

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS



Open Access, Refereed Journal Multi Disciplinary
Peer Reviewed Edition :

www.ijlra.com

DISCLAIMER

No part of this publication may be reproduced or copied in any form by any means without prior written permission of Managing Editor of IJLRA. The views expressed in this publication are purely personal opinions of the authors and do not reflect the views of the Editorial Team of IJLRA.

Though every effort has been made to ensure that the information in Volume 2 Issue 7 is accurate and appropriately cited/referenced, neither the Editorial Board nor IJLRA shall be held liable or responsible in any manner whatsoever for any consequences for any action taken by anyone on the basis of information in the Journal.

Copyright © International Journal for Legal Research & Analysis

IJLRA

EDITORIAL TEAM

EDITORS



Megha Middha

Megha Middha, Assistant Professor of Law in Mody University of Science and Technology, Lakshmangarh, Sikar

Megha Middha, is working as an Assistant Professor of Law in Mody University of Science and Technology, Lakshmangarh, Sikar (Rajasthan). She has an experience in the teaching of almost 3 years. She has completed her graduation in BBA LL.B (H) from Amity University, Rajasthan (Gold Medalist) and did her post-graduation (LL.M in Business Laws) from NLSIU, Bengaluru. Currently, she is enrolled in a Ph.D. course in the Department of Law at Mohanlal Sukhadia University, Udaipur (Rajasthan). She wishes to excel in academics and research and contribute as much as she can to society. Through her interactions with the students, she tries to inculcate a sense of deep thinking power in her students and enlighten and guide them to the fact how they can bring a change to the society

Dr. Samrat Datta

Dr. Samrat Datta Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Samrat Datta is currently associated with Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Datta has completed his graduation i.e., B.A.LL.B. from Law College Dehradun, Hemvati Nandan Bahuguna Garhwal University, Srinagar, Uttarakhand. He is an alumnus of KIIT University, Bhubaneswar where he pursued his post-graduation (LL.M.) in Criminal Law and subsequently completed his Ph.D. in Police Law and Information Technology from the Pacific Academy of Higher Education and Research University, Udaipur in 2020. His area of interest and research is Criminal and Police Law. Dr. Datta has a teaching experience of 7 years in various law schools across North India and has held administrative positions like Academic Coordinator, Centre Superintendent for Examinations, Deputy Controller of Examinations, Member of the Proctorial Board



Dr. Namita Jain



Head & Associate Professor

School of Law, JECRC University, Jaipur Ph.D. (Commercial Law) LL.M., UGC -NET Post Graduation Diploma in Taxation law and Practice, Bachelor of Commerce.

Teaching Experience: 12 years, AWARDS AND RECOGNITION of Dr. Namita Jain are - ICF Global Excellence Award 2020 in the category of educationalist by I Can Foundation, India. India Women Empowerment Award in the category of "Emerging Excellence in Academics by Prime Time & Utkrisht Bharat Foundation, New Delhi.(2020). Conferred in FL Book of Top 21 Record Holders in the category of education by Fashion Lifestyle Magazine, New Delhi. (2020). Certificate of Appreciation for organizing and managing the Professional Development Training Program on IPR in Collaboration with Trade Innovations Services, Jaipur on March 14th, 2019

Mrs.S.Kalpana

Assistant professor of Law

Mrs.S.Kalpana, presently Assistant professor of Law, VelTech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Avadi. Formerly Assistant professor of Law, Vels University in the year 2019 to 2020, Worked as Guest Faculty, Chennai Dr. Ambedkar Law College, Pudupakkam. Published one book. Published 8 Articles in various reputed Law Journals. Conducted 1 Moot court competition and participated in nearly 80 National and International seminars and webinars conducted on various subjects of Law. Did ML in Criminal Law and Criminal Justice Administration. 10 paper presentations in various National and International seminars. Attended more than 10 FDP programs. Ph.D. in Law pursuing.



Avinash Kumar



learning.

