AN INTERVENTION ON DIGITAL SAFETY: PROTECTING WOMEN FROM ONLINE HARASSMENT AND ABUSE IN THE ERA OF ARTIFICIAL INTELLIGENCE

Dr. Mariagoretti Ijeoma Obiakor & Prof. Nkechi P. Ikediugwu

Department of Educational Management snd Policy, Faculty of Education,
Nnamdi Azikiwe University Awka, Anambra State

ABSTRACT

With the rise of technology and the increasing use of artificial intelligence, the issue of online harassment and abuse against women has become a major concern. This paper addresses the need for digital safety measures to protect women from such forms of violence in the era of AI. Through the analysis of current research and data, we explore the various ways in which AI is utilized in perpetrating online harassment and abuse against women and the potential consequences it can have on their well-being. Additionally, we discuss the challenges and limitations in addressing this issue and propose solutions to promote a safer online environment for women. These solutions include stricter regulations for AI algorithms, enhanced user privacy settings, and increased awareness and education on digital safety practices. Our hope is that this paper will provide insight and guidance for policymakers, tech companies, and individuals to take action towards protecting women from online harassment and abuse in the ever-evolving world of artificial intelligence. Finally, Protecting women from online harassment and abuse requires a comprehensive approach that leverages technology, advocacy, and policy reform to address the unique challenges posed by digital platforms and AI advancements.

Keywords: Digital, Safety, Women, Harassment and Artifical Intelligence

Introduction

The rapid advancement of technology, particularly in the field of artificial intelligence (AI), has transformed the way individuals communicate, share information, and engage with online platforms. While these technological innovations offer substantial benefits, they also present significant challenges, particularly in issues related to digital safety for vulnerable populations. Among these populations, women face heightened risks of online harassment and abuse, which can manifest in various forms including cyberbullying, stalking, doxxing, and sexual harassment (Kelley et al., 2020). Research indicates that women are disproportionately targeted in online spaces, where they often encounter a range of hostile behaviors that can have severe emotional, psychological, and sometimes physical repercussions (Duggan, 2017).

The term "digital" refers to the representation of data in a binary format which can be processed by electronic devices. This transformation has fundamentally altered the way information is created, stored, and shared. In recent decades, the digital landscape has expanded significantly, influencing various aspects of daily life, including communication, education, and entertainment.

One of the primary characteristics of digital technology is its ability to facilitate instantaneous communication across vast distances. According to Castells (2012), the rise of digital communication technologies has created a "network society" where relationships and interactions are increasingly mediated by digital platforms. This shift has had profound implications for social dynamics and identity formation.

Furthermore, the digital revolution has transformed traditional industries. As noted by Brynjolfsson and McAfee (2014), businesses have adapted to digital technologies to improve efficiency and innovate their services. The proliferation of data analytics and digital marketing strategies has allowed companies to engage with consumers in unprecedented ways.

In education, the digital age has introduced new learning modalities. Online learning platforms and digital resources provide access to a wealth of information and opportunities for collaborative learning (Siemens, 2014). This democratization of knowledge has the potential to bridge educational gaps, although it also raises concerns about digital equity and access.

Therefore, the concept of digital encapsulates a wide range of technological advancements that have reshaped society. From communication to commerce and education, digital technologies continue to influence every facet of human life, fostering both opportunities and challenges especially online harassment which this paper will explore, surrounding the intersection of AI, gender, and online safety will be critical in informing effective strategies to combat online harassment and facilitate a more equitable digital environment.

Harassment

Harassment is a complex phenomenon that encompasses a range of unwanted behaviors aimed at individuals or groups, often leading to distress and adverse effects on their well-being. It can occur in various contexts, such as the workplace, educational institutions, or public spaces, and can manifest through verbal, physical, or psychological means.

One of the most recognized forms of harassment is sexual harassment, which is defined as unwelcome sexual advances or conduct that creates a hostile or intimidating environment (Mariagoretti, 2020). This type of harassment not only affects the victims' mental and emotional health but can also impact their professional lives and career progression.

In the workplace, harassment can take various forms, including bullying, discriminatory remarks, and pervasive hostile behaviors (Rayner et al., 2002). Research indicates that individuals who experience workplace harassment may face increased levels of stress, anxiety, and depression, ultimately leading to decreased job satisfaction and productivity (Hershcovis, 2011).

