iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** X **Month of publication:** October 2026

DOI: https://doi.org/10.22214/ijraset.2026.85301

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue X Oct 2026- Available at www.ijraset.com

Hate Speech and Cyber Ethics: A Critical Analysis Based on Cancel Culture and Echo Chambers

Wong TsunTing[1], Mohd Haizra[2]
Faculty of Business and Communication, INTI International University, Nilai, Negeri Sembilan, Malaysia

Abstract: ***In the modern digital media world, one of the most serious problems surrounding the ethical issue is that of hate speech. Most previous research and discussions surrounding the topic have focused on the individual phenomenon of cyberbullying, while entirely overlooking the concept of mass hate speech which often masquerades as "social justice." This paper provides an analysis of contemporary online situations of "cancel culture" and examines how platform algorithms and the echo chamber effect create an environment that creates more extreme forms of rhetoric. Using Albert Bandura's Moral Disengagement Theory, Cass Sunstein's Echo Chamber Effect, and John Suler's Online Disinhibition Effect as the basis of this analysis, the paper suggests that people today justify hate speech as being a form of moral justice because of the influence of algorithms in amplifying hate speech messages. It is hoped that all stakeholders in the digital realm will accept responsibility for this ethical challenge through the use of digital ethics in shaping a more positive public discourse.***
Keywords: Hate Speech; Cyber Ethics; Cancel Culture; Echo Chamber; Moral Disengagement; Algorithmic Gatekeeping.

I. INTRODUCTION

The use of digital social media has transformed how people engage with politics and society. However, the freedom of expression allowed in cyberspace has also led to a rise in hate speech. The world has witnessed the emergence of the term "cancel culture"; the use of online social media to "cancel" people who offer opposing views by means of public boycotts, name-calling, and/or doxxing. Scholars have suggested that cancel culture is a form of shaming through social media that can be both an example of accountability and cyber-bullying. People who engage in such behavior often believe they are serving a useful purpose, but their actions cross all moral boundaries and represent nothing more than a form of hate speech. This article will examine how algorithms have contributed to this ethical dilemma through the ways in which algorithm-generated recommendations influence collective action (e.g., mobilization) among social media users.Moreover, the rapid changes in architecture of modern platforms have altered the public sphere in ways that are never before seen in the history of media. Unlike the case with traditional broadcasting media where speakers were filtered through professional journalists, the digital platforms create a world of hyper-connectivity, fragmentation, and even alienation. In these settings, the way the user interacts with the content alters in the form of likes, retweets, or comments. Thus, it becomes mandatory for the users to post content that is provocative, emotionally charged, and moralized in order to get maximum personalized exposure due to the nature of the structure. As a result, one should note that "cancel culture" may not simply operate in the spontaneous formation of a community but rather functions as a digitally mediated process. Thus, the advancements in digital transformation of conflict into capital, makes a considerable impact on the ability of individuals to operate in the future.

II. LITERATURE REVIEW

The majority of the existing literature about online hate speech focuses on its psychological impact on its victims but not much research is conducted to ascertain how hate speech is legitimized "socially". Some authors argue that social media promotes sensational and controversial content since it works according to the attention economy [5] however, there aren't any studies that consider "cancel culture" in relation to platform algorithms through which to explore the problem from the media ethics perspective. This literature review aims to fill this gap by analyzing the impact of structural factors and mediation algorithms on the legitimization and normalization of hate speech.

III. THEORETICAL FRAMEWORK

All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

A. Theory of Moral Disengagement

Moral Disengagement Theory is an idea developed by Albert Bandura that explains how different people can do harmful things to each other on the internet without feeling guilty about it [1].

©IJRASET: All Rights are Reserved

In "cancel culture," social media users can use "moral justification" to justify the use of some hate speech while claiming to be "doing the right thing," therefore removing their own moral boundaries.

B. Echo Chamber Effect

According to the Echo Chamber Effect, algorithms that provide recommendations for online users give them data that supports their previously held beliefs [2]. Information segregation leads to group polarization, breeding hostility towards opposing opinions by users living in echo chambers and lowering the threshold for hate speech provision.

C. The Online Disinhibition Effects

In addition to the above theories, John Suler's definition of the "Toxic Online Disinhibition Effect" indicates that anonymity, invisibility, and asynchronous nature of communication in cyberspace serve to undermine social norms [7]. This psychological process encourages individuals to exhibit more aggressive behavioral patterns than they would do otherwise in personal interactions.

D. Algorithmic Amplification and Attention Economy

In order to understand how modern cyber hate functions, it is important to apply the attention economy concept. Communication platforms use a commercial model based on advertising revenue and user engagement. Algorithms are designed to detect, prioritize, and promote material that causes strong emotions (especially rage, fear, and righteous anger). For example, when individual moral disengagement combines with engagement-oriented recommendation engines, a positive feedback loop is created. Vicious and hostile speeches bring higher rates of clicks which leads to the amplification of these statements due to algorithms. As a result, algorithmic amplification is a means that helps transform local quarrels into harassment attacks on the platform level, making the digital attention a commodity.