Avinash Kumar has completed his Ph.D. in International Investment Law from the Dept. of Law & Governance, Central University of South Bihar. His research work is on "International Investment Agreement and State's right to regulate Foreign Investment." He qualified UGC-NET and has been selected for the prestigious ICSSR Doctoral Fellowship. He is an alumnus of the Faculty of Law, University of Delhi. Formerly he has been elected as Students Union President of Law Centre-1, University of Delhi. Moreover, he completed his LL.M. from the University of Delhi (2014-16), dissertation on "Cross-border Merger & Acquisition"; LL.B. from the University of Delhi (2011-14), and B.A. (Hons.) from Maharaja Agrasen College, University of Delhi. He has also obtained P.G. Diploma in IPR from the Indian Society of International Law, New Delhi. He has qualified UGC - NET examination and has been awarded ICSSR - Doctoral Fellowship. He has published six-plus articles and presented 9 plus papers in national and international seminars/conferences. He participated in several workshops on research methodology and teaching and

ABOUT US

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS

ISSN

2582-6433 is an Online Journal is Monthly, Peer Review, Academic Journal, Published online, that seeks to provide an interactive platform for the publication of Short Articles, Long Articles, Book Review, Case Comments, Research Papers, Essay in the field of Law & Multidisciplinary issue. Our aim is to upgrade the level of interaction and discourse about contemporary issues of law. We are eager to become a highly cited academic publication, through quality contributions from students, academics, professionals from the industry, the bar and the bench. INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS ISSN 2582-6433 welcomes contributions from all legal branches, as long as the work is original, unpublished and is in consonance with the submission guidelines.

CRITICAL ANALYSIS: HATE SPEECH IN INSTAGRAM

AUTHORED BY - CAROL FRANCIS JOSEPH

Abstract-

The researcher has aimed to critically analyse on the issue of Hate Speech in the specific online platform of Instagram. The study tackles with three main questions: Validity of using Section 66A of the IT Act (Information Technology Act 2000) on Instagram platform abusers, the problem in defining what a hate speech exactly is and rules and regulations that govern the issue of hate speech on Instagram.

The research methodology adopted for this particular paper is the doctrinal method of research wherein various kinds of literature articles were extensively reviewed by the researcher in coming to a conclusion.

This research goes in depth to analyse hate speeches on Instagram with an aim to examine its significance, challenges and legal implications or other rules that govern it. It looks into the validity of Sec 66A of the Information Technology Act 2000, challenges in defining hate speech and whether there are comprehensive laws that regulate hate speech on Instagram.

By critically analysing hate speech on Instagram, this research further gives insights to policy makers, administrators and users in maintaining a healthy digital environment that protects an individual's rights and respects the fundamental principle of freedom of speech even in an online platform.

Keywords-

Instagram, Instagram community guidelines, Hate speech, Social-media, Violative content,

Introduction:

It is quite fascinating how mere technology has brought us together without any physical interaction, through a website or an app with the help of internet. Social-media demonstrates various such technologies that facilitate communication, networking, sharing of ideas and thoughts, expressing opinions, connecting etc. on a broad scale. It helps users to post photos, videos, messages and engage in liking, sharing and commenting on others' posts as well. Instagram is one such platform that has evolved over time by bringing up creative ideas for young users in today's generation.¹ It acts as a source for youngsters to occupy themselves in the world of 'internet'. The day-by-day consumption of content, specifically various content has watched a large audience shift from other social media platforms to Instagram. However, this evolved technology has also brought about an alarming rise in the issue of hate speeches on the Instagram platform.

United Nations have defined Hate speech as, "... any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are..."² Hate speeches not only happen physically but is found to occur on social media platforms as well. It targets people based on certain attributes like race, religion, caste, colour, etc. With Artificial Intelligence becoming the core of every technology, we can see the future generation relying on various social media platforms which could further increase the risk of hate speeches in the online mode of communication. Instagram being one of the strongest platforms which could shape and influence the public at large in becoming a change for tomorrow, has to be well regulated and up-to-date with the new trends of the society. Hence, it becomes important to inspect the effectiveness of present laws, rules and regulations that govern hate speeches in the online mode, especially on Instagram.

Historical evolution of social media:

Historical changes in the use of social media begins in the late 1990s with the emergence of blogs and forums like 'SixDegrees' where people shared and discussed on written content. This helped in community building and user-generated content. In 2002 and 2003 'MySpace' and 'Friendster'

¹ Maya Dollarhide, "Social Media: Definition, Importance, Top Websites & Apps", (10th October, 2023, 6:14 p.m.), <https://www.investopedia.com/terms/s/social-media.asp>

