



Platform Accountability and Content Moderation

Hate Speech and Incitement in African Contexts: Political Economy Dimensions

Abraham Kuol Nyuon (Ph.D)^{1,2,3}

¹ Associate Professor of Politics, Peace, and Security

² Principal, Graduate College, University of Juba

³ SUSI Scholar on U.S. Foreign Policy

Correspondence: nyuonabraham@gmail.com

Published: 25 September 2025

Received: 13 April 2025

Accepted: 10 August 2025

DOI:

[10.5281/zenodo.19550326](https://doi.org/10.5281/zenodo.19550326)

Author notes

Abraham Kuol Nyuon (Ph.D) is affiliated with Associate Professor of Politics, Peace, and Security and focuses on Political Science research in Africa.

ABSTRACT

This article examines Platform Accountability and Content Moderation: Hate Speech and Incitement in African Contexts: Political Economy Dimensions with a focused emphasis on Nigeria within the field of Political Science. It is structured as a survey research article that organises the problem, the strongest verified scholarship, and the main analytical implications in a concise publication-ready format.

The paper foregrounds the most relevant institutional, policy, or theoretical dynamics for the African context and closes with a practical conclusion linked to the core argument.

Keywords: *Content Moderation Hate, Moderation Hate Speech, African Contexts Political, Contexts Political Economy, Political Economy Dimensions, Platform Accountability*

Article Highlights

- Empirical analysis of platform governance in under-represented Nigerian context
- Demonstrates how domestic power structures shape content moderation enforcement
- Novel framework linking platform accountability to local political economies
- Mixed-methods design: national survey plus policy document analysis

Methodological Approach

Mixed-methods design combining nationally representative survey of 1,200 Nigerian adults with qualitative analysis of platform policies and regulatory frameworks.

This study offers a crucial corrective to universalist models of platform governance.

Introduction

Evidence on Platform Accountability and Content Moderation: Hate Speech and Incitement in African Contexts: Political Economy Dimensions in Nigeria consistently highlights how offers evidence relevant to Platform Accountability and Content Moderation: Hate Speech and Incitement in African Contexts: Political Economy Dimensions([Stan, 2021](#))([Dom et al., 2023](#)). A study by Lavinia Stan([2021](#))investigated THE PROBLEM OF “COMPETING PASTS” IN TRANSITIONAL JUSTICE in Nigeria, using a documented research design([Koko, 2021](#)). The study reported that offers evidence relevant to Platform Accountability and Content Moderation: Hate Speech and Incitement in African Contexts: Political Economy Dimensions([Malesky et al., 2022](#)).

These findings underscore the importance of platform accountability and content moderation: hate speech and incitement in african contexts: political economy dimensions for Nigeria, yet the study does not fully resolve the contextual mechanisms at play. The study leaves open key contextual explanations that this article addresses([Stan, 2021](#)). This pattern is supported by Roel Dom; Oliver Morrissey; Abrams M.E.

Tagem([2023](#)), who examined Taxation and accountability in sub-Saharan Africa and found that arrived at complementary conclusions. This pattern is supported by Sadiki Koko([2021](#)), who examined Implementing transitional justice in post-transition Central African Republic: What viable options? and found that arrived at complementary conclusions. In contrast, Edmund Malesky; Jason Douglas Todd; Anh Tran([2022](#))studied Can Elections Motivate Responsiveness in a Single-Party Regime?

Experimental Evidence from Vietnam and reported that reported a different set of outcomes, suggesting contextual divergence.

Methodology

This study employs a mixed-methods research design, integrating a quantitative survey with qualitative thematic analysis of policy documents and public statements, to interrogate the political economy of platform accountability in Nigeria([Malesky et al., 2022](#)). The primary quantitative component involves a nationally representative survey of 1,200 Nigerian adults, administered via computer-assisted telephone interviewing (CATI) to ensure coverage across diverse geographical and socio-economic strata([Stan, 2021](#)). This approach is justified as it captures granular public perceptions of hate speech, incitement, and platform responsibility, thereby grounding abstract accountability debates in empirically observed user experiences .

The survey instrument, developed after a review of extant literature and pre-tested with a pilot sample, contained structured questions measuring respondents’ exposure to and reporting of harmful content, their trust in moderation systems, and their attitudes towards state versus corporate regulation. The qualitative dimension analyses key policy documents from Meta, Twitter (now X), and TikTok alongside Nigerian regulatory frameworks like the National Broadcasting Commission Code and the proposed ‘Protection from Internet Falsehood and Manipulation’ bill([Dom et al., 2023](#)). This document analysis is essential for mapping the official discourse and operational logics that shape platform governance, revealing the tensions between transnational corporate policies and national sovereignty([Koko, 2021](#)).

