

Article

Hate speech in higher education: Exploring student perspectives and its impact on democratic values

Kimete Canaj¹, Artan Nimani¹, Shqipe Husaj¹, Blerim Canaj^{2,*}¹ University of Gjakova “Fehmi Agani”, Gjakova 50000, Kosovo² UBT—University of Business and Technology, Prishtina 10000, Kosovo* **Corresponding author:** Blerim Canaj, blerim.canaj@ubt-uni.net

CITATION

Canaj K, Nimani A, Husaj S, Canaj B. (2025). Hate speech in higher education: Exploring student perspectives and its impact on democratic values. *Journal of Infrastructure, Policy and Development*. 9(1): 10918. <https://doi.org/10.24294/jipd10918>

ARTICLE INFO

Received: 16 December 2024

Accepted: 13 January 2025

Available online: 16 January 2025

COPYRIGHT



Copyright © 2025 by author(s).

Journal of Infrastructure, Policy and Development is published by EnPress Publisher, LLC. This work is licensed under the Creative Commons Attribution (CC BY) license. <https://creativecommons.org/licenses/by/4.0/>

Abstract: Hate speech in higher education institutions is a pressing issue that threatens democratic values and social cohesion. This research explores student perspectives on hate speech within the university setting, examining its forms, causes, and impacts on democratic principles such as freedom of expression and inclusivity. This research is extended to determine the debates and theories elaborated from different perspectives qualitative and quantitative analysis of data collected from 108 participants at Higher Education in Kosovo. From the communication standpoint, analyzing hate speech in the media and social media is key to understanding the type of message used, its emitter, how the message rallies supporters, and how they interpret message. The findings highlight the need for proactive policies and educational interventions to mitigate Research on hate speech in higher education in Kosovo is crucial for fostering social cohesion and inclusivity in its diverse society. Hate speech undermines the academic environment, negatively affecting students’ mental health, learning outcomes, and overall well-being, necessitating efforts to create safer educational spaces. The study aligns with Kosovo’s aspirations for European integration, emphasizing adherence to human rights and anti-discrimination principles. Despite the issue’s significance, there is a lack of empirical data on hate speech in Kosovo’s higher education, making this research vital for evidence-based policymaking. With a youth-centric focus, the study aims to educate and empower young people as future leaders to embrace respect and inclusivity. By addressing hate speech’s local challenges and global relevance, the research supports institutional reforms and offers valuable insights for post-conflict and multicultural societies. Hate speech while fostering a culture of mutual respect and democratic engagement.

Keywords: Hate speech; freedom of expression; democratic values; free speech; social media

1. Introduction

Hate speech, defined as expressions that incite hatred, discrimination, or violence against individuals or groups, poses significant challenges to the democratic ethos of higher education institutions (Council of Europe, 2020). Universities are expected to be location for free speech and intellectual debate, yet they are not immune to the pervasive influence of hate speech, both online and offline. This paper investigates the prevalence and impact of hate speech in higher education, focusing on student perspectives and the broader implications for democracy.

In the context of Kosovo, society face unique global challenges related to hate speech, particularly concerning ethnicity, gender, religion, and politics. Incidents of inflammatory remarks targeting specific ethnic groups, sexist language against women, and misuse of political rhetoric have been reported. Such examples underscore the urgent need to address hate speech within this specific sociopolitical framework. Hate

speech in this setting is often described as “a silent poison,” eroding the values of mutual respect and inclusivity that universities aim to uphold.

2. Methodology

A mixed-methods approach was adopted, combining quantitative and qualitative methods. 108 questionnaires were distributed across universities, focusing on students' experiences and perceptions of hate speech in society and through social media. The questionnaire included both closed and open-ended questions to capture diverse insights. Additionally, in-depth interviews with 20 students provided qualitative data, offering a deeper understanding of the personal and communal impacts of hate speech within the social media context. Research on hate speech in higher education (HE) in Kosovo employed a sequential mixed-methods design, integrating quantitative and qualitative phases from March to November 2024. In the first phase, surveys and questionnaires were distributed to 108 correspondents to measure the prevalence and patterns of hate speech, with data analysis highlighting key themes and outliers. Building on these findings, the second phase involved qualitative methods, such as in-depth interviews or focus groups with selected participants, to explore motivations and contextual factors. This phased approach ensured that quantitative results informed the design of qualitative inquiries, enriching the study with a deeper understanding. The integration of both methods provided comprehensive insights, facilitating targeted exploration and actionable outcomes. By combining baseline data with contextual analysis, the methodology established a robust foundation for shaping public policies and addressing hate speech effectively in HE.

The questionnaire in the study measured the prevalence, perception, and impact of hate speech in higher education (HE) using closed-ended and Likert-scale items for both quantitative and nuanced insights. Reliability was confirmed with a Cronbach's alpha score exceeding 0.80, indicating high internal consistency. Content validity was established through expert reviews from education, sociology, and linguistics specialists. Construct validity was verified by piloting the questionnaire with a smaller sample to ensure it measured the intended variables. Likert-scale items, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree), allowed for detailed measurement of attitudes and perceptions. The questionnaire was culturally adapted to ensure linguistic and contextual appropriateness for respondents in Kosovo. These psychometric properties ensured the tool's robustness for collecting reliable and valid data. The approach provided a strong foundation for qualitative analysis and informed policy recommendations.

2.1. Sample size determination

The study aimed for a statistically representative sample using a 95% confidence level and a 5% margin of error. Based on standard sample size calculation methods and the estimated population size of individuals exposed to or discussing hate speech in Kosovo, the target sample was 108 participants.

2.2. Sampling method

A stratified random sampling approach was utilized to ensure proportional representation across age groups, genders, and geographic locations. The strata included urban and rural populations, emphasizing young people active on social media, where hate speech is frequently encountered.

2.3. Questionnaire design

The questionnaire was adapted from established scales, such as the Hate Speech Tolerance Scale, and included:

- Structure: Closed-ended Likert-scale items measuring the prevalence, impact, and perceptions of hate speech, along with demographic questions for subgroup analysis.
- Psychometric Properties: Reliability: Cronbach's alpha ranged from 0.78 to 0.85, confirming internal consistency. Validity: Expert review ensured content validity and construct validity was confirmed via exploratory factor analysis.

This methodological framework ensures robust data collection and analysis for examining hate speech perceptions in Kosovo

3. Literature review

The existing body of literature highlights the dual challenge universities face in upholding free speech while countering hate speech. Studies indicate that hate speech in educational settings often correlates with broader societal divisions and reflects underlying issues such as discrimination and marginalization (UNESCO, 2015). Research also emphasizes the psychological and social impacts of hate speech on students, including diminished academic performance and reduced participation in campus activities (Parekh, 2012).

Specific studies in the Western Balkan countries after the last war have pointed to the role of historical and political tensions in exacerbating hate speech. In Kosovo, for instance, hate speech often manifests along ethnic lines, reflecting unresolved conflicts and societal divisions (European Commission, 2022). Similarly, the influence of patriarchal norms contributes to the prevalence of gender-based hate speech, targeting women who challenge traditional roles. As one report aptly states, "hate speech in these institutions often serves as both a mirror and a magnifier of societal inequities." Altman et al. (2012) explore the tension between freedom of expression and human rights law, using Holocaust denial as a critical case study to question the boundaries of permissible speech. Archakis et al. (2018) investigate how humor masks underlying racist attitudes in public discourse, emphasizing the need for critical analysis of anti-racist campaigns. Assimakopoulos et al. (2017) adopt a discourse-analytic approach to highlight how hate speech manifests within the European Union and its sociopolitical implications. Awan (2016) documents the rise of Islamophobia on social media platforms like Facebook, revealing the normalized hate targeting Muslims through qualitative analysis. Barlow and Awan (2016) examine the silencing of women and Muslims in academia via online threats, showcasing the intersection of gendered and religious hate in digital spaces. Ben-David and Matamoros-Fernández (2016) analyze the covert discrimination present in

the social media accounts of far-right political groups, indicating how hate speech adapts in online environments. Boeckmann and Turpin-Petrosino (2002) conceptualize the broader societal harms of hate crimes, arguing for proactive policy frameworks to mitigate their effects. Burnap and Williams (2015) use machine learning to study hate speech on Twitter, providing a foundation for data-driven policymaking against online hate. Cammaerts (2009) critiques the tension between radical pluralism and free speech in digital public spaces, suggesting the need for nuanced regulation to balance these ideals. Herz and Molnár (2012) present a comprehensive analysis of hate speech regulations across different contexts, urging policymakers to consider cultural and legal variations when crafting responses. They multifaceted nature of hate speech, revealing its evolution across platforms, cultural settings, and legal interpretations.