² Hate Speech: Turning the tide", (10th October, 2023,9:15p.m.), <https://www.un.org/en/observances/countering-hate-speech#:~:text=The%20Plan%20of%20Action%20defines,%2C%20race%2C%20color%2C%20descent%2C>

were the two initial platforms that acted as social networking sites to connect people in digital relationships. In the mid-2000s, we had 'Facebook' (launched in 2004) that redefined social networking and set a competition for other social media platforms. In 2006, 'Twitter' came into being with its creative feature of posting short updates that became a new platform for news and ongoing trends. By 2010 'Instagram' and 'Pinterest' came up with its idea of photo and video sharing. Along with this we also had 'YouTube' and 'TikTok' that allowed for video and photo sharing on the social media platform. By 2015 we entered the era of live streaming which allowed the broadcast of live videos for entertainment purposes. However, these platforms gradually developed into political tools that played a crucial role in various famous campaigns (e.g.- Black Lives Matter) and interrupted privacy and data security. In 2020 Facebook put forward its idea of bringing the physical and virtual world together through the concept of 'Metaverse' (now Meta Platforms)³

Recognition of hate speech:

Landmark case of *Elonis v United States*⁴ was one of the cases that brought hate speech in social media to the forefront. In this case Anthony Elonis was accused of posting threatening messages to his ex-wife and co-workers through Facebook. The question that the court addressed here was whether Elonis' messages could be considered as 'true threats' or if it could be considered his right to free speech under the fundamental principles. This case challenged the interpretation of potential harm through an online statement and questioned the limits of free speech. It also raised the question of differentiating between a hate speech and the expression of anger and frustration.

Hate speech on Instagram:

As mentioned before, Instagram is a social media platform that allows users to share photos, videos; like, share and comment on others' posts as well as connect with people to build relationships. It was launched in 2010 and gradually gained popularity amongst young people for its creative ways of interaction and social networking. With the increase in the number of people using this platform as well as the complexity involved in maintaining online connections gave rise to the issue of hate speech on Instagram. Like the other social media platforms, even

³ Alexandra Samur; Colleen Christison, "The History of Social Media in 33 Key Moments", (12th October, 2023, 6:35p.m.), <https://blog.hootsuite.com/history-social-media/>

⁴ "Facts and Case Summary- *Elonis v. United States*", (12th October, 2023, 8:40 p.m.), <https://www.uscourts.gov/educational-resources/educational-activities/facts-and-case-summary-elonis-v-us#:~:text=The%20court%20held%20that%20the,three%20years%20of%20supervised%20release>

Instagram faced a challenge to balance between the right to free speech and the need to provide with a safe and healthy online environment. Instagram has taken measures to address the issue of hate speech. Instagram's community guidelines specifically prohibits hate speeches made, content moderation targets violative content and erases them, comment filters enables users to hide or block offensive comments on user's posts, machine learning algorithms identify and restrict violative content, anti-bullying features like blocking, restricting and reporting accounts as well as partnerships and education wherein Instagram collaborates and comes up with awareness driven campaigns for the betterment and reduction of hate speeches.⁵

Guidelines by Instagram:

Last year, on 21st January 2022, Instagram (owned by Meta today) came up with new measures to address hate speeches on the platform.⁶ These measures include showing lower feeds and posts that may contain violative content. This only affects the post by penalizing only that specific post and not the whole account of the user. Instagram closely examines the captions, posts, feeds that display violative content and compare it with the previously reported offensive content. It also looks into the user's account history of being reported and having previous posts and feeds that have been flagged. Using algorithms and processes, Instagram lowers the violative content that have been repeatedly reported as problematic by the users. However, it is important to note that Instagram does not completely remove a content unless it is fact-checked by third parties and makes all of the content of an account difficult to be found by users if it is repeatedly posting offensive content. The community guidelines as provided by Instagram has specific section dedicated to addressing hate speeches and offensive content under the "Respect for other community members" paragraph of the Instagram community guidelines.⁷ It states clear intention of being against any hate speeches or offensive content that may encourage violence and attacks individuals on different attributes like race, religion, caste, etc. and the guidelines also specifies that if a hate speech is posted for the need to raise awareness, clear intention to make such a post should be mentioned. All these measures show how Instagram is aiming towards the safety of its online consumers. However, the effectiveness of these measures and guidelines in addressing hate speeches on Instagram depends on how it is enforced and how clearly it is communicated to

⁵ "An Update on our work to tackle abuse on Instagram", (14th October, 2023, 11:20 p.m.), <https://about.instagram.com/blog/announcements/an-update-on-our-work-to-tackle-abuse-on-instagram>

⁶ Tech Desk, "Instagram will show 'potential' hate speech lower down in your Feed, Stories", (15th October, 2023, 6:30 a.m.), <https://indianexpress.com/article/technology/tech-news-technology/instagram-to-show-potential-hate-speech-lower-down-in-your-feed-stories-7734969/>

⁷ "Community Guidelines", (16th October, 2023, 5:00 p.m.), <https://help.instagram.com/477434105621119>

the users.

