



OPEN ACCESS

EDITED BY

Graciela Padilla-Castillo,
Complutense University of Madrid, Spain

REVIEWED BY

David Del Pino Díaz,
Universidad Antonio de Nebrija, Madrid, Spain
Xabier Martínez-Rolán,
University of Vigo, Spain

*CORRESPONDENCE

Dilara Asardag
✉ dilara.asardag@tuni.fi

RECEIVED 03 June 2025

ACCEPTED 30 June 2025

PUBLISHED 06 August 2025

CORRECTED 07 August 2025

CITATION

Asardag D (2025) Feminist exploratory interpretive study of the content policy changes of Meta and the corresponding news coverage.

Front. Commun. 10:1640230.

doi: 10.3389/fcomm.2025.1640230

COPYRIGHT

© 2025 Asardag. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Feminist exploratory interpretive study of the content policy changes of Meta and the corresponding news coverage

Dilara Asardag*

TARU Research Group, Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland

Meta CEO Mark Zuckerberg recently made an announcement that there will be major changes to his company's content moderation policies. As part of this move, Meta is eliminating its third-party fact-checking program, and replacing it with a community based program called Community Notes as well as making changes to the "Hateful Conduct" policy document of Meta. Under conditions of rising anti-gender backlash and the emergence of post-truth societies transnationally with distorted notions of truth, these changes are deeply concerning for social media researchers, media studies researchers, gender studies researchers and policy scholars as well, as not only CEOs like Mark Zuckerberg but also, politicians like vice-president JD Vance, can be seen as launching a brutal ideological assault on Europe. Taking into account these political dynamics, in this article, I investigated (i) how are the proposed policy changes of Meta discursively constructed in major European online news outlets; (ii) what are the actual policy changes, taking into account women and LGBTIQ+ identities? (iii) what should/ could be the European level response to this situation considering policy and civil society level discussions? I carried out an exploratory reflexive thematic-discursive analysis of 14 articles from major European news sources published online between 7 January to 7 March 2025 related to the announcements of Mark Zuckerberg as well as responses to these announcements and Meta's official "Hateful Conduct" policy document and the official statement. Adopting an intersectional feminist standpoint theory, strong objectivity and group oppression theory, the exploratory reflexive thematic-discursive analysis was carried out through the lens of hateful speech. The emerging themes of this research are (i) retreat on human rights, LGBTIQ+ rights and freedom of expression, (ii) frightening developments for the rise of misinformation/ disinformation, (iii) Zuckerberg as "re-setting relations with Europe, ignoring Digital Services Act of EU and the need to debunk 'institutionalization of censorship claims.'"

KEYWORDS

exploratory approach, feminist standpoint theory, hate speech policies, news media, Meta, social media, content moderation, LGBTIQ+

1 Introduction

In a Facebook video shared on 7 January 2025, Meta CEO Mark Zuckerberg has recently announced that there will be major changes to his company's content moderation policies (Zuckerberg, 2025). After years of having our content moderation work focused primarily on removing content, he said, there will be an effort to "get back to our roots and focus on

reducing mistakes, simplifying our policies and restoring free expression to give people voice” (Zuckerberg, 2025). As part of this move, Meta is eliminating its third-party fact-checking program, and replacing it with a community based program called Community Notes. Similar to Elon Musk’s social platform X, Meta’s Community Notes will rely on users to submit notes or corrections to posts that are potentially misleading or lack context in Instagram, Facebook and Threads. Moreover, there are changes to the “Hateful Conduct” policy document of Meta (Zuckerberg, 2025). Similar to the so-called “free speech absolutist” Elon Musk, according to Zuckerberg, the main impetus for change is “blocking censorship,” “the desire to bolster free expression” and retreating from efforts to diminish the spread of what he calls hate speech (Zuckerberg, 2025).

Moreover, according to Zuckerberg, Europe has an ever increasing number of laws, institutionalizing censorship and making it difficult to build anything innovative there (Zuckerberg, 2025). Zuckerberg’s statement was timed with the release of a post by Meta Chief Global Affairs Officer Joel Kaplan, where he said “We are getting rid of a number of restrictions on topics like immigration, gender identity and gender that are the subject of frequent political discourse and debate” (Hendrix, 2025). Moreover, Kaplan said on FOX and Friends, “We got a lot of pressure around COVID in particular to take down more content, even things like humour and satire about the pandemic and vaccines...but there is a real opportunity here with President Trump coming into office with his commitment to free expression for us to get back to those values and really provide the space for people to have the kind of discourse and debate they want to have” (Hendrix, 2025). These announcements were timed with Mark Zuckerberg’s interview with the conservative Trump supporter podcaster Joe Rogan where Zuckerberg stated “companies now require more masculine energy and less diversity policy” (Piquard, 2025).

