2026 Volume: 02 Issue: 06 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-12

AI-ASSISTED SOCIAL MEDIA USAGE POLICIES AND CYBERBULLYING REDUCTION AMONG STUDENTS IN PUBLIC SECONDARY SCHOOLS IN RIVERS STATE, NIGERIA

***Ojujoh, Womemi, Nwachukwu, Ruth Michael**

Department of Educational Management, Faculty of Education University of Port Harcourt.

Article Received: 29 April 2026
Article Revised: 19 May 2026
Published on: 09 June 2026

***Corresponding Author: Ojujoh, Womemi**
Department of Educational Management, Faculty of Education University of Port Harcourt.

DOI: https://doi-doi.org/101555/ijrpa.1153

ABSTRACT

This study investigated the relationship between AI-assisted social media usage policies and cyberbullying reduction among students in public secondary schools in Rivers State, Nigeria. It was guided by two research questions, and two null hypotheses were tested. A correlational design was used. The population included 5,833 teachers from 320 public secondary schools in Rivers State, with a sample of 583 teachers representing 10% of the population, selected through stratified random sampling. A validated 4-point rating scale questionnaire, developed by experts, served as the data collection instrument. The reliability of the instrument was determined using Cronbach's Alpha, with coefficients of 0.79 for the AI-Assisted Social Media Usage Policies Assessment Scale (AASMUPAS) and 0.77 for the Cyberbullying Reduction Among Students Assessment Scale (CRASAS). Pearson Product-Moment Correlation (PPMC) statistics were used to answer the research questions and to test the hypotheses at a 0.05 level of significance. Findings revealed a high positive and significant relationship between content monitoring, parental control and teachers' productivity in public secondary schools in Rivers State. Based on the findings, it was concluded that the implementation of these AI-driven policies is essential for schools seeking to safeguard learners and foster a safer digital learning environment. Consequently, it was recommended, among others, that school administrators and policymakers should make content monitoring a mandatory component of AI-assisted social media usage policies in all public secondary schools, as it has been proven highly effective in significantly reducing cyberbullying among students.

KEYWORDS: Artificial Intelligence (AI), Social Media, Policies, Cyberbullying Reduction.

INTRODUCTION

The proliferation of social media has fundamentally reshaped how adolescents communicate, learn, and socialise. While these platforms offer educational and social benefits, they also pose serious risks, particularly cyberbullying, which can result in emotional distress, diminished academic performance, and long-term psychological harm (Akeusola, 2023; Wordu,Dan-Jumbo, & Mina, 2021). Cyberbullying, defined as repeated harm inflicted via digital means against vulnerable individuals, poses serious risks to pupils' mental health, academic achievement, and social well-being. Research indicates that around 55% of pupils report experiencing cyberbullying at some point during their schooling (Stone, 2025). This underscores the urgent need for effective interventions within school environments, as traditional reactive policies often fail to keep pace with evolving digital threats.

In response to these digital threats, policymakers and educators have begun crafting social media usage policies tailored to school environments, seeking to delineate appropriate conduct and define the boundaries of acceptable engagement. Some jurisdictions have enhanced anti-bullying codes of conduct to explicitly include AI-generated misuses, such as deepfakes and manipulated imagery, recognising their potential for psychological harm and reputational damage (Mani, School District Superintendent, 2024). These policy adaptations reflect growing awareness that conventional definitions of cyberbullying must evolve in the AI era. In other words, Artificial Intelligence (AI) promises hope in detecting and mitigating cyberbullying.

Artificial Intelligence (AI) has emerged as a powerful tool for proactive cyberbullying detection and management in educational settings. It refers to the ability of computer systems to perform tasks that typically require human intelligence, such as problem-solving, learning, reasoning, and language understanding. It enables machines to analyse data, recognise patterns, and make informed decisions with minimal human intervention (Russell & Norvig, 2021). Modern AI applications range from virtual assistants like Siri and chatbots to autonomous vehicles and medical diagnosis tools, reflecting its rapid advancement in recent years (Haenlein & Kaplan, 2019). AI continues to evolve through machine learning and deep learning models, which allow systems to improve their accuracy and efficiency over time.