The reviewed literature highlights the complexity of hate speech across media and societal contexts. Londo (2021) investigates the interplay between hate narratives and disinformation in Albanian online media, emphasizing their socio-political implications. McNamee et al. (2010) and Meddaugh and Kay (2009) explore the mechanisms of online hate groups, noting the ethical dilemmas and normalization of discriminatory ideologies within specific platforms. Murthy and Sharma (2019) present online hostility as a networked phenomenon, visualizing its prevalence through YouTube comments, while Nguyen (2023) suggests integrating public health frameworks and AI to combat hate speech effectively. Fortuna and Nunes (2018) provide a systematic overview of automated hate speech detection, offering insights into computational advancements and challenges in identifying hateful content.

4. Data analysis and research findings

The analysis of 108 respondents aged 18 and above highlights the interplay between hate speech and democracy. The sample consisted of 63% BA, 15% MA, and 20% PhD holders, reflecting diverse educational perspectives. Hate speech is widely perceived as a threat to democratic principles, particularly freedom of expression, tolerance, and equality. Respondents with higher academic qualifications (MA and PhD) demonstrated a deeper understanding of its systemic impact, emphasizing its role in undermining minority rights and polarizing society. Younger participants, primarily BA holders, focused on the emotional and personal effects, especially its prevalence on digital platforms. The findings underscore the need for civic education, stricter regulations on online hate speech, and policies that balance freedom of expression with inclusivity, ensuring democratic resilience against disinformation and divisive rhetoric.

The prevalence of hate speech within academic settings has emerged as a significant concern, particularly on social media, where 68% of surveyed respondents reported witnessing or experiencing it. Hate speech frequently targets individuals based on ethnicity, gender, and political beliefs, creating “invisible chains” that hinder open dialogue and mutual understanding. Ethnic minority students reported being subjected to derogatory remarks both in classroom discussions and online forums, with one stating, “It feels like our identities are used as weapons against us.” Similarly, female students noted frequent instances of sexist comments, especially in male-

dominated subjects, exemplified by one respondent's observation: "Being a woman with an opinion often makes you a target for ridicule." Religious and political affiliations also became tools for exclusion, particularly during election periods, as one student remarked, "The rhetoric during election seasons turns into a battlefield."

These experiences have profound implications for democratic values within universities. While 72% of students agreed that hate speech undermines democracy, only 48% believed their institutions had effective policies to address the issue. This gap reflects widespread confusion about the boundary between free speech and hate speech, underscoring the need for targeted educational initiatives. The impact extends beyond personal experiences, eroding trust among students, fostering hostility, and deterring participation in democratic processes. As one student shared, "I stopped attending debates because the atmosphere became toxic due to hateful comments." These findings highlight the urgent need for comprehensive policies and educational programs to counteract hate speech and promote a more inclusive and democratic academic environment.

The research on hate speech in higher education (HE) in Kosovo introduces several novelties. It provides localized insights by addressing the prevalence and impact of hate speech within Kosovo's specific multicultural and post-conflict context, an area previously under-researched. For instance, it examines how hate speech manifests in universities with diverse ethnic compositions. The sequential mixed-methods design combines quantitative surveys with qualitative interviews, allowing for a deeper understanding of patterns and motivations. For example, if quantitative data reveals high instances of hate speech in specific universities, qualitative interviews explore underlying reasons.

The study adopts a youth-centric approach, focusing on students as future leaders, and examines how hate speech affects their mental health and learning outcomes. By tailoring research tools to Kosovo's linguistic and cultural context, such as using culturally sensitive survey items, it ensures accuracy and relevance. It also generates the first extensive empirical dataset on hate speech in HE in Kosovo, filling a critical knowledge gap. With a focus on policy-driven outcomes, the findings aim to support reforms, such as introducing stricter codes of conduct and awareness campaigns. The research offers a holistic view of hate speech by integrating interdisciplinary perspectives from sociology, education, and linguistics. These novelties position the study as a significant contribution to fostering inclusivity and shaping actionable solutions in HE institutions in Kosovo.

Another student from Kosovo shared, "During election campaigns, political rhetoric on campus becomes divisive, turning discussions into personal attacks instead of constructive dialogue." To understand how youth perceive and comprehend the importance of definitions for terms such as misinformation, disinformation, malinformation, clickbait, and bait, a structured approach has been adopted. Such terms are presented in the **Table 1**.

Table 1. Definitions of mis/dis/mal-information, clickbait and bait.

Misinformation ¹	Disinformation ²	Mal-information ³	Clickbait ⁴	Bait ⁵
Misinformation is information that is false, but the person who is disseminating it believes that it is true.”	Disinformation is information that is false, and the person who is disseminating it knows it is false. It is a deliberate, intentional lie, and points to people being actively disinforming by malicious actors.”	Mal-information is information “that is based on reality, but used to inflict harm on a person, organization or country”	Something (such as a headline) designed to make readers want to click on a hyperlink, especially when the link leads to content of dubious value or interest.	To persecute or exasperate with unjust, malicious or persistent attacks (baiting organisation country.” or Something (such as a headline) designed to make readers want to click on a hyperlink, especially when the link leads to content of dubious value or interest. minority groups). To try to make angry with criticism or insults. (Baiting a politician during a debate.)

The role of the media in generating and spreading hate speech in the WBs is recognized in many reports. As stated in the RESILIENCE report ‘Hate Narratives in the Western Balkans and Turkey’ in the narrative strategies exposed in both media and user content, clear features of fascism, mainly racism, nationalism, intolerance to difference and misogyny, can be found (Hrvatín et al., 2021). Hate narratives are not isolated cases of specific individuals but results of devised campaigns that use different methods, such as biased reporting, mis/disinformation and spread of fear (Sokol, 2021). Even when the mainstream media do not produce hate speech narratives on their own, they become a means for carrying them (Londo, 2021). At the end, the media also play a negative role when they fail (due to lack of resources and/or lack of will) to adequately deal with those user comments in which hate narratives particularly escalate (Council of Europe 2013, 2020; Eurobarometer 2016).

- Hate speech and hateful narratives are used to mobilize support for nationalist and exclusionary ideologies targeting specific ethnic or religious groups and to dehumanize and delegitimize these groups, making them vulnerable to discrimination and violence.
- Hate speech in the WBs contributes to the perpetuation of discrimination and violence against targeted groups, and it can also radicalise society in a manner that can lead to violent extremism. Therefore, the response to hate speech, taking the complexity of the phenomenon into account, should be multifaceted and multisectoral. The media’s role in generating and disseminating hate speech in the Western Balkans is widely acknowledged in numerous reports. According to the RESILIENCE report “Hate Narratives in the Western Balkans and Turkey”, narrative strategies evident in both media content and user-generated content display clear features of fascism, including racism, nationalism, intolerance toward diversity, and misogyny (Hrvatín et al., 2021). Hate narratives are not the product of isolated individuals but rather the result of coordinated campaigns employing various tactics, such as biased reporting, misinformation, disinformation, and the propagation of fear (Sokol, 2021). Even when mainstream media do not create hate narratives themselves, they often serve as conduits for their dissemination (Londo, 2021). Moreover, the media play a detrimental role when they fail—whether due to limited resources or lack of willingness—to address user comments where hate narratives escalate significantly (RESILIENCE, 2021).

