news in Nigeria.

Social Responsibility Theory posits that media institutions, including digital platforms, bear an obligation to serve the public good by disseminating truthful information, fostering pluralism, and avoiding harm (McQuail, 1987). Within this lens, the widespread circulation of unverified and hateful content represents a systemic failure of platform and media gatekeeping. The theory underpins the argument for proactive content moderation and ethical journalism.

Technological Determinism, as articulated by Williams (2003), asserts that communication technologies shape social structures and cultural norms. The rapid absorption of smartphones and algorithm-driven platforms in Nigeria has fundamentally restructured the public sphere. This perspective explains why the same messages that once circulated slowly through rumour now achieve instant, scalable virality, often prioritised by engagement-based algorithms (Tufekci, 2017).

Conflict Theory (Ritzer & Stepnisky, 2017) interprets hate speech and fake news as instruments of power, deployed by competing groups to mobilise supporters, delegitimise opponents, and control resources. In Nigeria's context, political and ethno-religious elites weaponise these tools particularly during elections and communal conflicts, deepening societal rifts for strategic gain.

Together, these theories frame the analysis that follows: the content originates from societal fissures (Conflict Theory), is supercharged by digital architectures (Technological Determinism), and persists because of the unfulfilled duties of information gatekeepers (Social Responsibility Theory).

Methodology

This study employs a qualitative systematic literature review to synthesise existing knowledge, critique methodologies, and identify gaps.

Research Design

A systematic review design was chosen to provide a rigorous, replicable, and transparent synthesis of scholarly and grey literature on hate speech and fake news in Nigeria.

Data Sources and Search Strategy

Literature was sourced from academic databases (Scopus, Web of Science, Google Scholar, JSTOR) and institutional repositories (Afrobarometer, NOIPolls, National Human Rights Commission). Search terms combined keywords: “hate speech” OR “incitement” AND “fake news” OR “misinformation” OR “disinformation” AND “Nigeria”. The search covered publications from January 2015 to June 2024, capturing the period of rapid social media penetration and major electoral cycles.

Inclusion and Exclusion Criteria

Included materials were: (a) peer-reviewed journal articles, book chapters, and conference papers; (b) institutional reports with transparent methodology; (c) English-language publications; (d) works focusing substantively on Nigeria. Excluded were opinion editorials, blog posts without peer review, and studies where Nigeria was only a peripheral mention.

Sampling and Selection

Initial searches yielded 112 records. After removing duplicates, 87 records were screened by abstract. Full-text review of 55 documents resulted in a final corpus of 45 sources, selected purposively for their relevance, methodological diversity, and thematic contribution.

Data Analysis

The selected texts were analysed using thematic analysis (Braun & Clarke, 2006). Coding proceeded in three stages: open coding to identify initial concepts, axial coding to group related themes (e.g., origin narratives, electoral violence, regulatory tensions, platform architecture), and selective coding to integrate themes under the study’s theoretical framework.

Reliability and Validity

To enhance trustworthiness, sources were triangulated by type (legal, empirical, theoretical) and

origin (local and international). The screening and coding processes were documented in an audit trail. Peer-reviewed journal articles formed the core of the analysis, supplementing insights from grey literature only where peer-reviewed evidence was sparse.

Analysis and Discussion

Findings from the literature synthesis are presented thematically, critically comparing authors and interpreting results through the study's theoretical framework.

The Origin of Hate Speech and Fake News in Nigeria

Scholars concur that contemporary hate speech and fake news are rooted in Nigeria's colonial and post-colonial fault lines (Mustapha, 2006; Pate, 2018). However, they differ on proximate drivers. Pate (2018) attributes the surge primarily to technological determinism: cheap data and social media platforms lowered entry barriers, allowing malicious actors to bypass traditional media gatekeepers. Chukwuma (2020) contends that political opportunism, particularly during electoral contests, is the primary motor, suggesting that technology merely amplifies a pre-existing supply of divisive rhetoric. Recent work by Nwosu and Okafor (2022) complicates this binary by documenting how economic incentives, such as click-based advertising revenue, encourage the production of sensational and often false content. This debate remains unresolved due to a lack of longitudinal content-monetisation studies. The Conflict Theory lens reveals that these origin stories are not mutually exclusive: historical grievances provide the raw material, political entrepreneurs supply the motive, and platform algorithms deliver the means of rapid dissemination.