Modern AI-powered monitoring systems can scan across school-issued platforms like Google Drive, e-mails, and social media, and detect concerning behaviours by analysing both content and context, allowing timely de-escalation and support interventions. Advanced models now

utilise stream-based machine learning combined with Large Language Models (LLMs) to process social media data in real time, detecting abusive behaviours while providing explainability via dashboards to build trust and transparency (Milosevic, 2022). Another promising approach includes user-specific severity detection, which considers demographic, psychological, and behavioural attributes along with the content to classify bullying severity accurately (achieving up to 98% accuracy) and enhance interpretability using SHAP and LIME techniques (Milosevic, 2023).

Simultaneously, the rapid growth of artificial intelligence technologies presents unprecedented opportunities for both detection and deterrence of harmful behaviours. Platforms now increasingly rely on AI-driven tools that scan text, images, and videos to flag abuse or harassing content, offering new modalities for proactive intervention (British School Manila, 2025). Major social media companies, like Instagram, have experimented with classifiers that detect bullying in comments and visual media and even prompt users to reconsider borderline messages before posting, an approach that blends technology with behavioural nudging (TIME, 2019). AI on its own is insufficient unless woven into formal social media usage policies within schools. Effective AI-assisted policies should describe how AI tools are embedded, defining monitoring thresholds, flagging protocols, human review workflows, and escalation paths for severe incidents (Prama et al., 2025). According to Rees, Sabia and Kumpas (2022) and García-Méndez and De-Arriba-Pérez (2025), some of the AI-assisted policies are: content monitoring and filtering, user behaviour analytics, cybersecurity and privacy protection, cyberbullying detection and reporting mechanisms, digital citizenship and ethical use, parental control/stakeholder access, policy compliance and enforcement. However, this study focused on content monitoring and parental control/stakeholders' access.

Content monitoring, as an AI-assisted social media usage policy, refers to the continuous use of AI-powered algorithms to automatically scan, detect, and flag inappropriate, harmful, or abusive content shared on social media platforms as part of a school policy to promote safe digital interactions among pupils. As a core aspect of AI-assisted social media usage policies, content monitoring enables schools to proactively identify cyberbullying behaviours such as threats, hate speech or harassment, often before major psychological harm occurs (García-Méndez & De Arriba-Pérez, 2025; Stone, 2025). Several studies have demonstrated that when effectively implemented with human supervision, AI-based content monitoring significantly contributes to early intervention and a reduction in cyberbullying incidents in school environments (Milosevic, 2023; Nee, 2023). However, empirical reports from some

school contexts also highlight negative outcomes, including false positives, over-surveillance, and violation of pupil privacy, which may create fear or mistrust among pupils (Gaggle Software, 2025; Milosevic, 2022). In situations where monitoring systems are poorly structured or biased, they have been found to unfairly target certain pupil groups or miscode slurs and evolving slang known as algospeak (Smith, 2020; Gencoglu, 2020). Consequently, while AI-based content monitoring remains a promising tool for cyberbullying reduction in schools, its effectiveness depends on transparent policy guidelines, explainability mechanisms, and an ethical balance between pupil safety and digital rights (García-Méndez & De Arriba-Pérez, 2025; Milosevic, 2022).

Parental control, viewed broadly as tools, permissions, and oversight mechanisms enabling guardians and key stakeholders to monitor or influence children's online interactions, has become increasingly salient within AI-enhanced social media ecosystems. Understood as a policy of stakeholders' access, it encompasses everything from content filtering, time limits, contact restrictions to AI-driven moderation aids (e.g., supervised account features, chat filters) (Achuthan, 2022). On a positive note, empirical research demonstrates that higher perceived parental control significantly predicts lower adolescent cyberbullying behaviours, as shown in structural models across different cultural contexts such as Singapore and India ($\beta = .068$ and $\beta = .192$, respectively; $p < .01$) (Cassidy, 2018). Complementing this, solutions such as next-generation parental controls that enhance children's digital competence and cybersecurity awareness, by coupling oversight with education, show promise in reducing risky online behaviours. Educational studies also highlight that active parental engagement, dialogue, supervision of device usage, and school-parent collaboration can foster prevention of cyberbullying and promote 'cyber-kindness' among peers (Gnanasekaran, 2025). Additionally, well-designed, age-appropriate parental control tools, when clear, usable, and matched to developmental stages, can guard children from harmful content while preserving beneficial internet access (Helfrich, Doty,Su,Jacqlyn, & Gabrielli, 2020).