Hate speech and hateful narratives are used to mobilize support for nationalist and exclusionary ideologies, targeting specific ethnic or religious groups. They aim to dehumanize and delegitimize these groups, making them more susceptible to discrimination and violence. In the Western Balkans, hate speech contributes to ongoing discrimination and violence against marginalized groups and can radicalize society, potentially leading to violent extremism.

Given the phenomenon's complexity, addressing hate speech requires a multifaceted and multisectoral approach. Institutional efforts should be complemented by contributions from other stakeholders, particularly civil society organizations.

4.1. Discussion

The findings underscore the tension between protecting free speech and curbing hate speech in academic settings. Democratic values thrive in environments where diverse perspectives are respected, yet hate speech undermines this ideal by silencing marginalized voices (Parekh, 2012). In Kosovo's HEIs, ethnic and gender-based hate speech exacerbates societal divisions and hinders the development of inclusive academic communities. Hate speech has been described as "a rhetorical wildfire," spreading quickly and leaving a trail of damaged relationships and lost opportunities for dialogue.

Addressing these issues requires a nuanced approach that considers the local context. For example, efforts to counter ethnic-based hate speech must involve cross-cultural dialogues and reconciliation initiatives, while gender-based hate speech necessitates challenging entrenched patriarchal norms. Universities must serve as "fortresses of understanding," where dialogue and inclusivity stand as bulwarks against division. (Gracia – Caladin et al. 2023, Gallaher et al 2021, Gelber 2002)

By understanding student perspectives and fostering a culture of inclusivity and respect, universities can uphold their role as champions of democracy and intellectual freedom. In Kosovo, tackling hate speech requires addressing the unique intersections of ethnicity, gender, religion, and politics, ensuring that HEIs remain spaces for constructive dialogue and democratic engagement. "Universities," as one respondent noted, "should be sanctuaries of learning, not battlegrounds of division." (Slagle 2009).

4.2. Hate speech vs. freedom of expression and democracy's dilemma

This research delves into the intricate dynamic of freedom of expression within the contexts of modern democracies of freedom of expression with examples from Kosovo. As a fundamental pillar of democratic governance, freedom of expression stands as both a cornerstone of individual liberty and a potential source of challenges (Constitution of Kosovo Article 40 Freedom of Expression) and Law Nr. 04/L-44. This study aims to unravel the complexities surrounding this essential democratic tenet, exploring its historical foundations, legal frameworks, and evolving societal expectations.

The research scrutinizes the delicate balance between the preservation of free expression and the prevention of harm, focusing on the nuances presented by hate speech in media. Examining the impact of hate speech on democratic ideals, the study investigates how Kosova societies grapple with the tensions between safeguarding

individual liberties and fostering a harmonious and inclusive public sphere. “Idiot” (idiot), “thief” (hajn), “criminal” (kriminel), and “fraudster” (mashtues). All these epithets, at first glance, not only appear but are expressions that in principle seem difficult to use even in an informal communication, be it a coffee shop or similar discussions. However, these expressions have already become part of everyday life in public discourse, starting from social networks, and can be heard in almost every television show that is broadcast in “prime time” on some televisions.

This article will present some of the cases from the questions to the students and during debates in the audio-visual media in Kosovo. Through a comprehensive analysis of case studies, legal precedents, and contemporary examples, this research provides insights into the evolving nature of freedom of expression in the digital age. It evaluates the role of technology and social media in shaping the discourse, assessing the implications for democratic processes and the potential challenges posed by the rapid dissemination of information. Ultimately, this research contributes to a nuanced understanding of the multifaceted relationship between freedom of expression, hate speech, and the democratic values. By shedding light on the complexities inherent in balancing individual freedoms with societal well-being, it offers valuable perspectives for policymakers, scholars, and citizens engaged in the ongoing discourse on the preservation and enhancement of democratic principles. Hate speech refers to any form of communication, conduct, writing, or expression that offends, threatens, or insults an individual or group based on attributes such as race, ethnicity, religion, gender, sexual orientation, disability, or other characteristics. It often involves the use of derogatory language, stereotypes, or inflammatory remarks with the intent to discriminate, marginalize, or incite violence against the targeted individuals or groups.

4.3. Language and system barriers

Language evolves rapidly, especially among younger generations who often communicate through social networks, which requires ongoing research into hate speech datasets. Online platforms are actively removing hate speech content both manually and automatically. However, individuals who spread hate speech constantly seek new ways to bypass any imposed restrictions. For example, some users post hate speech as images containing hateful text, which can evade basic automated detection systems. While image-to-text conversion could address some of these challenges, it still faces limitations and does not fully solve the problem of automatic hate speech detection. Additionally, altering language structure presents another challenge, such as using unfamiliar abbreviations or mixing languages, for example: i) Writing part of a sentence in one language and the rest in another; ii) Writing a sentence’s phonetics using a different language (e.g., writing Albanian sentences using English letters).

Hate speech takes many forms, such as verbal insults, written statements, online content, images, or actions that reinforce bias and animosity (a strong feeling of dislike or hatred). It harms not only those directly affected but also creates a harmful social climate, deepening divisions and inciting hostility within communities. Addressing hate speech typically involves fostering education, raising awareness, encouraging tolerance, and, when needed, enforcing legal protections. Striking a balance between safeguarding free speech and preventing harm presents a significant challenge for

societies tackling this issue. Promoting public awareness, encouraging community involvement, and nurturing inclusive discussions are crucial steps in reducing the effects of hate speech and fostering more respectful and cohesive communities (Grant and Both, 2009).

Examines the justification for hate speech bans amid ongoing global debates and evaluates the state of free expression in contemporary democracies.

Distinguishes between rights-based and democratic models, emphasizing the democratic approach

Emphasizes the relationship between general principles of free speech and their historical contexts

As Kalven (1967) states:

“The instrument of dissent and criticism is the individual faculty member or the individual student. The university is the home and sponsor of critics; it is not itself the critic. ... The neutrality of the university as an institution arises then not from a lack of courage nor out of indifference and insensitivity. It arises out of respect for free inquiry and the obligation to cherish a diversity of viewpoints. ... there emerges, as we see it, a heavy presumption against the university taking collective action or expressing opinions on the political and social issues of the day ...”

This explains the justification for banning hate speech within global debates on free expression, focusing on how democratic and rights-based models frame the issue. It highlights the historical contexts shaping free speech principles, underscoring the importance of safeguarding diverse viewpoints. As Kalven (1967) articulates, universities should remain neutral institutions that prioritize free inquiry, fostering criticism through individual voices while refraining from collective stances on social and political matters.

4.4. Hate speech and education on democratic citizenship

Hate speech goes beyond merely disparaging individuals based on their inherent identity traits. It serves as a tool wielded by those in power to preserve their social, political, or economic dominance, relegating its victims to subordinate roles and obstructing true equality. This calls for a reevaluation of the broad protections granted to hate speech under the Kosovo Constitution Article 40 civilians should leverage existing civil legal mechanisms to address freedom of expression. Allowing civil penalties for the subjugation caused by hate speech would create opportunities for marginalized groups to advance toward equality. (Kosovo 2022 Human Rights Report).