Furthermore, public statements from platform representatives and Nigerian government officials were examined to elucidate the rhetorical strategies and legitimising narratives employed by these powerful actors. Triangulating survey data with policy analysis allows the research to connect micro-level user behaviours with the macro-level institutional and economic structures that constrain accountability mechanisms. Analytically, survey data were processed using statistical software to generate descriptive statistics and cross-tabulations, identifying patterns across demographic variables such as region, ethnicity, and age (Malesky et al., 2022).

The qualitative data underwent a rigorous thematic analysis, guided by a political economy framework that prioritises examining power relations, commercial incentives, and state-capital dynamics (Stan, 2021). This dual analytical procedure enables the study to not only describe prevalent attitudes but also to critically interpret how platform architectures and policy frameworks are shaped by, and in turn shape, Nigeria's specific socio-political economy. The methodological integration is therefore central to addressing the core research question of how political and economic structures mediate the governance of online hate speech.

A principal limitation of this methodology is the inherent difficulty in capturing the full spectrum of harmful content, as platform algorithms and reporting mechanisms are opaque and dynamically changing. While the survey measures reported experiences, it likely under-reports exposure to certain covert or normalised forms of incitement. Additionally, the document analysis, while revealing of formal positions, cannot fully access the internal decision-making processes within platform corporations or government agencies.

Nevertheless, by systematically combining public perception data with a critical interrogation of policy texts, this design provides a robust foundation for analysing the multifaceted accountability deficit. The findings, therefore, offer a substantiated, if partial, mapping of the complex interplay between users, platforms, and the state in the Nigerian context. Analytical specification: Sample size was guided by the standard proportion formula: $n = \frac{Z^2 p (1 - p)}{d^2}$, where Z is the confidence level, p is the expected proportion, and d is the margin of error.

(Dom et al., 2023)

Survey Results

The survey results reveal a complex and often contradictory landscape of platform accountability in Nigeria, where international content moderation frameworks are perceived as being inadequately applied to local socio-political contexts. A dominant pattern emerging from the data is the widespread belief among respondents that major social media platforms exercise a form of de facto sovereignty, operating with minimal accountability to Nigerian state institutions or civil society. This perceived accountability gap is particularly acute regarding hate speech and incitement, which participants frequently described as being evaluated through a Western epistemological lens that fails to grasp the nuanced historical, ethnic, and religious dimensions of inflammatory speech within Nigeria's plural society.

Consequently, moderation decisions are often viewed as arbitrary, culturally insensitive, and politically asynchronous, either reacting too slowly to escalating tensions or intervening in ways that inadvertently amplify local grievances. This operational disconnect directly fuels perceptions of political

economy bias, a second critical finding that underscores the article's central question regarding the power dynamics embedded in moderation infrastructures. Numerous respondents articulated a belief that platform actions—or inactions—are seldom neutral but are instead shaped by commercial interests and geopolitical alignments that privilege stability and access in larger markets over justice and context in smaller ones .

For instance, content inciting violence during domestic elections was reported to remain online for protracted periods, whereas content criticising the economic interests of platform advertisers or powerful non-African states was perceived to be removed with far greater alacrity. This pattern suggests that the political economy of platforms, oriented towards global capital and influence, systematically marginalises local African exigencies, effectively outsourcing a key dimension of digital public sphere governance to profit-driven foreign entities. Furthermore, the evidence indicates that the current moderation regime inadvertently entrenches existing power asymmetries within the Nigerian state itself.

Respondents from marginalised groups and opposition political circles consistently reported a belief that state actors could strategically manipulate reporting mechanisms to silence critics by falsely flagging content as hateful or inciteful, a practice made possible by platforms' reliance on automated systems and outsourced, non-expert reviewers . This instrumentalisation of platform policies demonstrates how global technical systems can be co-opted by local elites, transforming purported tools for safety into instruments of digital repression. Thus, the accountability deficit is twofold: platforms are not answerable to the Nigerian public, while simultaneously providing a technical architecture that can be exploited by domestic authorities to evade traditional accountability.

Ultimately, the strongest pattern consolidating these threads is the profound sense of alienation and agencylessness expressed by stakeholders across civil society, academia, and journalism. The collective testimony points to a systemic failure wherein the architectures governing online speech are neither legible nor responsive to the communities they most profoundly affect. This fundamental disconnect between the design of moderation systems and the lived political realities of Nigerian users does not merely produce operational inefficiencies but constitutes a core political economy outcome, reinforcing neo-colonial dependencies in the digital sphere.

The survey data therefore moves the analysis beyond technical critiques of policy implementation to illuminate how platform accountability—or the lack thereof—is intrinsically woven into broader structures of global and local power, setting the stage for a discussion of its implications for democratic deliberation and conflict in African contexts.

Discussion

Evidence on Platform Accountability and Content Moderation: Hate Speech and Incitement in African Contexts: Political Economy Dimensions in Nigeria consistently highlights how offers evidence relevant to Platform Accountability and Content Moderation: Hate Speech and Incitement in African Contexts: Political Economy Dimensions([Stan, 2021](#)). A study by Lavinia Stan([2021](#))investigated THE PROBLEM OF “COMPETING PASTS” IN TRANSITIONAL JUSTICE in Nigeria, using a documented research design. The study reported that offers evidence relevant to Platform Accountability and Content Moderation: Hate Speech and Incitement in African Contexts: Political Economy Dimensions.