Laws that govern Instagram internationally:

There is no specific law dedicated to govern Instagram, but like other social media platforms, it is dependent on various rules and laws related to intellectual property rights, privacy and data protection as well as other online activities.

Some important laws that apply to Instagram include: Data Protection and Privacy laws like General Data Protection Regulation (GDPR)⁸ in European Union, Personal Information Protection Law (PIPL) in China and California Consumer Privacy Act (CCPA) in United States⁹ for protection of data and privacy of users¹⁰, Copyright and Intellectual Property laws Digital Millennium Copyright Act (DMCA) in United States¹¹ for respecting copyrights and intellectual property rights of users who post a content, Cyberbullying and Harassment laws which varies across the States for stalking, harassment and bullying done online, Defamation laws for spreading false information intending to ruin one's reputation, Hate Speech and Discrimination laws for posting offensive content based on different attributes of an individual, Content Moderation laws specific in some countries for removing content that portrays hate speeches, violative comments or political speeches, Child Online Privacy and safety laws like Child Online Privacy Protection Act (COPPA) in the United States¹² for regulating online users who belong below a certain age (Instagram has its own policy for those under 13 years of age), E-Commerce and Advertising laws for users who indulge in online promotion of products and services to comply with consumer rights, labelling and other basic needs.

⁸ "How does the General Data Protection Regulation (GDPR) affect advertising on Instagram?" (18th October, 2023, 7:00 p.m.), <https://help.instagram.com/2000935033561463>

More

Computer Help

Android App Help

Basic Mobile Browser Help

iPad App Help

iPhone App Help

Mobile Browser Help

⁹Arielle Feger, "California Privacy regulations have implications beyond the state", (18th October, 2023, 7:30 p.m.), <https://www.insiderintelligence.com/content/california-privacy-regulations-have-implications-beyond-state>

¹⁰ Dr Preeti Goel, TOI, "Personal Data Privacy- Does India need regulations like Europe's GDPR and USA's CCPA", (18th October, 2023), <https://timesofindia.indiatimes.com/blogs/voices/personal-data-privacy-does-india-need-regulations-like-europes-gdpr-and-usas-ccpa/>

¹¹ "Copyright Guidelines", (18th October, 2023, 8:00 p.m.), <https://help.instagram.com/126382350847838>

¹² "Children's Online Privacy Protection Rule", (18th October, 2023, 10:30 p.m.), <https://www.ftc.gov/business-guidance/resources/complying-coppa-frequently-asked-questions>

Laws that govern Instagram in India:

In India, the Constitution itself provides with the Fundamental Right to freedom of speech and expression under Article 19(1)(a) and Article 19(2) imposes reasonable restrictions on this fundamental right. The constitution of India itself faces a challenge to balance between hate speech and the freedom of speech. Article 21 provides for the Right to Privacy which is also being infringed in huge numbers in the online platform. Article 21A talks about Right to education. Although not directly applied, education about the digital world, its significance and consequences may make a better world to live in for the future generations. Article 22 describes rights of people who are arrested which specifically comes into use when online harassments take place. The Indian Constitution came into being much before the invention of these digital platforms. Hence, it is important to consider that legal precedents play a significant role in determining these fundamental rights and duties imposed on the citizens.¹³

Along with this, laws like -Indian Penal Code (IPC)- Sections 153A (promotion of enmity between people based on different attributes), 295A (acts that disrespect others' religious feelings) and 505(promotion of enmity between classes) for penalizing hate speeches in the online platform, Information Technology (IT) Act 2000 along with the rules under it like Section 69A for restricting public access to certain information if it is deemed to be in the interest of the safety and security of the nation for friendly relations with foreign states and maintaining public order; previously, we had Section 66A of the IT Act that specifically dealt with hate speeches but this was struck down by the Indian court in the Shreya Singhal v Union of India case¹⁴; various rules provided by the IT Act to address defamation, hateful and other offensive content in the online platform, Protection of Children from Sexual Offences (POCSO) deals with special consideration to child abuse and hate speeches against children in the online platform, the Scheduled Caste and Scheduled Tribes (Prevention of Atrocities) Act for hate speeches against the so called 'oppressed' classes even today, (Protection of Civil Rights and Representation of People's Act also state about offensive content against individuals -though not directly since it does not apply in the online platforms), legal precedents which set a limit and examine the viewpoint of hate speeches online. It is important to note that although Instagram has its own content moderation