However, the policy also carries negative ramifications. Overly stringent or opaque control mechanisms risk undermining parent–child trust and adolescent autonomy, potentially prompting concealment of online misbehaviour rather than open communication. For example, qualitative insights reveal that top-down monitoring may be perceived by youths as a signal that 'my parents don't trust me,' prompting efforts to bypass restrictions and leading to reluctance to seek parental help when confronting disturbing online content (Milovidov& Oxford Research, 2025). From a privacy and security standpoint, many parental control technologies introduce elevated risks: some tools have security vulnerabilities that could be

exploited to leak sensitive data or enable adversarial control, ironically magnifying cyber-risks rather than reducing them (Ali,Elgharabawy, Duchaussoy, Mannan, & Youssef, 2020). Further, critiques of traditional parental control models suggest a shift towards community-based safety approaches, where oversight is extended beyond immediate family to include trusted community or school stakeholders, an approach that respects teen autonomy while embedding supportive networks, highlighting the limitations of control-centric models (Akter, Park, & Wisniewski, 2025; U.S. Department of Commerce, 2024).

Consequently, the above two policies discussed are essential to ensure stakeholders (teachers, parents, administrators) understand AI's decisions, reinforcing trust and accountability. Such policies must also be adaptive, capable of evolving with emerging threats, such as AI-enabled deepfakes that can be used for harassment and extortion. Indeed, incidents of AI-generated non-consensual explicit images among pupils have surged, prompting jurisdictions like South Australia to enact legal penalties (up to four years' imprisonment or fines up to $20,000) and mandate a deeper understanding and prevention of AI misuse in schools. AI-assisted policies should not stand in isolation. A comprehensive framework for internet safety includes proactive measures such as 'Safety by Design,' digital literacy initiatives, regulatory compliance, and empowerment tools to help pupils manage their online presence responsibly. Educational interventions should promote positive digital citizenship, teaching pupils about ethics, empathy, and consent in digital spaces, particularly in light of AI's rising role (Wordu et al., 2021).

Public secondary schools are uniquely positioned at the intersection of policy, technological capability, and pupil vulnerability. While technical AI solutions show promise, there's a notable gap in empirical research examining how AI-assisted social media usage policies relate to cyberbullying rates and pupil well-being within real-world school contexts. Hence, this study examined the topic 'AI-assisted social media usage policies and cyberbullying reduction among students in public secondary schools in Rivers State.'

Statement of the Problem

The rising cases of cyberbullying among students, amplified by unregulated AI-assisted social media interactions, have generated widespread concern in Nigerian public secondary schools. Many schools currently lack clear policies to guide students' use of AI-driven social media tools, thereby increasing the likelihood of online harassment, misinformation, and other harmful digital behaviours. As a result, parents, educators, and policymakers are becoming increasingly alarmed by the psychological and academic consequences associated

with cyberbullying, making the development of AI-assisted social media usage policies a critical public issue that demands urgent attention.

Despite the mounting concerns over adolescent cyberbullying in Nigeria, little empirical attention has been directed toward the potential of AI-assisted social media usage policies, specifically, content monitoring and AI-enabled parental control tools, to mitigate these harmful behaviours among secondary school students in Rivers State. Existing research underscores the broader utility of AI in detecting and preventing online harassment through automated identification of abusive content, while studies within Nigerian settings highlight the critical role of parental mediation, demonstrating how authoritative parenting correlates with lower rates of cyberbullying.