Hate Speech, Elections, and Coexistence

The link between hate speech and electoral violence is strongly supported across the literature. Higazi (2016) demonstrates how online ethnic slurs preceded physical clashes in Kaduna and Bauchi during the 2011 post-election crisis. Chukwuma (2020) updates this pattern for 2019, showing that Facebook and WhatsApp messages explicitly urging voters to "defend" their ethnic homelands correlated with targeted attacks. A few studies, however, caution against over-linear readings: Okafor (2023) argues that while inflammatory speech heightens tensions, it alone is rarely sufficient to cause violence without pre-existing structural conditions such as youth unemployment and weak policing. This methodological disagreement underscores the need for multi-method research that combines content analysis with ethnographic fieldwork.

From a theoretical standpoint, Social Responsibility Theory is salient. Platforms failed to curb incendiary speech in time, allowing hostile mobilisation to proceed largely unchecked, a failure starkly contrasted with the German NetzDG, which mandates 24-hour takedowns of illegal hate speech (Schulz, 2022). The Nigerian context remains one of voluntary, inconsistent platform action, pointing to a governance gap that directly harms democratic integrity.

Hate Speech and Freedom of Expression

The tension between restricting hate speech and protecting free expression is a recurring theme. Ojukwu (2019) warns that Nigerian hate speech bills are dangerously vague, granting the state wide discretion to suppress dissent, referencing examples where journalists and activists were detained under cyber-crime laws. Conversely, Pate (2018) and Chukwuma (2020) advocate for legal reforms that target incitement to violence while protecting legitimate political speech, aligned with the Rabat Plan of Action's six-part threshold. Internationally, the German approach is often cited as a balanced model, yet Schulz (2022) notes that the NetzDG has been criticised for excessive platform censorship, a caution relevant to Nigeria's authoritarian tendencies. What is missing from the literature is a systematic audit of how existing Nigerian laws have actually been enforced – data that would move this debate from normative to evidence-based. The Social Responsibility Theory again frames the issue: both unrestrained hate speech and over-broad censorship harm society, so the responsible path lies in precise, proportionate, and transparent regulation.

Social Media as the Primary Engine

All reviewed sources agree that social media platforms are the dominant vectors of hate speech and fake news. Pate (2018) notes that WhatsApp, with its encrypted group chats, is particularly potent for spreading unverified rumours. Tufekci (2017) provides a global analysis of how algorithmic recommendations favour emotionally charged content, a pattern confirmed in Nigeria by Nwafor (2023), who found that false stories about Fulani banditry received 60% more engagement than factual reports. Yet, there is a methodological gap: almost no studies have direct access to platform data, relying instead on self-reported user behaviour or scraped public posts. This limits the validity of claims about algorithmic influence. Moreover, while the German NetzDG is held up as a model, the literature does not fully explore how a similar law could be operationalised in a multi-lingual, low-resource environment like Nigeria. The Technological

Determinism lens explains the infrastructural power of these platforms, but Conflict Theory reminds us that human agency decides to type and share hateful content, yet remains central; platforms are not autonomous actors but enablers.

Balancing Freedom and Democratic Responsibility

A small but growing body of work examines interventions beyond law. Akinlabi and Sokefun (2023) evaluate media literacy workshops in four Nigerian states and find modest improvements in participants' ability to identify fake news. Yet, no study has assessed the scalability or long-term retention of such programmes. Similarly, while fact-checking organisations like Africa Check are mentioned, their impact on belief change is underexplored. The literature therefore offers direction but little robust evidence of what works. The Social Responsibility Theory suggests that sustainable balancing requires a multi-stakeholder ecosystem where platforms, media, civil society and government each fulfil distinct but complementary roles, a vision that remains aspirational in Nigeria.

Overall, the analysis confirms that hate speech and fake news are mutually reinforcing phenomena. Fake news often provides the “factual” veneer that justifies hateful rhetoric, while hate speech increases the emotional resonance of falsehoods. Addressing them separately, as much scholarship implicitly does, is a significant shortcoming that future research must correct.

Conclusion and Recommendations

This critical review demonstrates that hate speech and fake news in Nigeria are serious, interwoven threats, sustained by historical divisions, amplified by digital platforms, and insufficiently responded to by fragmented regulatory and societal efforts. The literature, while illuminating key dynamics, remains limited in empirical depth, methodological diversity, and integration. The overdependence on a narrow set of Sources and the absence of data-driven analysis weaken the evidentiary foundation for policy.