However, under conditions of rising anti-gender backlash (see. US and Turkey, for example; Abaday, 2024) and the emergence of post-truth societies transnationally with distorted notions of truth (Asardag and Donders, 2021), these changes are deeply concerning for social media researchers, media studies researchers, gender studies researchers and policy scholars as well. Not only CEOs like Mark Zuckerberg but also, politicians like vice-president JD Vance, can be seen as launching a brutal ideological assault on Europe, accusing its leaders of suppressing free speech, failing to halt illegal migration and running in fear from voters’ true beliefs (Wintour, 2025). As the concept of truth is being weaponized by the right wing politicians and CEOs, it is hypothesized that these implied policy moves of Meta are significantly geared towards right-wing pundits like president-elect Trump and other conservatives, as they are now allowed spreading more lies, hate speech, and conspiracy theories as facts, camouflaged as “free expression.”

Social media researchers, gender studies scholars, and relevant NGOs and advocacy networks (GLAAD, 2025) are deeply worried that with the change of the policies, Meta is directly endorsing women, LGBTIQ+ communities, immigrants and other marginalised groups to be the direct targets of hateful rhetoric, harassment and vitriol. Also to be noted, major technology company CEOs of Google, Microsoft, Amazon, Meta and Open AI have all donated 1 million dollars to Trump’s inaugural fund (Yang, 2025). As argued by Magalhães (2025), the sequence of events is well-documented. Initially, Zuckerberg’s efforts to implement stronger speech governance measures—aimed at appeasing policymakers in the so-called “post-truth” era following

Donald Trump’s first election in 2016—backfired. Republican accusations of censorship galvanised public opposition to Meta and other tech giants, leading to a backlash against platform regulation. In the wake of these reactions, many countries, following Europe’s example, significantly increased state control over digital platforms. Then, Trump won again (Magalhães, 2025). It is to President-elect Trump, his supporters, followers, and allies that Zuckerberg is saying sorry and pledging allegiance. And it’s not just Trump. Authoritarian leaders around the world, including Turkey and Russia are undoubtedly eager to harness an incredibly advanced global communication system for their own agendas (Magalhães, 2025).

While the European Union gets isolated in the international arena by the US and Russia in matters, for instance, related to the support for Ukraine, democratic values of EU overall are under attack under the current government of US. JD Vance was recently pointing out that the true threat to Europe stemmed not from external actors such as Russia and China, but Europe’s own internal retreat from some of its “most fundamental values,” he repeatedly questioned whether the US and Europe any longer had a shared agenda (Yang, 2025). While the existing self-regulation prospects of digital platforms has clearly been open for scrutiny and criticism by some scholars who argue that rather than being allowed to self-regulate (Pukallus and Arthur, 2024, p. 471), these companies are needed to be legally forced to systematically uphold their own community standards, it is hypothesized that the current changes proposed by Mark Zuckerberg points towards a direction that renders the current situation even worse. Although the notion of hate speech has been approached from legal and regulatory frameworks as well as hate speech studied as mundane everyday practice in media studies (Nikunen, 2023; Nikunen et al., 2025), this article, by using *the notion of hateful speech* that will be explained later in the article, will attempt to extend the analysis beyond the legal category of hate speech as well as hate speech as mundane every day practice in order to comprehend the potential discursive and material ramifications of *hateful speech* for women and LGBTIQ+.