¹³ DrishtiIAS, "Hate Speech", (19th October, 2023, 5:00 a.m.), <https://www.drishtias.com/daily-updates/daily-news-analysis/hate-speech-5>

¹⁴ "Shreya Singhal v Union Of India", (19th October, 2023, 6:45 a.m.), <https://www.legalserviceindia.com/legal/article-10124-shreya-singhal-v-union-of-india-air-2015-sc-1523.html#:~:text=In%20this%20case%20court%20found,such%20words%20lead%20to%20an>

policies, it is obligated to comply with these Indian laws as well.¹⁵

Various research and case studies show that hate speeches occur in huge numbers within Instagram and individuals are targeted for belonging in a particular community or group. These studies also examine the different kinds of contents and modes used to convey a hate speech—through images, posts, videos, comments, captions, etc. Many surveys conducted by researchers have concluded the effects on the emotional and psychological well-being of a person who is being targeted.

Types and impact of hate speeches:

Hate speech can take place in various forms which includes: racist comments, religious discrimination, gender-based harassment, body shaming, cyberbullying as well as social and political hate. All these kinds of hate speeches target individuals based on specific attributes and for belonging to a certain community or group.¹⁶

These have a negative impact on the users. Mental Health effects like stress, anxiety, Social-divisiveness which may lead to divisions and conflicts between different groups and its influence on the youth of the State.¹⁷

Validity of Section 66A of the IT Act:

Through the *Shreya Singhal v Union of India* case¹⁸, the Supreme Court of India struck down Section 66A of the Information Technology Act 2000 stating that it is unconstitutional and violated Article 19 of the Indian Constitution that provided with the fundamental right to freedom of speech and expression. Section 66A of the IT Act stated that any offensive information sent using a computer or any other electronic devices would be punishable (punishment for three

¹⁵Deepika Yadav, “Analysis on Hate Speech”, (19th October, 2023, 11:15p.m.), <https://www.drishtiiias.com/daily-updates/daily-news-analysis/hate-speech-5>

¹⁶ Babak Bahador, “Classifying and identifying the intensity of Hate Speech”, (20th October, 2023, 3:45 p.m.), <https://items.ssrc.org/disinformation-democracy-and-conflict-prevention/classifying-and-identifying-the-intensity-of-hate-speech/#:~:text=Meanwhile%2C%20speech%20refers%20to%20communication,demonization%2C%20and%20in%20citement%20to%20violence>

¹⁷ United Nations , “Hate Speech is rising around the world”, (20th October, 2023, 4:45 p.m.), <https://www.un.org/en/hate-speech>

¹⁸ “*Shreya Singhal v Union Of India*”,(21st October, 2023, 5:40 p.m.), <https://www.legalserviceindia.com/legal/article-10124-shreya-singhal-v-union-of-india-air-2015-sc-1523.html#:~:text=In%20this%20case%20court%20found,such%20words%20lead%20to%20an>

years). This section was significantly used against hate speeches in the online platforms but was struck down for being vague and needed interpretation. This section was being misused by the police officials for arresting people on mere expression of their thoughts. Therefore, its vagueness that led to the striking down of the section called the attention for an effective way to address the issue of hate speech on the online platforms. However, until a new way is established, it becomes Instagram's responsibility to include within its internal guidelines to protect freedom of speech while restricting hate speeches in its platform. This also consists of other laws, rules and regulations as well. To sum it up, with the absence of Section 66A of the IT Act, there should be a collective action and effective collaboration done by Instagram and other social media to fight hate speeches in the online platforms.¹⁹ Therefore, it is clear that Section 66A cannot be rendered and used since it is vague and has more negative repercussions than positive results.

Challenges in defining hate speech online:

What exactly constitutes a hate speech? This has been an unanswered question even in the physical mode of interaction. Hate speech does not have a particular definition. It varies across countries and across different legislations within a country. The United Nations takes efforts in giving a proper definition to hate speech. While this definition may form the basis to decide cases related to hate speech, it is not totally relied upon by every member States due to the difference in cultural and contextual factors. This year (2023) the Secretary General of United Nations commented stating that we should challenge the discrimination faced by individuals by addressing the issues of hate speeches that is spreading faster on the internet. Hence it is essential to pay attention that having no consistency in the definition of hate speech amounts to more unsolved problems.