Additionally, interventions employing content filtering in Nigeria have shown preliminary effectiveness in reducing exposure to harmful online material, yet their integration within AI frameworks and their direct impact on reducing cyberbullying in schools remain untested. These gapsare further contextualised by global policy analyses advocating for data-driven, age-appropriate AI tools to detect and curb online abuse. Consequently, the central problem this study addresses is: What is the nature of the relationship between AI-assisted social media usage policies, operationalised through content monitoring and parental control mechanisms, and the prevalence of cyberbullying among public secondary school students in Rivers State?

Aim and Objectives of the Study

This study aimed to examine the relationship between AI-assisted social media usage policies and cyberbullying reduction among pupils in public secondary schools in Rivers State. Specifically, the objectives of the study sought to:

1. Ascertain the relationship between content monitoringand cyberbullying reduction among students in public secondary schools in Rivers State.
2. Determine the relationship between parental controland cyberbullying reduction among students in public secondary schools in Rivers State.

Research Questions

The following research questions guided the study:

1. What is the relationship between content monitoring and cyberbullying reduction among students in public secondary schools in Rivers State?

2. What is the relationship between parental control and cyberbullying reduction among students in public secondary schools in Rivers State?

Hypotheses

The following hypotheses were tested at the 0.05 level of significance.

1. There is no significant relationship betweencontent monitoring and cyberbullying reduction among students in public secondary schools in Rivers State.
2. There is no significant relationship betweenparental control and cyberbullying reduction among students in public secondary schools in Rivers State.

METHODOLOGY

A correlational research design was adopted to examine the relationship between the independent variable (AI-assisted Social Media Usage Policies) and the dependent variable (Cyberbullying Reduction Among Students). The study population consisted of 5,833 teachers (2,844 male and 2,989 female) across 320 public senior secondary schools in Rivers State (Rivers State Senior Secondary School Board, 2024). A sample of 583 teachers, representing 10% of the population, which is acceptable for large populations according to Nwane in Kpee (2015), was selected using stratified sampling. Twenty-three schools were drawn using simple random technique from six Local Government Areas across the three senatorial districts (Rivers West, Rivers East, and Rivers South East) to ensure equal representation.Two sets of structured questionnaires served as instruments: the AI-Assisted Social Media Usage Policies Assessment Scale (AASMUPAS) and the Cyberbullying Reduction Among Students Assessment Scale (CRASAS). Each questionnaire comprised Sections A (demographic data) and B (items on AASMUPAS and CRASAS), using a four-point Likert scale (Strongly Agree to Strongly Disagree). The reliability of the instruments was confirmed using Cronbach's Alpha, yielding coefficients of 0.79 and 0.77, respectively. Data were analysed using the Pearson Product-Moment Correlation (PPMC) to answer research questions and test hypotheses at a 0.05 significance level. Out of 583 questionnaires distributed, 552 were duly completed and valid for analysis, giving a retrieval rate of approximately 95%.

RESULTS AND FINDINGS

Research Question 1: What is the relationship between content monitoring and cyberbullying reduction among students in public secondary schools in Rivers State?

Table 1: Pearson Product Moment Correlation (PPMC) showing the Relationship between Content Monitoring and Cyberbullying Reduction Among Studentsin Public Secondary Schools in Rivers State.

Variables	n	df	r	Decision
Content Monitoring Cyberbullying Reduction Among Students	552 552	550	0.611	High Relationship

Decision Rule: 0.00 – 0.19 = Very Low, 0.20 – 0.39 = Low, 0.40 – 0.59 = Moderate, 0.60 – 0.79 = High, 0.80 – 1.00 Very High

To answer research question 1, Table 1 results produced a correlation coefficient, 'r', of 0.611, which is 61% by percentage. This value shows a high and positive relationship between content monitoring and cyberbullying reduction among students in public secondary schools in Rivers State. This implies that any increase in content monitoring would lead to a corresponding increase in cyberbullying reduction among students in public secondary schools in Rivers State.

Research Question 2: What is the relationship between parental control and cyberbullying reduction among students in public secondary schools in Rivers State?

Table 2: Pearson Product-Moment Correlation (PPMC) Showing the Relationship between Parental Control and Cyberbullying Reduction Among Students in Public Secondary Schools in Rivers State.