In this article, therefore, considering these political dynamics as well, it will be attempted to be exploratively understood, (i) how are the proposed policy changes of Meta discursively constructed in major European online news outlets; (ii) what are the actual policy changes, taking into account women and LGBTIQ+ identities? (iii) taking into account these situations, it is also important to ask, what should/ could be the European level response to this situation considering policy and civil society level discussions? Having an exploratory approach (Casula et al., 2021; Swedberg, 2020), I carried out a reflexive thematic-discursive analysis of articles from major European news sources published online between 7 January to 7 March 2025 related to the announcements of Mark Zuckerberg as well as responses to these announcements and Meta’s official “Hateful Conduct” policy document and the official statement of GLAAD’s (2025). The mentioned articles, the Hateful Conduct policy document and GLAAD’s official statement were found through the search function of Google and from major European news sources such as TechPolicyPress, EuroNews, Reuters, Wired, BBC, Guardian, DW, Politico, PoliticoEU, LeMonde and the official website of Meta. After doing a thorough reading of the news pieces and Hateful Conduct policy document of Meta as well as GLAAD’s official statement, certain preliminary themes and first reflections emerged and reflected through intersectional feminist standpoint theory, strong objectivity

(Harding, 1986; Hartsock, 1983) and group oppression theory (Young, 1990). The emerging themes of this research are: (i) retreat on human rights, LGBTIQ+ rights and freedom of expression, (ii) frightening developments for the rise of misinformation/ disinformation, (iii) Zuckerberg as “re-setting relations with Europe, ignoring Digital Services Act of EU and the need to debunk “institutionalization of censorship claims.” This article is structured as follows: the materials and methods section that also include the relevant research literature and theoretical framework, results and emerging themes of the analysis and concluding remarks.

2 Materials and methods

2.1 Definition of hate speech and academic research on hate speech

There seems to be a no widely agreed-upon and universal definition of hate speech under international human rights law as well as current academic scholarship. For example, Carlson (2021, p. 2) adopts a broad definition of hate speech and stipulates that it “includes broad categories of speech, including racism, anti-semitism, homophobia, bigotry against the disabled, political hatred, rumour-mongering, misogyny, promotion of terrorism, cyberbullying, harassment, stalking, and the sale and promotion of online products.” Moreover, illegal hate speech is defined in EU law as the public incitement to violence or hatred on the basis of certain characteristics, including race, colour, religion, descent and national or ethnic origin. While the Framework Decision on combatting racism and xenophobia covers only racist and xenophobic speech, the majority of Member States have extended their national laws to other grounds such as sexual orientation, gender identity and disability (European Commission, 2016). To provide a unified framework for the United Nations to address the issue globally, the UN Strategy and Plan of Action on Hate Speech defines 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, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor” (United Nations, 2019). Moreover, hate speech has been defined as a conscious and willful public statement intended to denigrate a group of people (Delgado and Stefancic, 1995).

The European Commission's Recommendation against Racism and Intolerance, No 15 (on combatting hate speech) of 8 December 2015, also cites hatred toward, humiliation of, or contempt for, a “person” who belongs to a group (Paz et al., 2020). Other definitions of hate speech include identifying characteristics, such as race, color, religion, ethnicity or nationality (Tsesis, 2002, p. 211) as well as, gender, sexual identity, or sexual orientation (Lillian, 2007). The greatest challenge for the legal literature, which has addressed this subject most extensively, is to establish a clear differentiation between hate speech and hate crime in order to support the application of criminal sanctions. Prioritizing research on hate speech is affiliated with the increasing media coverage of this phenomenon and its growing presence on social media and the internet in general (Paz et al., 2020). As argued, scientific output on hate speech is not limited to a given field, but it can be found in journals of law, sociology, communication, and psychology, among others (ibid.).

It is comprehended that internet and social media provide the focus of hate speech studies in the communication and media studies area, where the emergent research on digital hate speech and its affective dynamics, reveal the relevance of the digital media environment for the growth and circulation of hate speech (Nikunen, 2023). So, beyond the legal discussions, the topic of hate speech has also been explored as a mundane everyday practice in media studies (Nikunen et al., 2025) where, hate speech is approached from a situational approach (Pohjonen and Udupa, 2017; Udupa et al., 2021) and with “the need to locate discrimination, incitement and hate speech historically within specific socio-political, economic and cultural contexts” (Banaji and Bhat, 2022). In a recent study on a meta-level, how various understandings of hate speech circulate within interactive everyday cultures of digital media, shaped by the digitalized media environment was examined (Nikunen et al., 2025). This mentioned research, theoretically brings together Titley's (2019, p. 3) framework of “debatability” that refers to the constant contestation of “what constitutes racism, as to whose ‘definition’ and voice counts, and to the consequences that should stem from these fractious forms of public recognition and denial” with the post-truth condition. Post-truth condition, in this case understood as the circulation of lies and false beliefs where a part of the circulation of false statements or ambiguous claims is unpurposeful, “a great deal of it is deliberate, strategically aimed at disinforming as a way to manage opponents and/or govern by capturing attention” (Nikunen et al., 2025; Harsin, 2018). Empirical research has investigated hate speech networks in a transnational, interactive, and convergent media environment (Nikunen et al., 2025; Banaji and Bhat, 2022; Govil and Baishya, 2018; Nikunen et al., 2021; Pöyhtäri et al., 2021; Siaperä, 2019; Sponholz and Christofolletti, 2018; Udupa and Pohjonen, 2019), while theoretically oriented research has focused on the definitions and understandings of hate speech (Brown, 2017a, 2017b; Udupa et al., 2021; Udupa and Pohjonen, 2019). The concept of hate speech in these latter studies, has become problematized as either too narrow, regulatory or an all-encompassing universal term (Brown, 2017a).