There is widespread agreement on the distinct features of hate speech on social media. These platforms have transformed the way racism, for instance, is represented—a transformation shaped by the specific characteristics of the platforms themselves. Unlike hate speech disseminated through traditional media, online racist discourse is often marked by a lack of emotional restraint. This can be attributed not only to the anonymity social media provides but also to the diverse methods available for expressing such ideas and the normalization of new cultural frameworks, such as the proliferation of post-racist narratives and approaches (Chiluwa, 2018; Harlow, 2015; Haxholli, 2022).

Twitter discourse is usually simple, impulsive, and offensive (Slagle, 2009; Waldron, 2012). Antagonistic and negative sentiments, such as those seen in Islamophobia, are expressed through complex identities that intersect with religion, ethnicity, politics, and gender (Vargo et al., 2020; Wodak, 2002). Exclusionary rhetoric is particularly prevalent, as exemplified by discourse surrounding refugees in Europe. Notably, European nationalist and racist groups frequently use Twitter to propagate socially acceptable forms of racism (Sugrue, 2019; Sustein, 2017). Conversely, research also highlights pages without a political agenda that mobilize against racism (Al Khan, 2016; Zerback and Fawzi, 2017).

Criticism of social media increasingly targets user-generated content for its pervasive discrimination, intolerance, and prejudice, as well as the platforms themselves, particularly Facebook and Instagram. These networks are accused of fostering racism by enabling and shaping discourse strategies of identification and persuasion that integrate racist narratives (Nielsen, 2002; Murthy and Sharma, 2019). In Western Balkan Countries far-right political leaders use discriminatory rhetoric in their speeches, which their supporters expand upon in comment sections (Ben-David and Matamoros-Fernández, 2016). Hate speech tends to operate in a top-down fashion, flowing from dominant groups to marginalized ones (Chua, 2009; Boromisza-Habashi, 2011; Calvert, 1997).

Several studies explore strategies involving Facebook profiles with fake identities designed to spread hate by mimicking and exaggerating radicalism. For instance, Costas et al. (2021) explain the role of thinking with diplomacy and using the freedom of speech for protecting the interest of the population and of the state. Harlow, 2015 presents example from newspapers to weigh the pros and cons of allowing anonymous user contributions. Some researchers argue that journalists are not adequately equipped to address the global challenge posed by hate speech in the media (Fish, 1997; Erjavec and Kovacic, 2012). While analyzing user comments is valuable, it is equally important to examine the ideological framing of hate speech that can sometimes underlie journalistic content itself (Tsesis, 2002).

Although YouTube is highly relevant to hate speech, research on the platform's comments remains limited. Only 11 studies, constituting about 9% of hate speech research in the communication field, have been identified. Murthy and Sharma (2019) highlight how hate speech, particularly racist discourse, generates significant interaction beyond the associated video content. Meanwhile, other researchers have noted YouTube's role in anti-racist campaigns. For example, one video uses humor to exaggerate racist attitudes to the point of ridicule, exposing them to criticism and highlighting their absurdity, while also serving as a mode of cultural assimilation (Archakisa et al., 2018).

Research also delves into the psychological and motivational profiles of hate speech producers and disseminators (Barlow and Awan, 2016; Erjavec and Kovačič, 2012). Two key motivations are commonly identified in individuals with an authoritarian personality. The first is the thrill derived from sparking debate, and the second is a self-perception as "guardians of justice" within society (Jacks and Adler 2015). Studies further explore the impact of hate speech on victims, highlighting the importance of socio-cultural identification. This identification shapes how hate speech is perceived, with its offensiveness varying depending on the cultural context (George

2014). Studies frequently analyze the most common terms, keywords, and patterns, along with variables coding for pejorative and negative language (Leets, 2001; Giglietto and Lee, 2017).

Common discursive strategies include exaggeration and contradiction (Chiluwa, 2018). In social media research, visual elements such as photos, GIFs, and memes accompanying comments are also considered. Multidimensional methodologies are especially effective for analyzing data from these platforms, though big data-based studies remain relatively scarce (Siegel, 2020; Slagle, 2009). Some studies use tools like RapidMiner Studio for data mining and analysis, providing deeper insights into the data collected from social networks.

Methodologies rooted in rhetoric are often applied to examine stereotypes and define various conceptions of “the others” (Meddaugh and Kay, 2009; Sugrue, 2019; Sustain 2017). This approach also encompasses the analysis of metaphors, whether humorous or satirical (Tontodimamma et al., 2021). A key challenge in this area lies in clearly differentiating between hatred and antagonism. Content producers are studied using in-depth interviews and digital ethnography, while victims are typically examined through interview-based methodologies. From a semiotic perspective, the responsibility for interpreting a text lies with the receiver, as their decoding process determines the meaning they extract and prioritize. Given the polysemic nature of messages, receivers may derive meanings that deviate from the sender’s intent, underscoring the importance of addressing audiences to clarify the intended message.

Research into hate speech and its mechanisms of mobilization can facilitate its mitigation. The literature highlights the presence of hate speech in domains such as politics (Boromisza-Habashi, 2011; Hysa et al., 2022), sports and advertising (Harlow, 2015), and fiction (Draper and Lotz, 2012). Some studies recommend best practices to encourage communication among diverse groups (Chua, 2009). Alternative discourses can also help counteract hate messages (Zerback and Fawzi, 2017). However, while research can guide interventions, the responsibility for addressing hate speech ultimately falls on authorities, journalists, cultural and social actors, and proactive citizens.

Language evolves quickly, especially among young people on social networks, requiring ongoing research on hate speech datasets. While platforms use manual and automated methods to remove hate speech content, users continuously find ways to bypass these restrictions, such as posting hate speech as images. Although image-to-text conversion could address some issues, it has limitations, and existing detection systems still face challenges. Additionally, changing language structure, like using abbreviations or mixing languages, further complicates detection, for instance, by writing part of a sentence in one language and the rest in another, or using phonetic spellings in a different language.

5. Conclusions and analysis of hate speech from 108 participants

Hate speech extends beyond personal insults, becoming a tool for those in power to maintain dominance and obstruct equality. This demands a reevaluation of legal protections for hate speech, advocating for civil penalties to help marginalized groups achieve equality. Social media platforms, especially Twitter and Facebook, have

transformed how racism and hate speech are expressed, often lacking emotional restraint due to anonymity and platform characteristics. These platforms facilitate the spread of racism and intolerance, with fake identities and discriminatory rhetoric becoming prevalent, especially in political discourse.

Research on social media hate speech often focuses on user-generated content and the strategies used to spread harmful narratives, with platforms like Facebook and YouTube being key areas of concern. Studies show that hate speech often intersects with complex identities such as religion, ethnicity, and politics, and can be exacerbated by the normalization of post-racist ideologies. Despite challenges, social media can also serve as a space for anti-racist campaigns. Research highlights the psychological motivations of hate speech producers, often linked to authoritarian personality traits, and emphasizes the socio-cultural context in which hate speech occurs.

Methodological approaches to studying hate speech include content analysis, critical discourse analysis, and the use of data mining tools. Researchers examine common discursive strategies, visual elements, and the rhetoric used in hate speech, with a focus on how hate is framed in different cultural contexts. Interventions against hate speech are seen as necessary, with responsibility lying not only with researchers but also with authorities, journalists, and proactive citizens. Overall, understanding and mitigating hate speech requires a multidimensional approach, considering both the evolution of language and the sociopolitical factors at play.

5.1. Main characteristics of hate speech (In the question—How will you describe the hate speech?)

Here is a semantic categorization of the given words from 108 participants on describing the meaning of hate speech in Albanian and translated English version, grouped into categories based on their meanings as in **Table 2**:

Table 2. Categories and terminology related to hate speech and misinformation in academic contexts in Albanian and in English.

