Social Media and Hate Speech: A Twitter Example

Hatice Köybaşı
Sakarya University, Sakarya, Serdivan

The internet has brought together people from diverse cultures, backgrounds, and languages, forming a global community. However, this unstoppable growth in online presence and user numbers has introduced several new challenges. The structure of the cyberspace panopticon, the utilization of big data and its manipulation by interest groups, and the emergence of various ethical issues in digital media, such as deceptive content, deepfakes, and echo chambers, have become significant concerns. When combined with the characteristics of digital dissemination and rapid global interaction, these factors have paved the way for ethical problems related to the production, proliferation, and legitimization of hate speech. Moreover, certain images have gained widespread acceptance as though they were real, despite having no factual basis. This recent realization that much of the information and imagery considered to be true is, in fact, a virtual illusion, is a commonly discussed truth. The alarming increase and growing legitimacy of hate speech within the digital realm, made possible by social media, are leading us toward an unavoidable outcome. This study aims to investigate the reality of hate speech in this context. To achieve this goal, the research question is formulated as follows: “Does social media, particularly Twitter, contain content that includes hate speech, incendiary information, and news?” The study’s population is social media, with the sample consisting of hate speech content found on Twitter. Qualitative research methods are intended to be employed in this study.

Keywords: internet, social media, Twitter, hate speech, digitalization, digital media.

Introduction

New media is embraced by users as a tool for self-realization, socialization, and gaining status. Additionally, it enables users to develop different styles, approaches, and language skills, allowing them to create new representations. Through social language and style flexibility, its impact on content increases, making it more influential in reaching the target audience.

While the extraordinary advantages of new media stand out as we transition into the Web 3.0 era, serious doubts still persist about it as a source of information in today’s world (Mehrabi, Hassan, Sham, & Ali, 2009, p. 138).

In traditional media, the presence of content responsibility, sanctions, and penalties, along with the transparency of channel ownership, provides a certain level of trust for everyone. However, the technical anonymity of new media, gaps in information technology law, and uncertainties regarding content responsibility create a complex media landscape. The real issue lies in people filling new media with unreal representations for the sake of likes and popularity. This unreality unwittingly leads many individuals into the swamp of hate speech. As the gap between truths and falsehoods widens, the new media landscape is becoming an increasingly challenging space for individuals to navigate and make sense of.

Hatice Köybaşı, Master's Student, Social Sciences Institute, Master of Communication Sciences, Sakarya University, Sakarya, Serdivan.
When we consider Plato’s allegory of the cave as a metaphor for media, media appears to be a reflection where objective reality is alienated and existence is reinterpreted according to thought (Renvoisé & Morin, 2013). Thus, the new reality created by media can sometimes contradict the methods of defining truth used by the old mind, create a disconnect between appearance and image (Saussure, 1998), and ultimately give rise to a moral paradox, leading to a conflict between the brain and the heart.

Even though the user of new media may not be well-versed in the philosophical foundations of this media, they are aware of its effects and the power of manipulation and disinformation. In other words, users understand that the content in new media may be fictional when making significant decisions and attempt to verify the accuracy of the content by searching for more sources. Therefore, the thesis that “the presence of a message in other channels creates trust in traditional media” (Hackett, 1998, p. 21) is not as applicable to new media. With new media, news, information, content, and images disseminated through these platforms can influence the masses in different ways. They can fuel hate speech within society or lead to unexpected negative outcomes. They can give rise to mass prejudice against a race or community.

Social media can take on various forms and serve different purposes. Some popular social media platforms include Facebook, Twitter, Instagram, LinkedIn, TikTok, and YouTube. These platforms allow individuals to connect with users worldwide, manage and curate news, build and maintain professional and social networks, and much more.

Social media is generally known for its fast-paced nature, where information is quickly disseminated. This environment is where the examination of the reality of hate speech is intended. Twitter, in particular, is one of the most significant platforms where information spreads rapidly across user bases. This platform presents numerous ethical challenges, one of which is hate speech.

This study aims to be conducted using a qualitative research method to explore the phenomenon of hate speech within this context.

This study is aimed at examining the approach of new media outlets towards hate speech and the problems arising as a result. The population of the study is social media, and the sample consists of hate speech content found on Twitter.

Within this study, we will delve into hate speech encountered in new media outlets. The issue of hate speech, particularly on social media platforms such as Twitter, will be addressed. In addition, the study will explore what constitutes hate speech, its ramifications, its rapid dissemination characteristics, and how it can be understood and comprehended.

The main problem statement of this study is formulated as follows: “Does social media, especially Twitter, contain content that includes hate speech, incendiary information, and news?” This question serves as the central focus of the study, aiming to shed light on the presence of hate speech in the realm of social media, particularly on Twitter, and its potential for spreading incendiary content and news.