Among other studies on the topic of hate speech approached from different perspectives, Twitter (Burnap and Williams, 2015; Ott, 2017) and Facebook (Farkas et al., 2018; Kus, 2016) are the platforms that received most attention, followed by YouTube (Murthy and Sharma, 2019). Among these studies, in addition to the refugee crisis, and the rise of extreme right-wing parties in Europe (Pradel, 2024; Blanco-Herrero et al., 2025), special focus is given to racist (Klein, 2012; Matamoros-Fernández and Farkas, 2021) and xenophobic discourse (Yamaguchi, 2013). Studies can also be observed to address Islamophobia (Awan, 2016), homophobia and transphobia (Mršević, 2013; Sánchez-Sánchez et al., 2024; Paz et al., 2020). There are also studies looking into how social media users perceive different forms of online hate speech (Schmid et al., 2022), algorithmic approaches for detecting hate speech (Laaksonen et al., 2020) as well as research on how to combat hate speech (Pukallus and Arthur, 2024). As mentioned before, anti-feminism, homophobia, transphobia, misogyny and transmisogyny are not always regarded under the term hate speech.

In media and communication studies, the research focus has been on the discursive and affective dimensions of hate speech, various definitions of hate speech, how hate speech gets constructed and spreads through social media as well as how it is perceived by different groups, how it can be detected through algorithmic and automated means as well as how it can be combatted (Pukallus and Arthur, 2024). However, there is a lack of research from media and communication

studies perspective about the news coverage and discourses about the existing/ newly emerging policies as well as, to what extent existing hate speech policies and the actual changes to existing policies can have a potential impact on women and LGBTIQ+ identities. Although there is an emerging research literature on anti-feminism, homophobia, transphobia, misogyny and transmisogyny coming primarily from feminist media studies, gender and queer studies and communication studies scholarship, this research, most of the time, is not regarded under hate speech research while research on racism, xenophobia, anti-immigration rhetoric usually are (Matamoros-Fernández and Farkas, 2021). Also, despite the connection between the mechanisms of post-truth and hate speech in the context of vernacular online debates about hate speech is drawn in recent research (Nikunen et al., 2025), how post-truth condition and disinformation becomes an integrative part of hate speech is not conceptualized and researched to the full extent.

Therefore, this article, by extending the legal category of hate speech and hate speech understood as an everyday practice, would introduce and adopt the lens of *hateful speech* to analyze the potential discursive and material ramifications of such speech for women and LGBTIQ+ identities from an intersectional lens. Understood as discourse that, without explicitly inciting hatred, reproduces symbolic and material exclusion through the normalization of degrading or pathologizing language as well as disinformation—hateful speech offers a conceptually pertinent lens for examining discursive practices such as the depiction of LGBTIQ+ identities as mentally ill or unworthy of access to public services.