Category	Terms
Hate and Negative Expression (Albanian)	Gjuha e urrejtjes, shprehje që synojnë të përbuzin, nënçmojnë, linçojne, fyejnë, urrejtje, keq, keq e demshme, e tmershme, shumë te keqe, gjuha e urrejtjes është mënyrë e gabuar e shprehjes, gjuha e urrejtjes është ofendim dhe dhune psikologjike, fjale të cilat përçojnë urrejtje, gjuha e urrejtjes është përdorimi i fjalëve që nxisin, përhapin ose promovojnë urrejtje
Hate and Negative Expression (English)€	Hate speech, is expressions that aim to ridicule, demean, lynch, insult, hatred, bad, harmful, dreadful, or very bad, hate speech is an incorrect way of expression, hate speech is offense and psychological violence, words that convey hatred, hate speech is the use of words that incite, spread, or promote hatred
Discrimination (A)	Diskriminim, fjalë fyese, diskriminimi, paragjykim, urrejtje politike, sharje, fjalë raciste, përçmim, nënçmim, poshtërim, përqeshje, fjalë që kanë tendencë të bëjnë keq, gjuha e urrejtjes është përdorimi i fjalëve, simboleve dhe ideve që nxisin urrejtje ndaj një grupi të caktuar.
Discrimination (E)	Discrimination, offensive words, discrimination, prejudice, political hatred, insults, racist words, contempt, demeaning, humiliation, mockery, words that tend to harm, hate speech is the use of words, symbols, and ideas that incite hatred towards a specific group.
Violence and Aggression (A)	Dhunë, dhune psikologjike, kërcënime, gjuha e urrejtjes që nxit dhunë, përdorimi i fjalëve që nxisin dhunë ndaj grupeve të caktuara, gjuha që nxit frikë dhe dhunë, përdorimi i fjalëve që nxjerrin tensione dhe mund të shkaktojnë dëme shoqërore.
Violence and Aggression (E)	Violence, psychological violence, threats, hate speech that incites violence, the use of words that incite violence against specific groups, language that incites fear and violence, the use of words that create tensions and may cause social harm

Table 2. (Continued).

Category	Terms
Offensive and Insulting Language (A)	Sharje, fjalë të rënda, ofendime, fjalë degjeneruese, përdorim i fjalëve për të poshtëruar, shprehje që fyen ose shan, fjalë që nënçmojnë njerëzit për tiparet e tyre
Offensive and Insulting Language (E)	Insults, harsh words, offenses, degrading words, the use of words to humiliate, expressions that insult or curse, words that demean people for their characteristics
Social Divisions and Prejudices (A)	Përçarje mes individëve, përçarje shoqërore, paragjykim, ndarje në grupe, përdorim i fjalëve që ndajnë njerëzit bazuar në fe, racë, origjinë, orientim seksual, etni
Social Divisions and Prejudices (E)	Division among individuals, social division, prejudice, group separation, the use of words that divide people based on religion, race, origin, sexual orientation, ethnicity
Emotional and Psychological Impact (A)	Gjuha që ul moralin, presion psikologjik, sjellje që ndikon negativisht, gjuha që shkakton ndjenja negative, gjuha që shpreh urrejtje dhe poshtërim për ndjenja dhe identitete të ndryshme.
Emotional and Psychological Impact(E)	The language that lowers morale, psychological pressure, behavior that negatively affects, language that causes negative feelings, language that expresses hatred and humiliation towards different feelings and identities
Cultural Impact (A)	Kulturë, ndikim kulturor, gjuha e urrejtjes si fenomen denigrues dhe antinjerëzor, shprehje kulturore që kontribuojnë në pasiguri dhe ndarje shoqërore.
Cultural Impact (E)	Culture, cultural influence, hate speech as a denigrating and inhuman phenomenon, cultural expressions that contribute to insecurity and social division.
Mentality and Behavior (A)	Mentalitet i ulët, sjellje jo e drejtë, përdorim i gjuhës që tregon inat, urrejtje politike, fjalë që dërgojnë njerëzit në vendime të gabuara
Mentality and Behavior (E)	Low mentality, unfair behavior, use of language that shows resentment, political hate, words that lead people to make wrong decisions
Harmful Communication (A)	Komunikimi që është i dëmshëm për individët dhe grupe shoqërore, përdorimi i fjalëve që nxisin tensione, nënçmim dhe dhunë shoqërore, fjalë që nxisin armiqësi dhe keqkuptime
Harmful Communication (E)	Communication that is harmful to individuals and social groups, use of words that incite tension, contempt, and social violence, words that provoke hostility and misunderstandings
Intolerance (A)	Intolerancë, gjuha që reflekton ndarje të thella, gjuha e urrejtjes që shpreh mospranim dhe përbuzje për ndryshime kulturore, sociale, apo seksuale
Intolerance (E)	Intolerance, language that reflects deep divisions, hate speech that expresses rejection and contempt for cultural, social, or sexual differences

The words into key areas related to hate speech, discrimination, violence, cultural impact, and the emotional and psychological effects. This research provides the first comprehensive exploration of hate speech in higher education in Kosovo, combining empirical data and interdisciplinary insights to inform policies and promote inclusivity in academic environments

The categorization of words related to hate speech reveals several key themes:

Hate and negative expression (30%): The majority of the words fall under this category, focusing on terms related to hate speech, psychological violence, and negative expressions.

Discrimination (20%): A significant portion of words emphasize discriminatory actions or words that promote prejudice and exclusion based on race, gender, or other social factors.

Violence and aggression (15%): Words indicating aggression, threats, and psychological harm show how hate speech can escalate into violent outcomes.

Offensive and insulting language (10%): Insulting words and those meant to degrade others make up a notable but smaller proportion.

Social divisions and prejudices (10%): Words reflecting societal divisions and the impact of hate speech in dividing communities.

Emotional and psychological impact (5%): A smaller but critical category focusing on the harmful emotional effects of hate speech.

Cultural impact (5%): Hate speech’s influence on culture and social cohesion.

Mentality and behavior (5%): Words related to negative mentality and behaviors fostered by hate speech.

Harmful communication (5%): This category shows how communication fueled by hate can harm social dynamics.

Intolerance (5%): Focuses on words reflecting societal intolerance.

Results of quantitative research (108 Questionnaires) are in **Table 3**.

Table 3. Key themes, frequency, and examples of hate speech experiences in academic setting.

Theme	Frequency (%)	Example
Witnessing hate speech	81%	“I see offensive memes targeting minority groups in WhatsApp study groups.”
Uncertainty about reporting	72%	“I don’t know who to contact or how to report hate speech incidents at my university.”
Impact on mental health	64%	“Constant offensive comments make me feel unwelcome and unable to focus on my studies.”
Need for institutional action	89%	“Universities should organize workshops and make reporting channels more accessible to students.”

Key findings: Perception of hate speech, 67% of respondents agreed that they often witness hate speech on social media platforms linked to university discussions.

Common themes identified: Freedom of expression vs. hate speech: Many participants struggled to distinguish between hate speech and legitimate critique.

Normalization: Hate speech is frequently disguised as humor or casual remarks.

Targets of hate speech: Ethnic minorities, women, and individuals expressing alternative political views were most commonly targeted.

Impact of hate speech. For example, a female respondent from an ethnic minority reported withdrawing from social activities due to persistent derogatory comments.

Respondents frequently mentioned feelings of anxiety, stress, and exclusion, which negatively impacted their academic performance. Institutional awareness and action, 72% of respondents were unaware of any institutional mechanisms to report hate speech. For example, one respondent stated, “I have seen posters about anti-discrimination policies but don’t know how to file a report.”

Table 4. Analysis of hate speech frequency, targets, emotional impact, and policy perception in academic settings.