Variables	n	df	r	Decision
Parental Control Cyberbullying Reduction Among Students	552 552	550	0.690	High Relationship

Decision Rule: 0.00 – 0.19 = Very Low, 0.20 – 0.39 = Low, 0.40 – 0.59 = Moderate, 0.60 – 0.79 = High, 0.80 – 1.00 Very High

To answer research question 2, Table 2 results produced a correlation coefficient, 'r', of 0.690, which is 69% by percentage. This value shows a high positive relationship between parental control and cyberbullying reduction among students in public secondary schools in Rivers State. This implies that any increase in parental control would lead to a corresponding increase in cyberbullying reduction among students in public secondary schools in Rivers State.

Test of Hypotheses

Hypothesis 1: There is no significant relationship between content monitoring and cyberbullying reduction among students in public secondary schools in Rivers State.

Table 4: Test of Hypothesis on the Significant Relationship between Content Monitoring and Cyberbullying Reduction Among Studentsin Public Secondary Schools in Rivers State.

Variables	n	df	R	p-value	alpha level	Decision	
Content Monitoring Cyberbullying Reduction Among Students	552 552	550	0.611	0.001	0.05	Reject Ho_1 (Significant)	

For hypothesis 1 tested, Table 4 reveals a significant correlation at r = .611, with a p-value of 0.001 (P<0.05). Since the p-value 0.001 is less than the alpha level of 0.05, we therefore reject the null hypothesis; thus, there is a significant relationship between content monitoring and cyberbullying reduction among students in public secondary schools in Rivers State.

Hypothesis 2: There is no significant relationship between parental control and cyberbullying reduction among students in public secondary schools in Rivers State.

Table 5: Test of Hypothesis on the Significant Relationship between Parental Control and Cyberbullying Reduction Among Students in Public Secondary Schools in Rivers State.

Variables	n	df	r	p-value	alpha level	Decision
Parental Control Cyberbullying Reduction Among Students	552 552	550	0.690	0.000	0.05	Reject Ho_2 (Significant)

For hypothesis 2 tested, Table 5 reveals a significant correlation at r = .690, with a p-value of 0.000 (P<0.05). Since the p-value 0.000 is less than the alpha level of 0.05, we therefore reject the null hypothesis; thus, there is a significant relationship between parental control and cyberbullying reduction among students in public secondary schools in Rivers State.

DISCUSSION OF FINDINGS

The first finding of the study showed a high positive relationship between content monitoring and cyberbullying reduction among students in public secondary schools in Rivers State. Also, the hypothesis tested established a significant relationship between content monitoring

and cyberbullying reduction among students in public secondary schools in Rivers State. These findings are in tandem with Milosevic (2023), Nee (2023), García-Méndez et al. (2025), and Stone (2025), whose studies found that a high positive link exists between content monitoring and cyberbullying reduction among students in schools. A possible explanation of these findingssuggests that effective AI-assisted content monitoring systems in schools may successfully identify and mitigate harmful online behaviours, leading to a noticeable reduction in cyberbullying among students.This implies that school policymakers and administrators in Rivers State should invest more in content monitoring technologies and integrate them into school regulations, as such systems evidently play a significant role in promoting students' safety and well-being.

The second finding of the study revealed a high positive relationship between parental control and cyberbullying reduction among students in public secondary schools in Rivers State. Also, the hypothesis tested established a significant relationship between parental control and cyberbullying reduction among students in public secondary schools in Rivers State. These findings agree with Cassidy (2018), Helfrich et al. (2020), and Gnanasekaran (2025), whose empirical studies showed thata high positive correlationexists between parental control and cyberbullying reduction among students in public secondary schools. Explanation to these findings indicates that when parents actively monitor and guide their childrens online activities, students are less likely to engage in or become victims of cyberbullying.This implies that schools should collaborate more closely with parents by incorporating effective parental control measures into policy frameworks to help further reduce cyberbullying among students.