The Concept of Hate Speech

According to the definition provided by the Turkish Language Association (http://www.tdk.org/), hatred is the desire for someone’s harm, unhappiness, or discomfort, accompanied by aversion. The term “speech” has a broad range of meanings and perspectives in sociological, philosophical, and political terminology. People can fundamentally hate many things. For example, they may not like scorpions, snakes, car horns, or pink phosphorescent shoes. However, when hatred is directed towards other human masses, the “others”, the course
of the situation changes, and it takes on a different form. There is a significant difference between hating scorpions, snakes, or pink phosphorescent shoes and harboring hatred based on ethnic origin. The latter relies on categories and generalizations that carry many biases against the “others” and tends to reduce individuals or communities to a negative point. For example, the prejudice that a particular ethnic origin is “dirty” directly or indirectly labels all individuals of that ethnic identity as “dirty” and reduces these individuals solely to this presumed characteristic. As a result, acting on the prejudice that a certain ethnic origin is “dirty” makes all individuals of that ethnic identity targets of hatred and subjects them to any harmful actions that may arise from such hatred (Ersoy, 2018, p. 119). The consequences of these harmful actions can lead to adverse outcomes for that ethnic identity or group.

Distinctive factors and conditions that differentiate societal structures include unique history, language, culture, needs, interactions, and the systematics of human relationships. When people define themselves and feel a sense of belonging in their society, they use the concept of “us”, but in this process, they also introduce the concept of “other”. Stereotypes, one of the steps that help identify or define the “other”, are a collection of generalized and categorized images in our minds related to individuals or groups (Öztürk, 2021, p. 144). This state of being the “other” or the act of labeling someone as the “other” can lead to hatred.

Discriminatory attitudes and behaviors often begin with the process of “othering”. The construction of the “other” manifests as a process that highlights differences, depicting those who exist outside the elements that make up our identity. Hate speech, which forms the mental and ideological foundations of hate crimes and these behaviors, emerges as the most severe form of discrimination. Hate speech, despite containing negative emotions, actually encompasses an ideology of exclusion, devaluation, and hostility among groups in society, pointing to a systematic hierarchy (Göregenli, 2012, p. 178).

The Hate Pyramid explains a process that starts with prejudiced attitudes and progresses step by step to prejudice, discrimination, violence, and ultimately, genocide. In this process, hatred begins with the formation of prejudices and then manifests in words and behaviors that can be described as hate speech. It can then escalate to engaging in discrimination and, finally, may lead to the commission of hate crimes—a dangerous progression.

There are certain distinctions between hatred and hate speech. Hatred represents a state of mind directed towards the “others”, often fueled by anger, fear, and a sense of the unknown or unfamiliar. On the other hand, hate speech is a negative form of expression that arises from all forms of intolerance, feeds on prejudices, and promotes, defends, or justifies hatred (Karadeniz, 2012, p. 1).

Within societal structures, media plays a significant role in the dissemination and distribution of desired values (Yanıkkaya, 2009, p. 11). Through various communication tools and platforms, the media can convey these values in different forms and missions by providing information, news, entertainment, and cultural content to society. Among the functions of the media are shaping the values of society, transmitting social norms, and generally influencing the thoughts and behaviors of the community. This can sometimes lead to positive outcomes, but at other times, it can result in negative consequences.

Especially with the emergence of second-generation internet technologies, social media has become a favorite as a new stage of mass communication. Social media, with its features like instant and two-way communication, strengthens the influence of the media day by day. Discriminatory attitudes that exist in daily life and hate speech generated in traditional media spread more easily and widely through the comprehensive features of social media. On social media platforms, anyone with internet access can create and share content, reinforcing the image of social media as independent, democratic, and decentralized, pushing it towards an
unstoppable point. The production of hate speech in social media environments and the incredible speed at which these speeches spread can serve as a precursor (Göregenli, 2012, p. 182). This situation can lead to minority groups and individuals from different ethnic backgrounds being easily subjected to hate speech within society.

Figure 1. Hate Speech Pyramid, (https://www.adl.org/sites/default/files/documents/pyramid-of-hate.pdf).

Hate Speech and Social Media

Social media refers to a new mass communication structure that emerged with the bidirectional and real-time information sharing capabilities provided by second-generation internet technologies (Web 2.0). Unlike traditional media, where communication is unidirectional and controlled by a specific class, social media has replaced passive viewers who cannot create content, intervene in produced content, or make comments with active viewers who can do all these things. In other words, we can say that social media is a structure that turns every individual with internet access into a potential mass communication tool. The content created by individuals, driven by concerns such as being followed, shared, and liked, has replaced the content created by traditional media, which often prioritizes interest groups and economic considerations.

In this context, it would not be accurate to say that hate speech spontaneously originates from social media. Hate speech is produced by individuals within the framework of society’s social, cultural, and political structure, its history, traditions, and customs (Varış & Avşar, 2022, p. 353).