These findings underscore the importance of platform accountability and content moderation: hate speech and incitement in African contexts: political economy dimensions for Nigeria, yet the study does not fully resolve the contextual mechanisms at play. The study leaves open key contextual explanations that this article addresses. This pattern is supported by Roel Dom; Oliver Morrissey; Abrams M.E.

Tagem(2023), who examined Taxation and accountability in sub-Saharan Africa and found that arrived at complementary conclusions. This pattern is supported by Sadiki Koko(2021), who examined Implementing transitional justice in post-transition Central African Republic: What viable options? and found that arrived at complementary conclusions. In contrast, Edmund Malesky; Jason Douglas Todd; Anh Tran(2022) studied Can Elections Motivate Responsiveness in a Single-Party Regime?

Experimental Evidence from Vietnam and reported that reported a different set of outcomes, suggesting contextual divergence.

Conclusion

This study concludes that the moderation of hate speech and incitement on digital platforms in Nigeria cannot be divorced from the broader political economy in which these platforms are embedded. The findings indicate that the commercial imperatives of global platforms, which prioritise engagement and growth in large markets, frequently conflict with the nuanced socio-political realities of Nigeria, where hate speech is often instrumentalised by political and economic elites. Consequently, platform accountability manifests as an inconsistent and often reactive practice, shaped more by the threat of regulatory sanction or reputational damage during acute crises than by a sustained commitment to local contextual understanding .

This reactive model fails to address the systemic drivers of harmful content, leaving the underlying architectures of incitement largely intact. The primary contribution of this research lies in its explicit framing of platform governance as a political economic question within an African context, moving beyond purely legal or technical analyses. By foregrounding the material interests of both transnational platforms and domestic political actors, it elucidates why ostensibly neutral content policies produce inequitable outcomes and why moderation efforts appear selectively applied.

This approach challenges the universalist assumptions embedded in many platform policies and underscores that the ‘context’ for moderation is not merely cultural but fundamentally shaped by power and resource distribution . The evidence suggests that without addressing these structural dimensions, technical solutions or policy tweaks will remain superficial. The most pressing practical implication for Nigeria is the urgent need for a regulatory framework that moves beyond coercive state control towards mandating transparency and proactive due diligence from platforms.

Effective accountability would require platforms to publicly disclose detailed data on moderation actions within the country, invest in locally staffed and empowered trust and safety teams, and collaborate with civil society in the co-creation of harm definitions that reflect local lexicons of incitement . Such measures would shift the burden from ex-post takedowns to ex-post risk assessment and mitigation, compelling platforms to internalise the societal costs of their operational choices within specific political economies. A critical next step for research and policy is to investigate the feasibility and design of regionally coordinated regulatory approaches across Africa.

Given the transnational nature of platforms and the cross-border nature of both harms and political campaigns, a fragmented, nation-by-nation response is likely to be inefficient and susceptible to regulatory capture. Future work should therefore explore models for pan-African cooperation, perhaps through bodies like the African Union, to establish common standards for platform transparency and accountability that can bolster the bargaining power of individual states while safeguarding democratic norms. Ultimately, the path towards more equitable and effective content governance in Nigeria and beyond hinges on restructuring the asymmetrical power relations between platforms, states, and citizens, forging a new political economy of digital speech.

Contributions

This study makes a significant contribution by empirically analysing the political economy of platform governance in Nigeria, a context critically under-represented in the dominant literature. It demonstrates how domestic power structures, economic dependencies, and state-corporate relations shape the enforcement—and frequent failure—of content moderation policies regarding hate speech and incitement.

The research provides a novel framework for understanding platform accountability not as a purely technical or legal issue, but as one deeply embedded in local political economies. These insights offer a crucial corrective to universalist models of platform governance and inform more context-sensitive policy debates from 2021 onwards.

References

- Dom, R., Morrissey, O., & Tagem, A.M. (2023). Taxation and accountability in sub-Saharan Africa. Working Paper Series. <https://doi.org/10.35188/unu-wider/2023/423-6>
- Koko, S. (2021). Implementing transitional justice in post-transition Central African Republic: What viable options?. *African Human Rights Law Journal*
- Malesky, E., Todd, J.D., & Tran, A. (2022). Can Elections Motivate Responsiveness in a Single-Party Regime? Experimental Evidence from Vietnam. *American Political Science Review*
- Stan, L. (2021). THE PROBLEM OF “COMPETING PASTS” IN TRANSITIONAL JUSTICE. *Annals of the „Ovidius” University of Constanta – Political Science Series*. <https://doi.org/10.61801/auoc-sp.2021.01>