IV. ETHICAL THREATS TO CANCEL CULTURE

A. The Illusion of Justice and Justified Evil

In the present online age, hate speech very often appears as ethical criticism. When somebody, public figure or ordinary citizen, makes an inappropriate statement, groups of people use insults and threats that often seen as "just punishment". There is a paradox that begs the question about how the authority of courts is used while other people can devalue the dignity of others for the sake of fulfilling their demands.

B. Algorithmic Segregation and Extreme Polarization

To keep users on the site, social networks use algorithms that promote extreme content that leads to anger and irritation. Glen Gillespie stresses that digital media perform the function of "guardians of the internet" although their content management is often shadowed with the idea of not stopping systemic oppression [8]. It goes without saying that from an ethical standpoint, this "engagement-based" algorithm framework is the violation of ethical principles as it promotes and rewards hate speech.

C. The Problem of Decontextualization and the Absence of Proportionality

A further serious ethical issue associated with online cancel culture is the tendency of digital decontextualization and the complete abandonment of proportionality of punishment. In the viral world, subtle yet crucial human communication, unfortunate historical mistakes, and/or trivial infractions are taken out of context, boiled down to a few seconds of video or a screenshot, and weaponized. When an intended target is framed within a simplistic binary system of oppressor versus oppressed, the digital mob demands instantaneous and complete social and economic ostracism without allowing the accused any legal protections or chances of redemption. Through vigilantism, there are no formalized legal or ethical standards for mob activities, and as such, this can lead to horrible outcomes in the real world, including extreme mental trauma, loss of employment, and social death. Thus, the lack of proportionality in online mobbing exemplifies the ways in which hate speech mediated through the digital world runs contrary to several of the key concerns of restorative justice and human dignity.

V. DISCUSSION

A. The Issue of Freedom of Speech Versus Hate Speech

The freedom of speech is fundamental to digital communication but does not grant people the right to engage in insults and slander. Criticism should be performed by rational discussion of a particular ideology; however, nowadays the term “cyber hate speech” refers to complete denigration of the person.

According to Citron and Norton, balancing free speech presupposes the establishment of "digital citizenship" with the cooperation of intermediaries and users in the struggle against the digital abuse (9). Media platforms need to find the moral balance between the protection of freedom of speech and the dignity of a person.

B. Outrage as a Marketable Product

Business models of these platforms create a commodity from “outrage” and “hatred” (5). Radical hate speech leads to the increase of advertising money. In regard to media ethics (3), media platforms assume the position of gatekeepers and contradict essential principles of Social Responsibility Theory (4) through the acceptance of hate speech for commercial purposes.

C. Responsibility of Institutions and the Effects of Legal Parameters

There are certain legal constraints with regard to hatred on the internet, which need to be addressed in light of the existing legal boundaries between freedom of expression and unlawful harassment. The conventional laws offer a challenge to officials dealing with rapid and decentralized, as well as anonymous communication across the internet. Liability in cases of internet mobbing becomes so broadly distributed among the different instigators, anonymous commentators, and corporations. Viewing corporations only as neutral technical conduits relieves companies from bearing social responsibility and provides them with exemption as per the laws. However, at the same time, the enforcement of strict laws regarding the communication on the internet may create chilling effect and inhibit legitimate dissent and criticism. Consequently, it is imperative to develop an ethical framework, which will allow developing and implementing of a transparent procedure of content moderation to achieve the different treatment of harsh public criticism and malicious and targeted hate.

VI. RECOMMENDATIONS

1) Algorithmic Governance of Platforms: Platforms should embed ethical accountability into their recommendation algorithms. According to the data ethics framework, technology design has to inherently embrace human rights principles and social stability in [[10]]. Algorithms should not only facilitate AI-driven content moderation but also filter out high profile controversial words, eliminate the echo chamber effect, and ensure that users are exposed to moderated resources that provide balanced perspectives.

2) Digital Empathy for Users: Society needs to learn to practice "digital empathy." All social media users have to verify facts and prove intellectual integrity prior to joining the arena of "cancel culture" or public critiques.