Moreover, the implications of harassment extend beyond the immediate harm to individuals. Organizations can suffer significant consequences, including legal liabilities, reputational damage, and a decline in employee morale (Robinson & O'Reilly, 1998). These factors highlight the necessity for effective policies and training aimed at preventing and addressing harassment in all environments.

Therefore, harassment is a detrimental behavior that can have profound psychological, social, and organizational implications. It is crucial for individuals and organizations to recognize and

address harassment to foster a healthy, respectful environment.

Artificial Intelligence

Artificial intelligence (AI) refers to the simulation of human intelligence processes by machines, particularly computer systems. Key capabilities of AI systems include learning, reasoning, problem-solving, perception, and language understanding (Russell & Norvig, 2020). Over the years, AI has evolved into a broad field with various applications that enhance productivity and efficiency across different domains, such as healthcare, finance, education, and transportation.

One of the fundamental components of AI is machine learning, which enables systems to improve their performance through experience and data input. Machine learning algorithms analyze patterns within large datasets and adapt based on their findings. According to Jordan and Mitchell (2015), advancements in machine learning have significantly contributed to the development of AI, transforming industries and leading to innovations such as autonomous vehicles and personalized medicine.

Another critical aspect of AI is natural language processing (NLP), which enables machines to understand, interpret, and respond to human language in a valuable manner. This capability is evident in applications such as virtual assistants, chatbots, and translation services (Jurafsky & Martin, 2020). NLP enhances human-computer interaction and allows for a more intuitive experience when engaging with technology.

Despite its many benefits, the rise of AI also poses ethical and societal concerns. Issues surrounding privacy, job displacement, and the potential for bias in AI algorithms require careful consideration and regulation (Binns, 2018). As AI systems become increasingly integrated into daily life, it is vital for stakeholders to address these challenges to ensure the responsible development and deployment of AI technologies.

In otherwords, artificial intelligence represents a significant technological advancement that simulates human cognitive functions. While AI offers numerous opportunities for improvement across various sectors, it is essential to navigate the ethical implications and societal impacts that accompany its growth.

The advent of AI technologies has further complicated the landscape of online harassment. AI algorithms are increasingly utilized to moderate content, analyze user behavior, and tailor experiences on social media platforms. However, these technologies can be flawed and may inadvertently contribute to the perpetuation of gender-based violence and harassment. Studies have found that AI systems may replicate and exacerbate existing biases, leading to inadequate responses to harassment incidents and a failure to protect victims (Lamb et al., 2021). Furthermore, the anonymity afforded by online platforms often emboldens perpetrators, creating a hostile environment where women may feel unsafe to express themselves freely (Binns et al., 2018).

In light of these challenges, there is an urgent need for comprehensive digital safety measures aimed specifically at protecting women from online harassment and abuse. This necessitates a all-around approach that includes the development of robust technological solutions, advocacy for policy changes, and the promotion of digital literacy and support systems for victims. Stakeholders, including policymakers, technology developers, and civil society organizations, must work collaboratively to create safer online spaces where women can

engage without fear of harassment (Binns, 2018; Chokri et al., 2022).

Problem Statement

Despite the significant advancements in technology and the growing prevalence of digital communication platforms, women continue to face alarming rates of online harassment and abuse. This issue is exacerbated in the context of artificial intelligence (AI), where algorithms and automated systems often fail to adequately protect vulnerable users, perpetuate biases, and inadvertently facilitate harmful behaviors. Current measures for combating online harassment are frequently insufficient, leading to a climate of fear and restriction of expression for women in digital spaces. Moreover, the lack of comprehensive strategies that incorporate both technological and human elements hinders effective interventions. Therefore, there is an urgent need to critically examine the intersection of AI technologies and online harassment, focusing on identifying effective protective measures and promoting a safer online environment for women. This study seeks to address the gaps in current knowledge and practices by exploring how AI can be leveraged to enhance digital safety for women while simultaneously illuminating the systemic issues that contribute to their victimization in online spaces. These problems are randomly listed below:

Prevalence of Online Harassment: Studies by Kelley, O'Connor, & Mowery, (2020)show that women are disproportionately affected by online harassment, with threats of violence, sexual harassment, and doxxing being common. This pervasive issue contributes to a hostile online environment that silences voices and stifles participation. Women are disproportionately affected by various forms of online harassment, including cyberstalking, doxing, and hate speech. The anonymity provided by the internet often exacerbates these issues, making it easier for perpetrators to attack without consequences. These issue needs to be resolved.