2.2 Hate speech vs. freedom of speech debate

From a legal perspective, Catherine O'Regan and Stefan Theil of the Bonavero Institute of Human Rights argue that (O'Regan and Theil, 2020), the business model of most social media companies is established on gaining attention and as offensive speech often attracts attention, it can become more widely generated on social media than it might on traditional social media. Through social media platforms (such as Facebook, Twitter, YouTube, Instagram and Snapchat), 3.19 billion users communicate and interact with each other by generating and sharing content. As argued, the urge to exclude or oppress “the enemy” emerges significantly in polarised spaces on social media in the form of coarsened public language, the circulation of extremist views, conspiracy theories and disinformation as well as the rejection of truth and facts. This is facilitated by social media's affordances which are designed to help tech companies achieve their monetary goals by attracting ever-increasing audiences and gathering ever-more data (Pukallus and Arthur, 2024). As pointed out, it is precisely on these digital communicative spaces that hate speech for the first time got artificially and mechanically amplified and disseminated with unprecedented speed, reach and force supported by algorithmic megaphones (Reich et al., 2023). As there exist rising amount of hate speech and offensive speech online, many countries have been considering whether they should be regulating this form of speech and if so how (O'Regan and Theil, 2020). However, it is important to point out that, the regulation of harmful speech in online spaces, requires drawing a fine line between “legitimate freedom of speech” and “hate speech.” Freedom of speech can be perceived to be protected in the constitutions of most countries around the world, as well as in the major

international human rights treaties. However, despite this widespread protection, many countries do not provide effective protection for freedom of speech. One of the dangers of regulating hate speech online (O'Regan and Theil, 2020; O'Regan, 2018), but also what I would argue de-regulating, is that it can act as an excuse for repressive regimes to further limit the rights of their citizens. Accordingly, it can be observed and hypothesized that, for example, this is happening currently in US, where under changes to Hateful Conduct policy document, in the name of endorsing freedom of speech, Meta is hypothesized to allow for the spread of further hate speech against, especially women and LGBTIQ communities aligned with the views of the conservative republican right-wing government in US.

When it comes to freedom of speech, the United States is regarded as being different from other jurisdictions being assessed in some important respects. Accordingly, as also Mark Zuckerberg mentioned in the Facebook video (Zuckerberg, 2025), “The First Amendment of the US Constitution” prohibits the restriction of free speech by government and public authorities. There are narrow exceptions for hate speech, understood as speech that is likely to incite imminent violence. The First Amendment, however, does not prevent private actors, like social media platforms, from imposing their own restrictions on speech. Social media platforms are further protected from private litigation because they are not considered publishers of the content posted to their sites in terms of section 230 of the Communications Decency Act 1996 (O'Regan and Theil, 2020). Hence, as can be observed from the discourses of CEOs who align with right-wing politicians, like Mark Zuckerberg, it can be hypothesized that the notion of truth, for example, the notion of “freedom of speech” and “freedom of expression” can be discursively weaponized to allow the ground for the permissance and spread of more hate speech through the change of respective Meta policies and this might potentially have a detrimental impact especially on women, LGBTIQ+ identities and immigrant groups.

2.3 Freedom of speech vs. current gender equality debates

Globally, it is viable to argue that debates on free speech have become particularly acrimonious around the issue of gender identity and feminism, with a clash between transgender rights and women's rights at its centre (Patel et al., 2022). On one side of the debate, there are ‘gender critical’ feminist voices (regarded as “trans-exclusionary radical feminists,” TERFs) of those who believe that biological sex is significant in categorising a woman; it is a biological reality that is different from a person's gender identity. According to them, patriarchal oppression is grounded partly in biological sex, which is why sex-based rights and protections need to be permitted in law (Patel et al., 2022, p. 14). They argue that the logic of what they call, “the trans gender identity ideology” where sex is regarded as a social construct and assigned at birth-is to erase the legal and political category of sex altogether. According to them, this creates important ramifications for women's rights in all areas of life such as health, sports, education, women only services, prisons and in the collection of government data in the census that informs policy and laws (ibid, p. 14).

While there is acceptance of a spectrum in biological sexual development, whether or not biological sex is an immutable category remains a subject of debate. Judith Butler, in her groundbreaking work in gender studies scholarship, in the book *Gender Trouble* (Butler,