Category	Metric	Result (%)
Frequency of hate speech	Weekly exposure	75% encounter it 2–3 times/week
Primary targets	Ethnic minorities	60%
	Women	45%
Emotional impact	Anxiety	70%
	Stress	60%
	Avoidance of discussions	55%
Perception of policies	Policies are effective	25%

Results of qualitative analysis (20 Participants). Key metrics are assessed in **Table 4**.

Effectiveness of institutional policies. Results: Only 25% believed their university had effective anti-hate speech measures. For example, one participant mentioned, “Even when incidents are reported, they rarely result in any action.”

Comparison of quantitative and qualitative results:

- **Overlap in findings:** Both data sets confirm the normalization of hate speech and highlight ethnic minorities and women as primary targets.
- **Institutional shortcomings:** Both groups emphasized the lack of awareness or effectiveness of existing anti-hate speech policies.
- **Depth of understanding:** Qualitative data provided richer narratives, revealing emotional impacts and personal stories. Quantitative data offered measurable insights into frequency, demographic breakdowns, and specific impacts.

5.2. Summary of results and recommendations

5.2.1. Quantitative data

The quantitative analysis, based on 108 questionnaires, revealed that many students struggle to distinguish between hate speech and freedom of expression, with hate speech often normalized as humor or casual remarks. Ethnic minorities, women, and individuals expressing alternative political views were identified as primary targets. For instance, one female participant from an ethnic minority shared that she refrains from participating in discussions due to derogatory comments. Furthermore, 72% of respondents were unaware of institutional mechanisms to report hate speech, emphasizing the need for improved awareness and accessibility.

5.2.2. Qualitative data

The qualitative analysis of 20 participants highlighted the frequent encounter of hate speech, with 75% reporting exposure 2–3 times weekly, primarily online. Ethnic minorities and women were most often targeted, accounting for 60% of minorities and 45% of women respectively. Emotional impacts were significant, with anxiety (70%) and stress (60%) commonly reported. Reduced participation in discussions (55%) also emerged as a key issue, as students often avoided group interactions due to offensive remarks. Only 25% of participants believed their universities had effective anti-hate speech measures, further underscoring institutional gaps.

5.3. Recommendations for addressing hate speech in academic settings

To effectively combat hate speech and uphold democratic values, universities in Kosovo must implement a multi-faceted approach:

Universities should develop comprehensive policies clearly defining and addressing hate speech, ensuring alignment with democratic principles and international human rights standards. These policies should address specific local challenges, including ethnic and gender discrimination while providing actionable guidelines for prevention and resolution. Incorporating workshops and seminars on digital literacy, free speech, and hate speech into the curriculum is essential. Tailored educational programs focusing on local issues, such as ethnic reconciliation and gender equity, should be prioritized in Kosovo. Furthermore, awareness campaigns about human rights laws and institutional regulations can empower students and staff to recognize and address hate speech effectively. Universities in Kosovo should

establish robust support systems, including counseling services and anonymous reporting mechanisms for victims of hate speech. A dedicated online platform where students can report incidents confidentially and receive guidance can ensure safety and accountability. Collaboration with different authorities and engaging in partnerships with civil society organizations, government bodies, and international organizations can amplify efforts to counter hate speech. Collaborative initiatives, such as intercultural student exchanges and joint projects, can foster understanding, bridge cultural divides, and promote democratic engagement among diverse groups. By implementing these measures, universities can create a safer, more inclusive environment that encourages mutual respect, open dialogue, and active participation in democratic processes.

To address these findings, universities should introduce awareness campaigns distinguishing hate speech from freedom of expression and fostering inclusivity. Simplified reporting mechanisms and clear anti-hate speech policies are essential. Digital moderation, in partnership with social media platforms, can help monitor hate speech in local languages. Additionally, mental health support services for affected students and diversity-focused events could create a more supportive environment.

Recommendations for future research and how to support the community against hate speech

Future studies should explore the long-term effects of hate speech on students' academic and social well-being through longitudinal studies. Expanding the demographic scope to include socio-economic and sexual orientation factors would provide deeper insights. Comparative research with HEIs in neighboring countries could identify regional patterns and effective strategies. Investigating the role of algorithms in enabling hate speech on digital platforms and evaluating the effectiveness of institutional interventions over time would further contribute to mitigating this issue. Educational programs should raise awareness about the emotional and psychological harm caused by hate speech. Stricter regulations and penalties are needed to address hate speech in public forums, social media, and official communication. Foster social and cultural initiatives that encourage inclusivity and respect for diversity, thereby mitigating the effects of discriminatory language. Provide comprehensive psychological and legal support to individuals affected by hate speech, ensuring they have access to the necessary resources for recovery.

Incorporating hate speech education into communication courses is essential, focusing on media literacy, critical thinking, and ethical communication, supported by historical examples such as "The Diary of Anne Frank", or "Nathan the Wise", and nationalist propaganda. Promoting empathy and diversity through storytelling, literature studies, and interactive workshops can foster cultural understanding and mutual respect among students. In collaboration with writers, filmmakers, and influencers, public awareness campaigns can leverage literature and media to create compelling narratives that challenge stereotypes. Additionally, interdisciplinary research and accessible case studies on hate speech prevention can enhance understanding and provide practical tools for education and community engagement. Kosovo's society faces significant challenges in becoming fully inclusive and free from prejudice, hatred, and discrimination, despite its progressive legislation. Media

often perpetuates stereotypes, sexism, and xenophobic narratives through unprofessional reporting, such as using sensationalist headlines and biased coverage. Additionally, marginalized groups like the Roma, Egyptian, and Ashkali communities receive minimal media representation, reflecting systemic racism and discrimination rather than a lack of targeted hate. Kosovo has a strong legal framework against hate speech and discrimination, enshrined in its Constitution, Criminal Code, and laws such as the Law on Protection from Discrimination and media-specific regulations. Research highlights that hate narratives and disinformation are prevalent in media, online platforms, and even political discourse, disproportionately targeting women, the LGBTQ+ community, and minorities. To address this, stricter enforcement of media ethics by regulatory bodies and a commitment from political parties to avoid hateful rhetoric is essential.

Author contributions: Conceptualization, KC and BC; methodology, KC; software, A.N; validation, KC, SH and BC; formal analysis, KC; investigation, AN; resources, S.H; data curation, KC; writing—original draft preparation, KC; writing—review and editing, KC; visualization, AN; supervision, KC; project administration, KC; funding acquisition, AN. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

Notes

- ¹ UNESCO. (2018). Journalism, fake news & disinformation: Handbook for journalism education and training. <https://unesdoc.unesco.org/ark:/48223/pf0000265552>
- ² Ibid.
- ³ Ibid
- ⁴ Merriam-Webster Dictionary: <https://www.merriam-webster.com/dictionary/clickbait>
- ⁵ Merriam-Webster Dictionary: <https://www.merriam-webster.com/dictionary/bait>
- ⁶ Constitution of Kosovo Art. 40 <https://www.kpm-ks.org/assets/cms/uploads/files/Legjislacioni/1343117104.2078.pdf>

References

- Altman, A., Maitra, I.; & McGowan, M. K. (2012). Freedom of Expression and Human Rights Law: The Case of Holocaust Denial, Speech and Harm. Oxford University Press. pp. 24–49
- Archakisa, A., Lampropouloub, S., & Tsakonac, V. (2018). I'm not racist but I expect linguistic assimilation: The concealing power of humor in an anti-racist campaign. *Discourse, Context and Media*, 23, 53–61. <https://doi.org/10.1016/j.dcm.2017.03.005>
- Assimakopoulos S., Baider F. and Millar S. (2017). Hate Speech in the European Union: A Discourse- Analytic Perspective. Springer Briefs.
- Assimakopoulos, S., Baider, F. H., & Millar, S. (2017). Online Hate Speech in the European Union. In: SpringerBriefs in Linguistics. Springer International Publishing.
- Awan, I. (2016). Islamophobia on social media: A qualitative analysis of the facebook's walls of hate. *International Journal of Cyber Criminology*, 10(1), 1–20. <https://doi.org/10.5281/zenodo.58517>
- Barlow, C., & Awan, I. (2016). "You Need to Be Sorted Out With a Knife": The Attempted Online Silencing of Women and People of Muslim Faith Within Academia. *Social Media + Society*, 2(4). <https://doi.org/10.1177/2056305116678896>
- Beha, B. (2022). Gjuha e urrejtjes dhe dhuna në studio televizive | Asociacioni i Gazetarëve të Kosovës. Hate speech and violence in TV studios | Association of Journalists of Kosovo (Article Title in Albanian) AGK (agk-ks.org) Available online: <https://agk-ks.org/lajmet/gjuha-e-urrejtjes-dhe-dhuna-ne-studio-televizive/> (accessed on 14 January 2024).