Lack of Awareness and Education: Many individuals, like Chokri et al (2022) Posited especially young women, are unaware of the risks associated with online platforms and the potential consequences of sharing personal information. There is a pressing need for education about digital safety, privacy, and reporting mechanisms.

Insufficient Support Systems: Victims of online harassment often lack access to adequate support systems and resources. Many are unsure of how to report incidents or seek help, leading to feelings of isolation and helplessness.

Cultural Stigma: Stigma around discussing online harassment may prevent victims from coming forward, further perpetuating a culture of silence and enabling the continuation of abusive behavior.

Inadequate Policy Frameworks: Existing legal and institutional frameworks often fail to address online harassment effectively, leaving victims without adequate protection or recourse.

Here are the some recorded incidents with date

There have been several notable incidents of online harassment and abuse against women in Nigeria that illustrate the growing challenges in the digital landscape. Here are a few examples:

Blessing Okoro Incident (2019): Blessing Okoro, a social media influencer, faced severe backlash after posting a photo in front of a house she falsely claimed was hers. The backlash quickly escalated into personal attacks and threats across various platforms. The incident

highlighted how quickly online narratives can spiral into harassment.

EndSARS Protests (2020): During the EndSARS protests against police brutality, many women activists faced online harassment, including doxxing and threats. Social media platforms were used to intimidate and silence women who spoke out, with personal information being shared and threats of violence issued.

Uku Ndukwu Case (2021): Uku, a Nigerian journalist and advocate, faced a coordinated attack on social media after reporting on issues related to gender-based violence. She was subjected to aggressive trolling, which included derogatory comments and threats, showcasing the risks female journalists face when tackling sensitive subjects.

Sadiya E. Mohammed: A young Nigerian woman who gained attention for her advocacy against early marriage and gender inequality faced significant online harassment, including threats and derogatory remarks, after sharing her views on social media.

Tiwa Savage Leak Incident (2021): Tiwa Savage, a prominent Nigerian musician, became the target of harassment after a private video of her was leaked. The situation led to significant online abuse, with many people blaming her for the leak and making various derogatory comments. This incident stirred discussions about privacy, consent, and the treatment of women in the public eye.

Sexual Harassment at Universities: Female students in Nigerian universities have reported online harassment related to issues of sexual consent and personal safety, where their personal information and images are shared without consent to shame or threaten them. Sexual cyber harassment of female students in

Nigerian universities is a serious issue. A 2018 World Bank survey found that 70% of female graduates reported experiencing sexual harassment during their studies. This harassment can include unwanted advances, inappropriate messages, and coercion for sexual favors, negatively affecting victims' mental health and academic performance.

The rise of AI and digital technologies has worsened the situation, leading to increased incidents of cyber harassment, such as online stalking and cyberbullying. Victims often face stigma and blame, which discourages them from reporting these incidents.

To combat this issue, various civil society organizations are advocating for the Anti-Sexual Harassment Bill, aimed at protecting students and promoting ethical standards in universities. It is essential for educational institutions to establish strong policies and support systems to prevent and address sexual harassment, both online and offline, to ensure a safe and respectful learning environment for all students.

Celebrity Backlash: Female celebrities like Genevieve Nnaji and Funke Akindele often face harassment and threats for their political views or personal choices. Social media users frequently attack them through derogatory comments and personal threats, reflecting a broader trend of silencing women.

Influencer Targeting: Several social media influencers, particularly those who advocate for women's rights or engage in feminist dialogue, have reported receiving threats and profound harassment, often fueled by trolls who deliberately undermine their campaigns.

These incidents are part of a broader pattern of gender-based harassment that's intensified by the anonymity and reach of social media,

especially with the rising influence of AI in content moderation and dissemination.

Methods of Intervention

Aims and Objectives

The "Digital Safety: Protecting Women from Online Harassment and Abuse" initiative aims to achieve the following objectives

Raise Awareness: To educate women and marginalized communities about online harassment, its forms, and the importance of digital safety.

Empower Victims: To provide resources and tools that empower victims of online harassment to report incidents, seek help, and advocate for them.

Develop Educational Programs: To create and implement workshops and training sessions focused on digital literacy, safety practices, and emotional resilience.

Establish Support Networks: To facilitate peer support groups and community networks where victims can share experiences and receive emotional support.