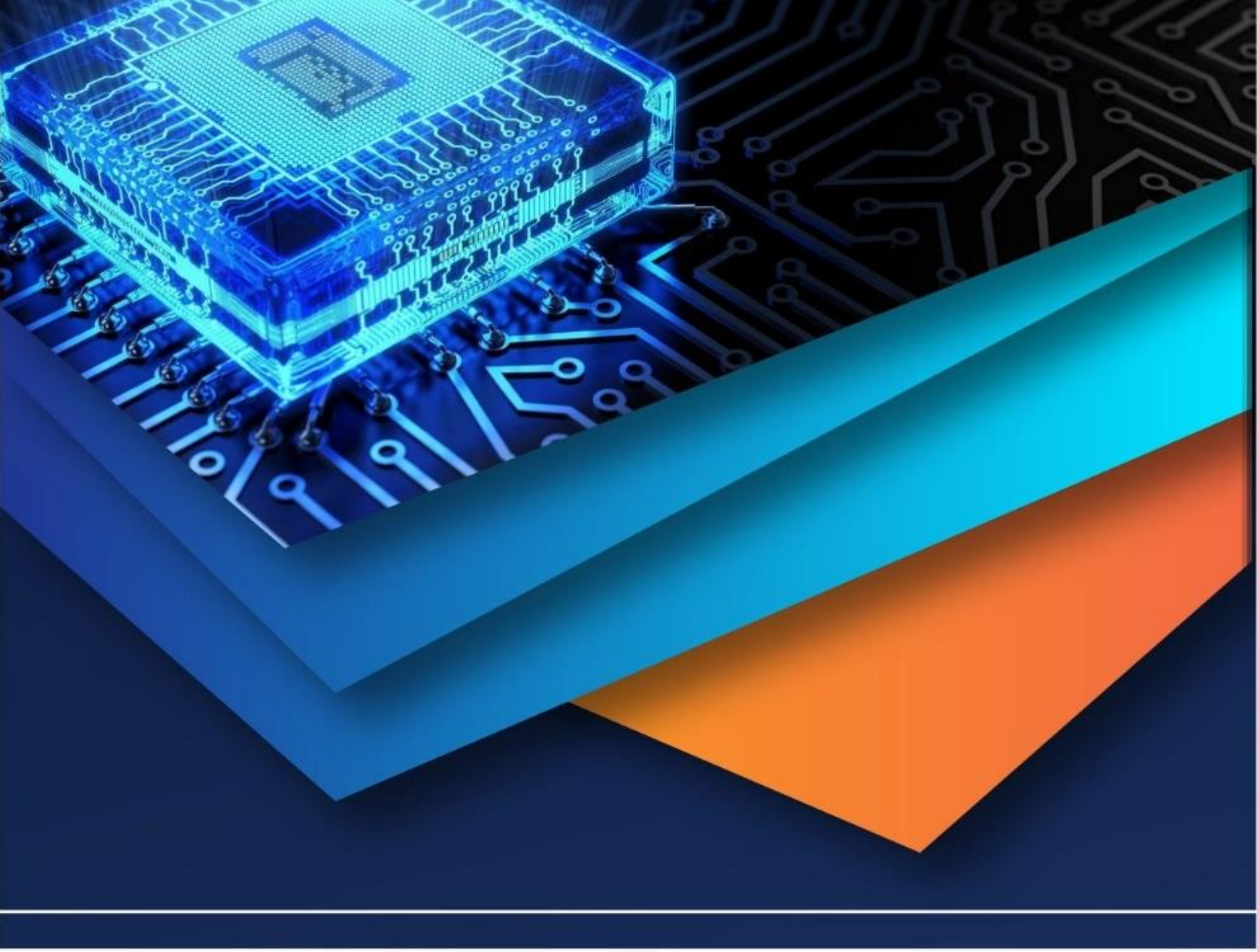
3) Legal Certainty for Regulators: The regulatory authorities have to define legal boundaries distinguishing "permissible negative criticism" from "prohibited hate speech." Through publishing regular reports on transparency, social media platforms will be held accountable for allowing the dissemination of hate speech.

VII. DISCUSSION

The issue of hate speech in the digital age has evolved from being simply an emotional outlet to something which has become a larger societal ethical dilemma as a result of various factors including the use of digital media algorithms, the disinhibition effect of the medium, and cancel culture; as the evidence from this study suggests reliance on self-regulation of users is not enough to solve the problem of hate speech online as it requires social media platforms to abandon an engagement-based logic that privileges profit-making and instead accept the role of ethical algorithmic gatekeepers.

REFERENCES

[1] A. Bandura, "Selective moral disengagement in the exercise of moral agency," Journal of Moral Education, vol. 31, no. 2, pp. 101-119, 2002.
[2] C. R. Sunstein, #Republic: Divided democracy in the age of social media. Princeton, NJ: Princeton University Press, 2018.
[3] S. J. A. Ward, Ethics and the media: An introduction. Cambridge, UK: Cambridge University Press, 2019.
[4] D. McQuail, McQuail's mass communication theory, 6th ed. London, UK: Sage Publications, 2010.
[5] S. Zuboff, The age of surveillance capitalism: The fight for a human future at the new frontier of power. New York, NY: PublicAffairs, 2019.
[6] M. D. Clark, "DRAG THEM: A brief etymology of so-called 'cancel culture'," Communication and the Public, vol. 5, no. 3-4, pp. 88-92, 2020.
[7] J. Suler, "The online disinhibition effect," CyberPsychology & Behavior, vol. 7, no. 3, pp. 321-326, 2004.
[8] T. Gillespie, Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media. New Haven, CT: Yale University Press, 2018.
[9] D. K. Citron and H. Norton, "Intermediaries and hate speech: Fostering digital citizenship for our information age," Boston University Law Review, vol. 91, pp. 1435-1484, 2011.
[10] L. Floridi and M. Taddeo, "What is data ethics?" Philosophical Transactions of the Royal Society A, vol. 374, no. 2083, p. 20160360, 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)