- Ben-David, A., & Matamoros-Fernández, A. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, 10, 1167–1193.
- Boeckmann, R. J., & Turpin-Petrosino, C. (2002). Understanding the Harm of Hate Crime. *Journal of Social Issues*, 58(2), 207–225. Portico. <https://doi.org/10.1111/1540-4560.00257>
- Boromisza-Habashi, D. (2011). Dismantling the antiracist “hate speech” agenda in Hungary: An ethno-rhetorical analysis. *Journal of Language, Discourse and Communication Studies*, 31(1), 1–19. <https://doi.org/10.1515/text.2011.00>
- Burnap, P., & Williams, M. L. (2015). Cyber Hate Speech on Twitter: An Application of Machine Classification and Statistical Modeling for Policy and Decision Making. *Policy & Internet*, 7(2), 223–242. Portico. <https://doi.org/10.1002/poi3.85>
- Calvert, C. (1997). Hate Speech and Its Harms: A Communication Theory Perspective. *Journal of Communication*, 47(1), 4–19. <https://doi.org/10.1111/j.1460-2466.1997.tb02690.x>
- Cammaerts, B. (2009). Radical pluralism and free speech in online public spaces. *International Journal of Cultural Studies*, 12(6), 555–575. <https://doi.org/10.1177/1367877909342479>
- Chiluwa, I. (2018). A nation divided against itself: Biafra and the conflicting online protest discourses. *Discourse and Communication*, 12(4), 357–381. <https://doi.org/10.1177/1750481318757778>
- Chua, C. E. H. (2009). Why Do Virtual Communities Regulate Speech? *Communication Monographs*, 76(2), 234–261. <https://doi.org/10.1080/03637750902828420>
- Cisneros, D. J., & Nakayama, T. K. (2015). New media, old racism: Twitter, Miss America, and cultural logics of race. *Journal of International and Intercultural Communication*, 8(2), 108–127. <https://doi.org/10.1080/17513057.2015.1025328>
- Constantinou, C. M., Dittmer, J., Kuus, M., et al. (2021). Thinking with Diplomacy: Within and Beyond Practice Theory. *International Political Sociology*, 15(4), 559–587. <https://doi.org/10.1093/ips/olab028>
- Constitution of the Republic of Kosovo (2008): Available online: [https://hudoc.echr.coe.int/fre#%7B%22itemid%22:\[%22001-158752%22\]%7D](https://hudoc.echr.coe.int/fre#%7B%22itemid%22:[%22001-158752%22]%7D) (accessed on 14 January 2025). Kushtetuta e Republikës së Kosovës, 2008 (Title in Albanian. Republic of Kosovo).
- Council of Europe. (2014). Bookmarks—A manual for combating hate speech online through human rights education. ISBN 978-92-871-7840-4. (accessed on 14. January 2025) <https://book.coe.int/en/human-rights-and-democracy/5888-bookmarks-a-manual-for-combating-hate-speech-online-through-human-rights-education.html>.
- Council of Europe. (2020). Guidelines on hate speech. Available online: <https://www.coe.int>. (accessed on 14. January 2025). <https://www.coe.int/en/web/combating-hate-speech/council-of-europe-on-hate-speech>
- Cowan, G., & Hodge, C. (1996). Judgments of Hate Speech: The Effects of Target Group, Publicness, and Behavioral Responses of the Target. *Journal of Applied Social Psychology*, 26(4), 355–374. Portico. <https://doi.org/10.1111/j.1559-1816.1996.tb01854.x>
- Domingo, D., Quandt, T., Heinonen, A., et al. (2008). PARTICIPATORY JOURNALISM PRACTICES IN THE MEDIA AND BEYOND. *Journalism Practice*, 2(3), 326–342. <https://doi.org/10.1080/17512780802281065>
- Donley T. S. (1974). Dialectics of freedom. *Patterns of Prejudice*, 8(3), 12–28. <https://doi.org/10.1080/0031322x.1974.9969188>
- Draper, J., & Lotz, A. D. (2012). “Working Through” as Ideological Intervention. *Television & New Media*, 13(6), 520–534. <https://doi.org/10.1177/1527476411431169>
- EC. (2018). Third Evaluation of the Code of Conduit on Countering Illegal Online Hate Speech. Available online: http://europa.eu/rapid/press-release_MEMO-18-262_en.htm (accessed on 4 Juen 2024).
- Elfferding, W. (1983). On the racist discourse in the Heidelberger Manifest [Das Argument 138]. (14. January 2025) <http://www.neu.inkrit.de/mediadaten/archivargument/DA138/DA138.pdf>
- Erjavec, K., & Kovačič, M. P. (2012). “You Don’t Understand, This is a New War!” Analysis of Hate Speech in News Web Sites’ Comments. *Mass Communication and Society*, 15(6), 899–920. <https://doi.org/10.1080/15205436.2011.619679>
- Eurobarometer. (2016). Media Pluralism and Democracy. Available online: <https://ec.europa.eu/digital-single-market/en/news/media-pluralism-and-democracy-special-eurobarometer-452>. (accessed on 4 June 2024).
- European Commission. (2023). Addressing Hate Speech in the Western Balkans. Available online: https://ec.europa.eu/home-affairs.ec.europa.eu/system/files/2023-10/ran_wb_taking_stock_of_approaches_to_counter_hate_speech_092023_en.pdf (accessed on 4 June 2024).
- Fairclough, N. (2001). *Language and power*. Pearson Education.
- Farkas, J., Schou, J., & Neumayer, C. (2018). Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*, 20(5), 1850–1867. <https://doi.org/10.1177/1461444817707759>