Advocate for Policy Change: To engage with policymakers to develop and implement stronger protections against online harassment and abuse.

Enhance Reporting Mechanisms: To improve awareness and accessibility of reporting mechanisms within online platforms and law enforcement agencies.

Proposed Activities

Workshops and Training Sessions: Conduct interactive workshops aimed at teaching women about digital safety, privacy settings, and how to identify and respond to online harassment.

Development of Educational Materials: Create informational resources, including pamphlets, videos, and online courses, that address digital safety and harassment prevention.

Online Safety Campaigns: Launch awareness campaigns on social media platforms to educate the public about the prevalence of online harassment and promote the importance of digital safety.

Support Group Formation: Establish peer-led support groups where individuals affected by online harassment can share experiences, provide mutual support, and learn coping strategies.

Policy Advocacy: Collaborate with advocacy organizations to engage policymakers in discussions about improving legal protections against online harassment.

Monitoring and Evaluation:

Implement a monitoring framework to evaluate the effectiveness of the initiative, collect feedback, and adjust programming as needed.

Partnerships and Collaboration

For effective implementation of this initiative, we will seek partnerships with the following entities:

Local NGOs and Community Organizations: Collaborate with organizations that focus on women's rights, digital safety, and mental health to leverage existing resources and networks.

Educational Institutions: FAWE should Partner with schools and universities to integrate digital safety education into existing curricula and reach a wider audience.

Technology Companies: FAWE should work with tech companies to promote safer online

environments and improve reporting mechanisms for harassment.

Law Enforcement Agencies: FAWE should collaborate with local law enforcement to ensure proper training on handling cases of online harassment and abuse.

Media Outlets:FAWE should engage media partners to amplify awareness campaigns and educate the public about the issue of online harassment.

Expected Outcomes

The "Digital Safety: Protecting Women from Online Harassment and Abuse" initiative is designed to produce several key outcomes that will contribute to a safer and more equitable online environment for women and marginalized communities. These outcomes include:

Increased Awareness and Understanding: Women and marginalized individuals will have a heightened awareness of the various forms of online harassment, how to recognize them, and the impact they can have on victims. This knowledge will empower them to take proactive measures to protect themselves and others.

Empowerment of Victims: Participants will gain the confidence and skills necessary to report incidents of online harassment. By providing access to resources and support networks, individuals will feel more empowered to seek help and take action against their abusers.

Enhanced Digital Literacy: The initiative will improve the digital literacy of participants, equipping them with essential skills to navigate online platforms safely. This includes understanding privacy settings, recognizing harmful content, and knowing how to use reporting mechanisms effectively.

Formation of Support Networks: Establishing peer-led support groups will create a safe space for individuals affected by online harassment to share their experiences, receive emotional support, and learn coping strategies. These networks will foster a sense of community and solidarity.

Increased Community Engagement: Awareness campaigns and community events will encourage greater dialogue around online harassment, leading to increased community involvement in addressing this issue. This collective engagement will help normalize discussions about digital safety and harassment prevention.

Stronger Advocacy for Policy Changes: The initiative will facilitate ongoing discussions with policymakers and advocacy organizations, resulting in the development and implementation of stronger legal protections against online harassment. This could lead to more robust laws and policies that prioritize the safety and rights of women online.

Improved Reporting Mechanisms: By raising awareness about existing reporting mechanisms and advocating for improvements, victims will have better access to support and resources when they experience online harassment. This could lead to higher reporting rates and more responsive actions from platforms and law enforcement.

Reduction in Incidents of Online Harassment: Over time, the initiative aims to contribute to a measurable reduction in reported incidents of online harassment within the community. By fostering a culture of respect and accountability, we hope to create a safer online environment for all individuals.

Evaluation and Continuous Improvement: The implementation of a monitoring and

evaluation framework will provide valuable insights into the initiative's effectiveness. Regular assessments will help identify areas for improvement and ensure that the program remains responsive to the needs of participants.

Conclusion

The Digital Safety: Protecting Women from Online Harassment and Abuse" initiative addresses a critical and pressing issue in today's digital landscape. As online harassment continues to disproportionately affect women and marginalized groups, it is essential to implement comprehensive strategies that not only raise awareness but also empower individuals to take action. Through targeted educational programs, supportive resources, and community engagement, this initiative aims to foster a safer online environment where women can freely express themselves without the threat of harassment or abuse.