CONCLUSION

Based on the findings, the study has shown that AI-assisted social media usage policies, particularly content monitoring and parental control, are strongly and significantly related to the reduction of cyberbullying among students in public secondary schools in Rivers State. Therefore, the implementation of these AI-driven policies is essential for schools seeking to safeguard learners and foster a safer digital learning environment.

Recommendation

Given the findings and conclusions of the study, the following are recommended:

1. School administrators and policymakers should make content monitoring a mandatory component of AI-assisted social media usage policies in all public secondary schools, as it has been proven highly effective in significantly reducing cyberbullying among students.
2. School management should strengthen collaboration with parents by integrating structured parental control measures into social media policies, since active parental involvement has been shown to significantly reduce cyberbullying among students.

REFERENCES

1. Achuthan, K. (2022). Theoretical perspectives of parental influence on adolescent cyber behaviour: A TPB approach across India and Singapore.*International Journal of Adolescence and Youth*. PMC9703604.
2. Akeusola, B. N. (2023). Preventing Cyberbullying in Nigeria: The Effectiveness of Social Media Literacy Education for Young People.*Journal of Current Social and Political Issues*,7, 17-34
3. Akter, M., Park, J. K., & Wisniewski, P. J. (2025). Moving beyond parental control toward community-based adolescent online safety. *arXiv* Press
4. Ali, S., Elgharabawy, M., Duchaussoy, Q., Mannan, M., & Youssef, A. (2020). Betrayed by the guardian: Security and privacy risks of parental control solutions. *arXiv* Press
5. Cassidy, W. (2018). What parents can do to prevent cyberbullying: Students' and educators' perspectives. *Social Sciences*, *7*(12), 251.
6. García-Méndez, S., & De-Arriba-Pérez, F. (2025, April 7). *Promoting security and trust on social networks: Explainable cyberbullying detection using large language models in a stream-based machine learning framework*. Active Press
7. Gnanasekaran, V. (2025). *Next-generation parental controls: Facilitating digital competence and cybersecurity awareness*. Computers & Security Publishers.
8. Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5–14.
9. Helfrich, E. L., Doty, J. L., Su, Y. Y., Jacqlyn, L., & Gabrielli, J. (2020). Parental views on preventing and minimizing negative online experiences. *Children and Youth Services Review*, 109, 104668.
10. Mani, S. (2024). Why schools need to wake up to the threat of AI-deepfakes and bullying. *International Journal of Educational Security,* 7(2), 78-91.

11. Milosevic, T. (2022). *Artificial intelligence to address cyberbullying and harassment*. Jets Publishers
12. Milosevic, T. (2023). *Effectiveness of artificial intelligence–based cyberbullying interventions: Children's perceptions*.SAGE Publications
13. Milovidov, E., & Oxford research (2025). *"*Kids can bypass anything…"*The Guardian*, March 16.
14. Nee, C. N. (2023). *The digital defence against cyberbullying: A systematic review of AI-based detection*. Cogent Social Sciences Publishers.
15. Prama, T. T., Amrin, J. F., Anwar, M. M., & Sarker, I. H. (2025, March 4). AI-enabled user-specific cyberbullying severity detection with explainability. *Journal of Policy Analysis and Management*, 9(4) 89-102
16. Rees, D. I., Sabia, J. J., &Kumpas, G. (2022). Anti-bullying laws and suicidal behaviors among teenagers. *Journal of Policy Analysis and Management*, 7, 98-112
17. Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson
18. Stone, A. (2025, February 10). *Cyberbullying and the technology to stop, prevent and combat it in schools*. EdTech Magazine and Papers.
19. Suler, J. (2004). The online disinhibition effects. *Cyber Psychology & Behavior*, 7(6), 73-91.
20. TIME. (2019, July 8). *Inside Instagram's war on bullying.*TimePress.
21. U.S. Department of Commerce (2024). Encouraging industry to provide age-appropriate parental control tools. *NTIA Online Safety Taskforce Guidance*, 9,7.
22. Wordu, H., Dan-Jumbo, I., & Mina, A. D. (2021). Effects of social media addiction on academic performance of students in secondary schools in Rivers State, Nigeria. *International Journal of Advanced Education and Research,* 5(3), 63-81