1990), argues that sex is socially constructed as well as gender—this book significantly advances our perception about “the subject” of feminism and the category of “woman.” So, on the other side of the debate, are the progressive transgender rights campaigners, transfeminists and intersectional feminists, who believe that the group of women is already diverse and adding one more group—trans women (that constitute a small percentage of women who have been largely marginalized)—would simply add to this diversity (Patel et al., 2022). Moreover, if a person feels very strongly that the body they were born with is not right for them, it is their right to be able to identify with the one that feels right for them. Speech that disparages trans people potentially damages and erodes their dignity. According to trans activists, a dominant form of disparaging and vilifying trans people include the statement that what counts as a woman is shaped primarily through their biological sex (Patel et al., 2022). The issue between ‘gender critical’ feminists and trans/transfeminist/intersectional feminist activists is thus primarily about what constitutes being a woman. Although there are variations, the viewpoint adopted by gender-critical feminist voices (TERFs) is generally similar to the one adopted by right-wing actors and religious-conservative groups regarding gender equality debates, including in the US. This situation creates a controversy between gender-critical feminist voices and trans/transfeminist/intersectional feminist activists on what constitutes freedom of speech in the public arena surrounding women’s rights and gender equality debates where the right-wing actors are hypothesized to get advantage from the controversy by endorsing their own conservative agendas regarding especially the rights of transgender people. We can observe this more closely from the change of hateful conduct policies below.

2.4 Feminist standpoint theory, strong objectivity, and group oppression theory

Taking into account the aforementioned hate speech vs. freedom of speech and freedom of speech vs. current gender equality debates, and by adopting the notion of *hateful speech*, the Meta “hateful conduct” policy document and the European news coverage related to the policy changes will be interpreted through feminist standpoint theory (Hartsock, 1983), strong objectivity (Harding, 1986) and group oppression theory (Young, 1990). Similar to all standpoint theories, feminist standpoint theory starts with the assumption that society is structured by power relations that generate unequal social locations; one location occupied by members of the dominant group, and other locations are inhabited by members of subordinate groups (Wood, 2009). To develop a feminist standpoint, individuals must engage in an intellectual struggle to recognize, analyze, and contest broad power relations, that account for the subordinate status of girls, and women and the activities they are expected to pursue (Wood, 2009). Connected to feminist standpoint theory, the notion of strong objectivity was first articulated by feminist philosopher Sandra Harding (1986). Strong objectivity builds on the insights of feminist standpoint theory that argues for the importance of starting from the experiences of those who have been traditionally left out of the production of knowledge (Harding, 1986; Naples, 2015). By beginning from the lived experiences of women and others who have been traditionally left outside of the institutions in which knowledge about social life is generated and classified, more objective and more relevant

knowledge can be produced (Naples, 2015). Harding (1986) and Hartsock (1983) argue that knowledge produced from the point of view of subordinated groups may offer stronger objectivity because of the greater motivation of these groups to comprehend the views and perspectives of those in positions of power.

Therefore, a scholar who approaches the research process from the point of view of strong objectivity, is interested in producing knowledge for use as well as in revealing the relations of power, that are hidden in traditional knowledge production processes. Hence, strong objectivity acknowledges that the production of power is a political process and that greater attention paid to the context and social location of knowledge producers, will contribute to a more ethical and transparent result. In fact, Harding (1991) argues that, an approach to research and knowledge production that does not comprehend the role that power and social location play in the knowledge production process must be understood as offering only a weak form of objectivity. Reflexivity is another practice that contributes to strong objectivity. Harding argues for a self-reflexive approach to theorizing in order to foreground how relations of power may be shaping the production of knowledge in different contexts (Naples, 2015). In this way, the knowledge production should involve a collective process rather than individualistic, top-down and distanced approach that typifies the traditional scientific method.

For Harding (1986), strong objectivity involves, analysis of the relationship between the subject and object of inquiry. So, in contrast to traditional models of value-free objectivity, strong objectivity insists that researchers critically examine the role of social situation in the formation of knowledge, exposing background assumptions and cultural agendas. It offers the methodological advice that researchers “start thought from marginalized lives” (Grasswick, 2016). The field of policy analysis in media studies has historically been dominated by top-down approaches to research, so in this study, I am adopting a feminist standpoint theory and strong objectivity to counteract this phenomenon. This would mean, investigating and uncovering, how the power holders, such as CEOs and right-wing politicians can have tremendous power over discourse, policy, governance and action and how this can have an impact on material reality of women and LGBTQ+ identities.