By focusing on awareness, empowerment, and advocacy, we can break the cycle of silence surrounding online violence and build a culture of respect and dignity. Collaboration with local organizations, educational institutions, technology companies, and policymakers is vital for the success of this initiative. Together, we can develop and implement effective solutions that provide immediate support and long-term protection for victims of online harassment.

As we embark on this journey, we recognize the importance of continuous monitoring and evaluation to assess the impact of our efforts and adapt to the evolving challenges posed by online harassment. By investing in digital safety initiatives, we are not only protecting the rights of women but also creating a more equitable digital space for all.

We urge stakeholders, community leaders, and potential funders to join us in this essential work. Together, we can empower women, promote digital safety, and ensure that everyone has the right to participate in the online world without fear of violence or harassment. Your support will be instrumental in making this vision a reality.

Summary

The discussion on Digital Safety: Protecting Women from Online Harassment and Abuse in the Era of AI focuses on the increasing prevalence of online harassment and abuse targeting women, especially in the context of rapidly evolving digital technologies and AI. AI can both mitigate and exacerbate online harassment. On one hand, AI tools can be employed to detect and prevent abusive content. Additional, algorithms may inadvertently perpetuate bias and discrimination, leading to further victimization of marginalized groups. Strategies for protecting women online include implementing stronger policies on social media platforms, enhancing user privacy, and fostering digital literacy among women. Encouraging reporting mechanisms and providing support resources are also critical to empower victims.There is a need for comprehensive legal frameworks that specifically address online harassment and abuse, incorporating definitions of digital violence and ensuring accountability for offenders. Grassroots organizations and advocacy groups play a crucial role in raising awareness about digital safety issues and providing resources for women experiencing online abuse. Collaboration between authorities, tech companies, and civil society is essential to create safer online spaces.

Finally, protecting women from online harassment and abuse requires a multifaceted approach that leverages technology, advocacy, and policy reform to address the unique

challenges posed by digital platforms and AI

advancements.

References

American Psychological Association. (2020). Violence and harassment. https://www.apa.org/topics/violence/harassment

Binns, A. (2018). Fairness in machine learning: Lessons from political philosophy. Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency, 149-158.

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. In Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency* (pp. 149-158). https://doi.org/10.1145/3287560.3287598

Binns, R., Malesky, N., & Kremling, J. (2018). Gender-based online harassment: A systematic review of the literature. Journal of Communication, 68*(2), 289-310.

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Castells, M. (2012). *Networks of outrage and hope: Social movements in the internet age*. Polity Press.

Chokri, A., Naoui, F., & Boughanmi, A. (2022). Tackling online abuse against women through artificial intelligence: Current approaches and future directions. New Media & Society, 24 (2), 295-312.

Duggan, M. (2017). Online harassment 2017. Pew Research Center. Retrieved from https://www.pewresearch.org/fact-tank/2017/07/11/online-harassment-2017/

Hershcovis, M. S. (2011). Incivility, social undermining, bullying, and harassment in the workplace: A clarification of constructs and recommendations for measurement. Journal of Occupational Health Psychology, 16 (2), 121-134. https://doi.org/10.1037/a0021687

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260. https://doi.org/10.1126/science.aaaa8412

Jurafsky, D., & Martin, J. H. (2020). Speech and language processing (3rd ed.). Pearson.

Kelley, K., O'Connor, J., & Mowery, D. (2020). The impact of online harassment on women in public life. Journal of Computer-Mediated Communication, 25(3), 448-465.

Lamb, R., & Stoecker, C. (2021). The role of AI in scaling harassment: A systematic review. Computers in Human Behavior, 121, 106791.

Rayner, C., Hoel, H., & Cooper, C. L. (2002). Workplace bullying: What we know, who is to blame, and what can we do?. The Leadership & Organization Development Journal, 23(4), 202-208. https://doi.org/10.1108/01437730210437502

Robinson, S. P., & O'Reilly, J. (1998). Violent behavior in the workplace: Investigation of the causes of an organization. Journal of Business and Psychology, 12(2), 291-305. https://doi.org/10.1023/A:1020518812915

Russell, S., & Norvig, P. (2020). Artificial intelligence: A modern approach (4th ed.). Pearson

Siemens, G. (2014). Learning in the digital age: The role of technology in education. _Educational Technology Magazine, 54_(2), 12-16.