Social media is moving towards a new phase in terms of the spread and impact of hate speech. One of the most significant effects of social media on hate speech is its ability to facilitate the rapid spread of such speech. The technological structure of the internet clearly shows the relationship between the production and dissemination of hate speech. Hate speech on the internet is not a new phenomenon, but compared to traditional media, it has gained more visibility and a wider reach (Karaman & Işikli, 2016, p. 3).
Social media platforms enable people to easily create and share content and interact with other users. This allows hate speech to reach a broader audience and spread more quickly. Furthermore, the algorithms on social media platforms direct users to specific content and facilitate the coming together of individuals with similar views, significantly increasing the speed of hate speech dissemination. Individuals with similar negative views can play a leading role in the spread of hate speech through these platforms.

The Reality of Hate Speech on Twitter

Twitter, a social media platform, was introduced into our lives for the first time in July 2006 by Jack Dorsey, Noah Glass, Biz Stone, and Evan Williams. Twitter was created, taking inspiration from SMS technologies, and allows users to write short messages limited to 280 characters, known as “tweets”. In 2013, Twitter achieved the distinction of being one of the top 10 most visited websites (Vayes, 2019). Twitter has become a part of our daily lives as a social network that enables users to share their thoughts instantly, follow news, interact with other users, and share content. Users can view, like, retweet, and comment on the tweets of the people they follow. The platform also allows users to track topics and see trends through the use of hashtags. Twitter is a highly popular platform where users globally share up-to-date information and express their opinions.

Twitter serves as a platform that not only hosts a wide range of different ideas and content but also contributes as much as possible to the spread of positive support mechanisms, such as assistance, philanthropy, and the swift resolution of legal cases. However, at times, alongside hostile expressions that affirm incorrect, derogatory, and prejudiced feelings, publications about an individual, group, or community can also be observed. With the emergence of prevalent social networks in recent times, the spread of extremism associated with hate speech against minority groups has become easier (Williams & Burnap, 2016, p. 212). The rate of encountering hate speech related to topics such as race, belief, and gender on Twitter is more frequent compared to mainstream media. This can be attributed to the platform’s characteristics, such as openness, rapid dissemination, and wide access to audiences (Karaman & Işıklı, 2016, p. 11).

Twitter’s communication process can lead to different interpretations due to its allowance of asynchronous communication and anonymity. Users can comfortably and confidently act with a certain level of narcissism behind the keyboard, thanks to the absence of a continuous feedback loop and the anonymity it offers. This can lead to the use of discriminatory, aggressive, or harsh language. However, this mode of usage may not reflect the language they would use in their concrete social lives and may emerge due to the influence of crowd psychology. Therefore, the harshness and discrimination in the language used on Twitter, along with the asynchronous and anonymous nature of communication, can lead to different interpretations and actions (Varış & Avşar, 2022, p. 356). This can negatively affect the attitudes of masses towards specific groups. Individuals, groups, or communities facing hate speech may find themselves feeling helpless, isolated, and fearful.

Method

Social media is generally an environment where news is quickly obtained and has a rapid dissemination feature. The aim of this study is to examine the reality of hate speech within this environment. Twitter is one of the most important platforms where any information can spread rapidly among masses. This platform is rife with ethical issues, and one of them is hate speech. This study aims to be conducted using a qualitative research method.

This study aims to examine the approach of new media outlets to hate speech and the problems arising as a result. The scope of the study is social media, with a focus on hate speech content found on Twitter.
In this study, hate speech encountered in new media outlets will be investigated, particularly the issue of hate speech on social media, especially on Twitter. The study will delve into what hate speech entails, its consequences, its rapid dissemination characteristics, and how it can be understood and comprehended.

The primary research question for this study is, “Does social media, especially Twitter, contain provocative information and news that includes hate speech?”

**Findings**

*Figure 2.* Twitter News Sample-1.

*Figure 3.* Twitter News Sample-2.
Figure 4. Twitter News Sample-3.

Figure 5. Twitter News Sample-4.

Figure 6. Twitter News Sample-5.

Conclusion

Social media, especially Twitter, can be influenced by the agenda of traditional media. In the context of freedom of expression, institutions and individuals who have the potential to influence public opinion should carefully choose their statements on such platforms. It is especially crucial for high-profile figures such as
politicians, artists, or journalists with a significant number of followers on Twitter to handle expressions containing hate speech with sensitivity.

At this point, the responsibility of social media users, particularly those who use platforms like Twitter as influential figures, is to share more sensible and calming content rather than inciting hate speech through news or comments. Journalists, politicians, artists, and other relevant individuals should refrain from posting information on Twitter that could fuel hate speech and should take steps to prevent it.

In this way, contributions can be made to establish a consensus and a peaceful environment in society, and efforts can be made to combat hate speech.

News, comments, and evaluations made in traditional media should use language that does not evoke hate speech. This is important because these contents can be viewed not only by a specific audience of traditional or social media followers but also by a broader audience and can lead to irreversible negative consequences.

Therefore, it is crucial that news, comments, and evaluations are prepared with sensitivity that promotes cultural and ethnic peace and does not encourage violence and hatred.

References