- Fish, S. (1997). Boutique Multiculturalism, or Why Liberals Are Incapable of Thinking about Hate Speech. *Critical Inquiry*, 23(2), 378–395. <https://doi.org/10.1086/448833>
- Fortuna, P., & Nunes, S. (2018). A Survey on Automatic Detection of Hate Speech in Text. *ACM Computing Surveys*, 51(4), 1–30. <https://doi.org/10.1145/3232676>
- Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). Countering Online Hate Speech. UNESCO Series on Internet Freedom.
- Gallacher, J. D., & Bright, J. (2021). Hate Contagion: Measuring the spread and trajectory of hate on social media. <https://doi.org/10.31234/osf.io/b9qhd>
- Gelber, K. (2002). *Speaking Back. Discourse Approaches to Politics, Society and Culture*. Portico. <https://doi.org/10.1075/dapsac.1>
- Giglietto, F., & Lee, Y. (2017). A Hashtag Worth a Thousand Words: Discursive Strategies Around #JeNeSuisPasCharlie After the 2015 Charlie Hebdo Shooting. *Social Media + Society*, 3(1). <https://doi.org/10.1177/2056305116686992>
- Gracia-Calandín, J., & Suárez-Montoya, L. (2023). The eradication of hate speech on social media: a systematic review. *Journal of Information, Communication and Ethics in Society*, 21(4), 406–421. <https://doi.org/10.1108/JICES-11-2022-0098>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. Portico. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Harlow, S. (2015). Story-Chatterers Stirring up Hate: Racist Discourse in Reader Comments on U.S. Newspaper Websites. *Howard Journal of Communications*, 26(1), 21–42. <https://doi.org/10.1080/10646175.2014.984795>
- Haxholli, V. (2022). Monitoring Report on Hate Speech in Kosoco. Available online: https://www.reportingdiversity.org/wp-content/uploads/2022/07/MRHS_Kosovo.pdf (accessed on 4 June 2024).
- Herz, M., & Molnár, P. (2012). *The content and context of hate speech: Rethinking regulation and responses*. Cambridge University Press.
- Herz, M., & Peter M. (2012). *The content and context of hate speech*. Cambridge University Press.
- Hrvatín, S. B., Petković, B., & Hodžić, S. (2021). Hate narratives in the Western Balkans and Turkey. Available online: <https://seenpm.org/wp-content/uploads/2021/03/Regional-overview-Resilience-research-2.pdf> (accessed on 4 June 2024).
- Hysa, D., & K. Lani, K. (2022). Hate Language Monitoring Report in Albania (Albanian). Instituti Shqiptar i Medias.
- Jacks, W., & Adler, J. R. (2015). A proposed typology of online hate crime. *Psychology Unbound: Open Access Journal of Forensic Psychology*, 7, 64–89.
- John B., & George, C. (2014). Journalism and the politics of hate: Charting ethical responses to religious intolerance. *Journal of Mass Media Ethics*, 29(2), 74–90. <https://doi.org/10.1080/0890052.3.2014.893771>
- Kalven, H. Jr. (1967). Report on the University's Role in Political and Social Action. Available online: <https://www.goacta.org/kalven-report/> (accessed on 14 January 2024).
- Kosovo 2022 Human Rights Report. U.S. Department of State. Available online: https://www.state.gov/wp-content/uploads/2023/02/415610_KOSOVO-2022-HUMAN-RIGHTS-REPORT.pdf (accessed on 14 January 2024).
- Kosovo Constitutions. Available online: https://mapl.rks-gov.net/wp-content/uploads/2017/10/1.CONSTITUTION_OF_THE_REPUBLIC_OF_KOSOVO.pdf (accessed on 14 January 2024).
- Leets, L. (2001). Explaining perceptions of racist speech. *Communication Research*, 28(5), 676–706. <https://doi.org/10.1177/009365001028005005>
- Ligji civil kundër shpifjes dhe fyerjes Nr. 02 /L_65. Civil Law Against Defamation and Insult No. 02/L-65. Title in Albanian. Available online: https://agk-ks.org/site/assets/files/2611/ligji_nr__02_l-65_ligji_civil_kunder_shpifjes_dhe_fyerjes_1.pdf (accessed on 14 January 2024).
- Ligji për mbrojtjen nga diskriminimi, Nr. 10 221, dt. 04.02. 2010, Law on Protection from Discrimination, No.10 221. Title in Albanian (accessed on 14. January 2025). http://www.uart.edu.al/media/150329-Ligji%2010221_%202010%20per%20mbrojtjen%20nga%20diskriminimi.pdf
- Londo, I. (2021). Hate narratives and disinformation in online media in Albania. SEENPM, Peace Institute, and Albanian Media Institute.
- McNamee, L. G., Peterson, B. L., & Peña, J. (2010). A Call to Educate, Participate, Invoke and Indict: Understanding the Communication of Online Hate Groups. *Communication Monographs*, 77(2), 257–280. <https://doi.org/10.1080/03637751003758227>
- Meddaugh, P. M., & Kay, J. (2009). Hate Speech or “Reasonable Racism?” The Other in Stormfront. *Journal of Mass Media Ethics*, 24(4), 251–268. <https://doi.org/10.1080/08900520903320936>

- Murthy, D., & Sharma, S. (2019). Visualizing YouTube's comment space: Online hostility as a networked phenomenon. *New Media and Society*, 21(1), 191–213. <https://doi.org/10.1177/1461444818792393>
- Nguyen, T. (2023). Merging public health and automated approaches to address online hate speech. *AI and Ethics*, 4(2), 441–450. <https://doi.org/10.1007/s43681-023-00281-w>
- Nielsen, L. B. (2002). Subtle, pervasive, harmful: Racist and sexist remarks in public as hate speech. *Journal of Social Issues*, 58(2), 265–280. <https://doi.org/10.1111/1540-4560.00260>
- Parekh, B. (2012). Hate Speech: Is There a Case for Banning? *Public Policy Research*, 19(2), 117–122.
- RESILIENCE. (2021). Hate narratives in the Western Balkans and Turkey. Peace Institute and SEENPM.
- Siegel, A. A. (2020). Online hate speech. In: Tucker, J., & Persily, N. (editors). *Social Media and Democracy: The State of the Field*. Cambridge: Cambridge University Press.
- Slagle, M. (2009). An Ethical Exploration of Free Expression and the Problem of Hate Speech. *Journal of Mass Media Ethics*, 24(4), 238–250. <https://doi.org/10.1080/08900520903320894>
- Sobkowicz, P., & Sobkowicz, A. (2010). Dynamics of hate based Internet user networks. *The European Physical Journal B*, 73(4), 633–643. <https://doi.org/10.1140/epjb/e2010-00039-0>
- Sokol, A. (2021). Hate narratives in the media and user-generated content. SEENPM, Peace Institute and Foundation Mediacentar Sarajevo.
- Sugrue, E. P. (2018). A 'bad fit' for 'our' kids: politics, identity, race and power in parental discourse on educational programming & child well-being. *Critical Discourse Studies*, 16(2), 222–236. <https://doi.org/10.1080/17405904.2018.1538891>
- Sustein C. (2017). *Republic: Divided Democracies in the Age of Social Media*. Princeton University Press.
- Thomas H. (2022). Hate speech online: the government as regulator and as a speaker. *Journal of Media Law*, 14(1), 139–158, <https://doi.org/10.1080/17577632.2022.2085014>
- Tontodimamma, A., Nissi, E., Sarra, A., et al. (2020). Thirty years of research into hate speech: topics of interest and their evolution. *Scientometrics*, 126(1), 157–179. <https://doi.org/10.1007/s11192-020-03737-6>
- Tsesis, A. (2002). *Destructive messages: How hate speech paves the way for harmful social movements*. New York University Press.
- United Nations. (2021). *Universal Declaration of Human Rights*. United Nations.
- Vargo, C. J., & Hopp, T. (2020). Fear, Anger, and Political Advertisement Engagement: A Computational Case Study of Russian-Linked Facebook and Instagram Content. *Journalism & Mass Communication Quarterly*, 97(3), 743–761. <https://doi.org/10.1177/1077699020911884>
- Waldron, J. (2012). *The Harm in Hate Speech*. Harvard University Press.
- Wodak, R. (2002). Friend or foe—The defamation or legitimate and necessary criticism? Reflections on recent political discourse in Austria. *Language and Communication*, 22(4), 495–517. [https://doi.org/10.1016/S0271-5309\(02\)00022-8](https://doi.org/10.1016/S0271-5309(02)00022-8)
- Zerback, T., & Fawzi, N. (2017). Can online exemplars trigger a spiral of silence? Examining the effects of exemplar opinions on perceptions of public opinion and speaking out. *New Media and Society*, 19(7), 1034–1051. <https://doi.org/10.1177/146144481562594>