Feminist standpoint theory is more valuable, when it embraces intersectionality (Crenshaw, 1991). From an intersectional perspective (Crenshaw, 1991), it is important to acknowledge that women+ who live with multiple intersecting social identities—differences based on sexual orientation, ethnic identity, age, gender, class, disability, race, tend to experience multiple systems of oppression. Furthermore, it is argued, and I agree that hate speech can be defined as mechanically enabled and amplified language with the intention to harm (Pukallus and Arthur, 2024). I also agree that most prominent group harm that can be brought about by hate speech is that of group oppression as conceived of by Young (1990). For Young (1990), oppression needs to be understood as structural and systematic, with its causes “embedded in unquestioned norms, habits and symbols, in the assumptions underlying institutional rules, and the collective consequences following those rules” (Young, 1990, p. 41).

For Young, the group element is fundamentally important, because group “meanings partially constitute people’s identities, in terms of the cultural forms, social situation, and history that group members know as theirs, because these meanings have been either forced upon them or forged by them or both” (Young, 1990, p. 44).

Hate speech is comprehended as fostering oppression because, it supports “a conceptualization of group difference in terms of unalterable essential natures that determine what group members deserve or are capable of, and that exclude groups so entirely from one another that they have no similarities or overlapping attributes” (Young, 1990, p. 47). Hence, Young (1990) identifies five faces of oppression: exploitation, marginalisation, powerlessness, cultural imperialism, and violence (where dominant groups and meanings render certain groups’ perspectives invisible and negatively stereotype them) while seeking to ensure that all aspects of social life are subordinated to a dominant ideological view (Pukallus and Arthur, 2024). Adopting these feminist theoretical views, this study has an exploratory approach with a working hypothesis that the implied “hateful conduct” policy changes have potential detrimental consequences for women and LGBTIQ+.

2.5 Exploratory approach and thematic-discursive analysis

Exploratory research in social sciences can be defined in different ways, but at its core, it consists of an attempt to discover something new and interesting, by working your way through a research topic (Swedberg, 2020). As argued, exploratory studies have taken a number of different forms in sociology, depending on their goals and means. The two forms that have been the most common are the following: (1) a topic that has not been researched before, is given a first tentative analysis; and (2) an already existing topic is explored in order to produce new ideas and hypotheses, but without being able to properly verify these (Swedberg, 2020). This research adopts the exploratory study approach because it unconventionally attempts to bring together discourse and policy, an analysis of discourses in relation to policy changes as well as the actual policy changes, in this case, regarding the Meta Hateful Conduct Policy change document, which as an approach is novel and understudied. As argued, in order to highlight that certain kinds of applied research lend themselves more easily to deductive approaches, to address the downsides of exploratory qualitative research, and to ensure qualitative rigor in exploratory research, a significant body of work on deductive qualitative approaches has emerged (Casula et al., 2021) (see for example, Gilgun, 2005, 2015; Hyde, 2000). According to Gilgun (2015, p. 3) the use of conceptual frameworks derived from comprehensive reviews of literature and *a priori* theorizing were common practices in qualitative research prior to the publication of Glaser and Strauss’s (1967) *The Discovery of Grounded Theory*. Gilgun (2015) coined the terms Deductive Qualitative Analysis (DQA) to arrive at some sort of “middle-ground” such that the benefits of *a priori* theorizing (structure) and allowing room for new theory to emerge (flexibility) can co-exist simultaneously. According to Gilgun (2015, p. 14) “in DQA, the initial conceptual framework and hypotheses are preliminary. The purpose of DQA is to come up with a better theory than researchers had constructed at the outset (Gilgun, 2005, 2009). Indeed, the production of new, more useful hypotheses is the goal of DQA.” It is pointed out that (Gilgun, 2015, p. 4) “conceptual frameworks are the sources of hypotheses and sensitizing concepts.” As indicated, sensitizing concepts frame the exploratory research process and guide the researcher’s data collection and reporting efforts. According to the authors (Casula et al., 2021, p. 1704), *working hypothesis* can act as a

useful framing tool, where the applicability of the working hypothesis as a tool provides more structure during the design and implementation phases of exploratory research. The working hypothesis is first and foremost a hypothesis or a statement of expectation that is tested in action (Casula et al., 2021, p. 1709). The term “working” suggest that these hypotheses are subject to change, are provisional and the possibility of finding contradictory evidence is real. In addition, a “working” hypothesis is active, it is a tool in an ongoing process of inquiry (ibid, p. 1709). Hence, I think that an exploratory approach with a working hypothesis is well-suited to discourse and policy research in tandem.