The dynamic nature of Instagram requires the platform to abide by the changing trends, Nuanced expressions through memes or other subtle forms of conveying messages need nuanced understanding of the circumstances and context, Subjectivity determines that hate speeches differ from individual to individual; what is considered a hate speech for one person may not be so for another person, Cultural relativity is present wherein hate speech may vary from different cultures as well, Regional-nuances dictates how historical, social, political or economic events affects the

¹⁹Law Audience Journal, Volume 5, Issue 2, Varun Kumar, "A critical Analysis of the invalidated Section 66A of the IT Act for Striking a balance between Free speech and Hate Speech Online", (21st October, 2023, 6:50 p.m.), <https://www.lawaudience.com/a-critical-analysis-of-the-invalidated-section-66a-of-the-it-act-for-striking-a-balance-between-free-speech-and-hate-speech-online/>

interpretation of a hate speech, Algorithmic bias may happen where only certain content is displayed which can form a biased content against a particular group; it is also very difficult to differentiate between a hate speech and a legitimate expression of opinions or thoughts through these algorithms, and lastly, Evolution of language in the online platform like the use of short forms and symbols makes it more difficult to identify a hate speech. These show the complexity of defining a hate speech. Therefore, it can be said that a precise definition of hate speech may restrict and deprive a person in performing their legal rights and duties even in the normal way. It is further noticed that it is better to interpret hate speeches from case to case depending on the gravity of the issue which will help in attaining a satisfactory justice system to all individuals.²⁰

Regulatory framework:

Regulatory framework is necessary to curb or rather control the issue of hate speeches that occur in the online platform. Instagram's content moderation and community guidelines provide with policies that each country has to abide with. These rules are also related to hate speeches. The key aspects of Instagram's guidelines include:

1. Prohibition of hate speech,
2. Reporting mechanisms wherein users can report contents or accounts,
3. Human moderators who review the reports made and decide if it is violative or offensive in nature,
4. AI and machine learning algorithms help restrict unessential content as well as preventive measures like Instagram's comment filters, anti-bullying measures,
5. Artificial intelligence driven cautions or warnings to prevent hate speeches on Instagram.

The effectiveness of these policies is still a topic of debate. It faces several challenges like algorithm bias where the machine is fed with certain information and does not have the ability to detect hate speeches which are not fed into its systems. Evolving tactics of how a hate speech is expressed which makes it more difficult to adapt and regulate it. Complexity of context where a simple sentence may sound offensive to one person but may not sound so to another person and the broader scale of the platform of Instagram with its vast number of users which makes it

²⁰ United Nations, "Understanding Hate Speech", (22nd October, 2023,8:40 p.m.), <https://www.un.org/en/hate-speech/understanding-hate-speech/what-is-hate-speech#:~:text=To%20provide%20a%20unified%20framework,person%20or%20a%20group%20on>

challenging to provide with daily regulation and inquire about the pending cases. Further, to improve these challenges in the regulatory framework we can reduce biases, enhance user education, transparency in content moderation and collaboration with experts.²¹

Problems in regulating Instagram specific laws:

- Jurisdictional complexities for a globally operating social media platform like Instagram makes it hard for a single set of laws to govern the issues faced.
- Varied definition of hate speech gives no consistency in regulation.
- Making the correct balance between content removal vs freedom of expression while addressing a hate speech in the laws is challenging.
- Enforcement and accountability of a vast platform like Instagram is difficult.
- Evolving nature of social media needs laws that evolve along with it which is slow to progress.

With change in trends, new ways of communication occur and with it the issue of hate speeches keep rising in huge amounts. Therefore, the already existing laws need strict implementation practically to see fruitful effects.