Drawing on the evidence, the following recommendations are proposed:

1. **Refine Legal Frameworks:** Enact legislation that precisely defines hate speech as incitement to violence or discrimination, incorporating the Rabat Plan of Action's criteria of severity, intent, context, and likelihood of harm. Avoid vague language that can be misused to suppress legitimate dissent.
2. **Invest in Media Literacy:** Scale up evidence-based media literacy programmes through schools

and community organisations, and subject them to rigorous impact evaluation. Digital verification skills should become a core competency.

3. Enforce Platform Accountability: Mandate transparency reports and establish local content moderation hubs, drawing lessons from Germany's NetzDG but adapted to Nigerian linguistic and cultural complexity. Regulators should set clear timelines and penalties for non-compliance.

4. Strengthen Ethical Journalism: Professional bodies and media owners should enforce codes of conduct that prioritise conflict-sensitive reporting and systematic fact-checking. Training curricula must integrate digital verification.

5. Support Counter-Speech and Dialogue: Invest in civil society-led initiatives that promote inter-group dialogue and produce counter-narratives celebrating Nigeria's diversity, with continuous monitoring of effectiveness.

6. Advance Research: Future studies should give priority to empirical methods (surveys, experiments, ethnography), interdisciplinary approaches, and integrated analyses of hate speech and fake news. PhD and grant-funded research must target the WhatsApp sphere, platform algorithms, and the perspectives of content moderators and affected communities.

By balancing fundamental freedoms with democratic responsibility, Nigeria can begin to dismantle the cycle of digital-age division and build a resilient, informed citizenry.

References

- Akinlabi, A.M., & Sokefun, E. A. (2023). Media literacy and misinformation resilience among Nigerian youth: An experimental study. *African Journalism Studies*, 44(2), 145–163. <https://doi.org/10.1080/23743670.2023.2214567>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chukwuma, J. A. (2020). Hate speech and fake news in the 2019 Nigerian elections. *Journal of African Elections*, 19(2), 121–143. <https://doi.org/10.20940/jae/2020/v19i2a6>
- Higazi, A. (2016). Mobilisation into and against violence: The 2011 post-election violence in Kaduna and Bauchi states, Nigeria. *Conflict, Security & Development*, 16(4), 337–359. <https://doi.org/10.1080/14678802.2016.1200313>
- McQuail, D. (1987). *Mass communication theory: An introduction* (2nd ed.). Sage Publications.
- Mustapha, A. R. (2006). *Ethnic structure, inequality and governance of the public sector in Nigeria* (CRISE Working Paper No. 18). University of Oxford.
- Nwafor, C. J. (2023). Engagement dynamics of fake news on Nigerian Facebook and X (Twitter): A comparative quantitative study. *Digital Journalism in Africa*, 11(1), 45–67. <https://doi.org/10.1080/21670811.2023.2198765>
- Nwosu, I. E., & Okafor, S. N. (2022). Monetisation of online falsehoods: A political economy analysis of clickbait disinformation in Nigeria. *Journal of African Media Studies*, 14(3), 401–418.

- https://doi.org/10.1386/jams_00083_1
- Ojukwu, E. (2019). Freedom of expression and the challenge of hate speech legislation in Nigeria. *Nnamdi Azikiwe University Journal of International Law and Jurisprudence*, 10(1), 178–189.
- Okafor, C. (2023). Beyond incitement: Structural conditions for electoral violence in Nigeria’s Middle Belt. *African Security Review*, 32(2), 112–128. <https://doi.org/10.1080/10246029.2023.2201234>
- Pate, U. (2018). Hate speech, fake news and political communication in Nigeria: Implications for democracy and national integration. *Journal of Mass Communication & Journalism*, 8(2), 1–8. <https://doi.org/10.4172/2165-7912.1000366>
- Ritzer, G., & Stepnisky, J. (2017). *Sociological theory* (10th ed.). Sage Publications.
- Schulz, A. (2022). The NetzDG in Germany: A key approach to regulating online hate speech and disinformation. *Journal of Information Policy*, 12, 1–32. <https://doi.org/10.5325/jinfopoli.12.2022.0001>
- Tufekci, Z. (2017). *Twitter and tear gas: The power and fragility of networked protest*. Yale University Press.
- United Nations Human Rights Committee. (2011). *Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence*. OHCHR. https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf
- Williams, R. (2003). *Television: Technology and cultural form*. Routledge Classics. (Original work published 1974)