Case Studies and examples:

Understanding the gravity of the issue of hate speeches on Instagram is better possible through case laws and examples. Starting with the landmark judgement made in the Supreme court of India- *Shreya Singhal v Union of India*²²- that struck down Section 66A of the Information Technology Act, 2000 (IT Act). This case originated when individuals were being arrested for making critical comments on social media against politicians and other public figures. The primary issue addressed here was the constitutional validity of Section 66A of the IT Act, 2000 (Punishing a person for sending offensive messages through computers or other electronic devices). This section violated the fundamental right of freedom of speech and expression under Article 19(1)(a). Another issue addressed was whether intermediaries like social media could be

²¹ UNESCO, UNO on Genocide Prevention and the Responsibility to Protect, Ch-3, "Addressing Hate Speech on Social Media: Contemporary challenges", (22nd October, 2023, 11:40 p.m.), <https://unesdoc.unesco.org/ark:/48223/pf0000379177>

²² Case Analysis, "*Shreya Singhal v Union Of India*", (23rd October, 2023, 6:45 a.m.), <https://www.legalserviceindia.com/legal/article-10124-shreya-singhal-v-union-of-india-air-2015-sc-1523.html#:~:text=In%20this%20case%20court%20found,such%20words%20lead%20to%20an>

held responsible for offensive posts that were allowed to be made against another individual. The supreme court in its decision held that this section was very broad and vague and hence held it unconstitutional. And it even held that intermediaries like social media platforms cannot be held criminally liable unless it abides by the IT Act laws and rules and that it should not be entrusted with censoring and monitoring duties. The judgement here was significant since it emphasized the freedom of speech provided under the Constitution.

A few examples related to this include:

- Racial hate speech- when a user posts memes that shows racial discrimination. The user creates a hostile environment targeting an individual's ethnicity through images and contents that are offensive in nature.
- Gender-based harassment- a user making bigotry comments on the post of a public figure, aiming at his gender and appearance.
- Homophobic content- wherein multiple comments are made on the post of a LGBTQ+ activist or individuals, attacking the sexual orientation of the LGBTQ+ community.

Instagram's responses to these were the implementation of mixed approaches to tackle such hate speeches. It included content removal, user tools, warnings and report systems as mentioned earlier in this paper. Along with these, there are legal implications like civil suits to seek damages, criminal charges for severe cases to punish a culprit and platform responsibility in content moderation.

Implications and recommendations for the regulation of Hate

Speech on Instagram:

Implications For Instagram Users-

- Awareness- users are to be made vigilant of hate speech made online and should report such offensive messages.
- Safety precautions- safety tools as provided by Instagram should be made of good use by the users to prevent hate speeches.
- Victim support- Instagram should give support to the victims of hate speech and report such accounts that engage in posting offensive contents.

Implications For Policy Makers-

- Policy reform- laws should be adapted and refined along with the changing times.
- Collaboration- there should be global collaborations done to maintain healthy international relations who encourage regulation of laws regarding hate speeches made online.
- Transparency- content moderation procedures should be transparent in nature, allowing users to understand the process of decision making and making online platforms like Instagram accountable.

Implications For Social Media Platforms-

- Algorithm improvements- to get accuracy in detection of hate speeches, it is necessary to improve algorithms.
- User Education- it is important to spread awareness and educate people about hate speeches and its negative impacts on the society.
- Collaboration- collaboration with experts globally will enhance the working of content moderation and will teach about the changing trends of hate speech which can then be resolved.

Recommendation-

- Enhance transparency- Instagram should make its content moderation and guidelines transparent and give its users a clear warning of the after effects of engaging in hate speeches online.
- Reduce algorithm bias- mitigate the algorithmic biases and ensure all users are treated equally.
- User reporting- simplify the report mechanisms for reporting accounts, posts, captions, or any other messages that are offensive which could bring a sense of responsibility.
- Consistent standards- aim for consistent global standards which respects all cultural and regional differences.
- Continuous monitoring- to make better policies, it is important to keep checking on the new trends that have emerged.
- Education- educate people about the mishaps of engaging in abusive speeches in the online platform by conducting various campaigns.

- Protection of vulnerable groups- the minority communities are to be protected through the implementation of specific measures exclusively to protect them.
- Collaboration- in order to gain insights, it is necessary to collaborate with experts across countries to come up with effective regulations to govern hate speech made online.²³

Conclusion:

The critical analysis provided in this paper highlights several aspects of the topic, “Critical Analysis: Hate Speech on Instagram.” It delves deep into the dynamic nature of Instagram in which a hate speech occurs in various forms and kinds targeting an individual’s different attributes like race, religion, gender, caste, etc. The study also shows that due to its dynamic nature where different jurisdictions interpret hate speech in different ways, defining hate speech is very difficult. A hate speech is influenced by its cultural and contextual factors and is therefore interpreted accordingly. Even when social media platforms like Instagram have adopted precautions and measures to battle with issues of hate speech, there still exist many challenges that interrupt the regulation of laws made to govern these issues globally.

The research problem of “Hate Speech on Instagram” is essential to be educated about in today’s digital world. Instagram is a wide platform used by uncountable users and serves as the future mode of communication. This makes it necessary for us to know the impacts of hate speech on society. Finding and applying solutions to deal with hate speech while maintaining freedom of speech is essential to achieve a safe online environment.

This paper talks about Section 66A of the IT Act 2000, which was previously used against people who passed any kind of offensive messages through any electronic devices. The landmark case of *Shreya Singhal v Union of India* that struck down this section on the basis of it being vague and police officials misusing it is to be noted. The case showed how there should be a fine line between providing freedom of speech and restricting hate speech.

The study then talks about the dynamic and evolving nature of social media and the ways in which a hate speech is communicated online. Through various literature it may be concluded that we

²³ Poni Alice JameKolok (UNESCO), “5 ways to counter hate speech in the Media through Ethics and Self-regulation”, (25th October, 2023, 8:45 a.m.), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9133687/>

cannot give a specific definition to what constitutes a hate speech in the online platform since it keeps changing from time to time.

Further, the research explains the regulatory framework and the challenges it faces in governing issues in the Instagram platform. The paper provides with Instagram's community guidelines that abide with the changing nature of the issues faced as well as gives few recommendations to improve on laws to regulate Hate Speech on Instagram. There should be practical implementation of the laws and practical awareness spread across the globe regarding the issue of hate speech on Instagram.

To conclude, this research paper describes the broader challenges faced in the digital world by focusing on the issue of Hate speech on Instagram. It takes the efforts of collaborators, policy makers and the users together to understand and act cautiously and responsibly in order to effectively address these issues. Hence, we must take steps in building a safe and respectful online environment for the betterment of the society at large in the digital space where there is equal balance between giving freedom of speech and preventing offensive content online.

References:

- Zachary Laub, Hate Speech on Social Media: Global Comparisons, Council on Foreign Relations, 2-10, (2019).
- Catherine O'Regan, Hate Speech Online: an (Intractable) Contemporary Challenge?, 71 Curr. Legal Probs. 403 (2018).
- Alam, R.L. Raina & F. Siddiqui, Free vs Hate Speech on Social Media: The Indian Perspective, 14 J. Info. Comm. & Ethics in Soc'y 350 (2016).
- Debarati Halder, A Retrospective Analysis of Section 66A: Could Section 66A of the Information Technology Act be Reconsidered for Regulating 'Bad Talk' in the Internet?, 1 Indian Student L. Rev. 99 (2015).
- Ashraf, Mehvish, Online Hate Speech in India: Issues and Regulatory Challenges, 3 Int'l J.L. Mgmt. & Human. 919 (2020).
- Mainack Mondal, Leandro Araújo Silva & Fabrício Benevenuto, A Measurement Study of Hate Speech in Social Media, in Proceedings of the 28th ACM Conference on Hypertext and Social Media (HT '17) 85-94 (2017).
- Gábor Kovács, Patricia Alonso & Rajiv Saini, Challenges of Hate Speech Detection in

Social Media, 2 SN COMP. SCI. 95 (2021).

- Z. Waseem & D. Hovy, Hateful symbols or hateful people? Predictive features for hate speech detection on Twitter, in Proceedings of the NAACL student research workshop 88-93 (2016).
- MacAvaney, S., Yao, H.R., Yang, E., Russell, K., Goharian, N. & Frieder, O., Hate speech detection: Challenges and solutions, Vol 14, Plus One, (2019).
- Jose, A.R., Freedom of Speech and Expression and Social Media: An Exigency for Balancing, Part 1, 2 Indian J. Integrated Rsch. L. 2-14 (2022).
- C.R. Scott, Benefits and Drawbacks of Anonymous Online Communication: Legal Challenges and Communicative Recommendations, 41 Free Speech Y.B. 127 (2004).
- Sarah Jameson, Cyber Harassment: Striking a Balance between Free Speech and Privacy, 17 COMMLAW Conspectus 231 (2008).
- Arun, Chinmayi and Nayak, Nihal, Preliminary Findings on Online Hate Speech and the Law in India, Berkman Klein Center Research Publication, 2-15 (2016).
- Tiwari, S. & Ghosh, G., Social Media and Freedom of Speech and Expression: Challenges Before the Indian Law, 2-15 (2014).
- Peter Coe, Anonymity and Pseudonymity: Free Speech's Problem Children, Media & Arts Law Review, Forthcoming, 12-24, (2018).

IJLRA