Based on this approach, I carried out a reflexive discursive-thematic analysis of articles from major European news sources published online between 7 January to 7 March 2025 related to the announcements of Mark Zuckerberg as well as responses to these announcements and Meta’s official Hateful Conduct policy document and the official statement of GLAAD’s (2025). Many of the news articles emerged during January 2025, however the timeline is kept between 7 January and 7 March to allow a fuller corpus of data. The mentioned articles and the policy document were found through the search function of Google from major well-known European English language news sources such as TechPolicyPress, EuroNews, Reuters, Wired, BBC, Guardian, DW, PoliticoEU, Politico and LeMonde and the official website of Meta and GLAAD. The corpus was selected through the Google search function by using keywords such as “Meta,” “hateful conduct,” “policy change,” “human rights,” “LGBTIQ+,” “women,” “fact-checking,” “disinformation” and “7 January.” Majorly, the search was made by using English language. Initially the articles were saved to the “reading list” tab of researcher’s MAC computer. The researcher did two or three initial readings of the documents and when it was realized that same/similar patterns kept repeating, it was understood that the data was representative and reached a saturation point at 14 articles—so in total, 14 articles were analyzed. The reflexive, in-depth thematic-discursive analysis was made through the lens of *hateful speech*.

Comprehended as discourse that, without explicitly inciting hatred, reproduces symbolic and material exclusion through the normalization of degrading or pathologizing language and disinformation, hateful speech offers a conceptually pertinent lens for examining discursive practices such as the depiction of LGBTIQ+ identities as mentally ill or unworthy of access to public services. Moreover, it attempts to understand how certain forms of misinformation/disinformation can have specific and more severe impact on women and LGBTIQ+ identities. After doing a thorough reading of the news pieces and the policy document, I carried out, what I call a discursive-thematic analysis (Herzog et al., 2019) where the data was categorized according to emerging themes. Thematic analysis is regarded as one of the most straightforward ways of deducing patterns of meaning—referred to as themes—from qualitative data. In its essence, thematic analysis consists of the analytical construction of: (a) codes, (b) themes in qualitative verbal expressions, and (c) patterns of recurrence, evaluation or associations within these themes (Herzog et al., 2019). In the analytical procedure, in addition to the thematic focus, the discursive orientation of the articles was distinguished where a systematic table (in Appendix) featuring key variables such as outlet, type of text, publication date, thematic focus, and discursive orientation were listed and presented. After the analytical procedure, certain preliminary themes emerged

as a result of this exploratory interpretive approach that was *reflective, reflexive and immersive*.

3 Results

3.1 News coverage and policy change analysis in tandem

While the hateful conduct policy document changes are comprehended to be “blocking of censorship and restoration of freedom of expression” by the CEOs like [Zuckerberg \(2025\)](#) and Meta Chief Officer, Joel Kaplan, emerging themes demonstrate that, in the major news outlets coverage, changes are discursively articulated as a (i) retreat on human rights, lgbtiq+ rights and freedom of expression (ii) frightening development for fuel of disinformation/misinformation, (iii) ignorance of the Digital Services Act of EU, Zuckerberg as “re-setting relations with Europe,” and the need to debunk “institutionalization of censorship” claims. Although the discursive articulations in the news coverage touch upon some characteristics of these changes and the ramifications for women and LGBTIQ+ identities, they need to be considered in tandem with actual policy changes to reflect and elaborate more in-depth on these potential dynamics and material and structural ramifications for women and LGBTIQ+ identities.

3.2 Retreat on human rights, LGBTIQ+ rights and freedom of expression

On one level, these policy moves of Meta are considered to be in breach of human rights. The justification provided by Mark Zuckerberg about “protecting free speech,” although referring to the fundamental right of freedom of expression, is not considered to be staying valid under human rights standards as it becomes a weaponization of the concept of truth, in this case, “freedom of expression” and “freedom of speech.” In the news coverage of TechPolicy, [Rosenblat \(2025\)](#), a policy advisor on technology and law states that Zuckerberg’s decision to prioritize free speech at the expense of countering hate speech and incitement to discrimination is out of step with human rights standards and a breach of his public-facing commitments. She says ([Rosenblat, 2025](#)), “While freedom of expression is a widely recognized human right, it is subject to certain limitations that protect other rights or societal interests, such as protecting against imminent violence.” International Covenant on Civil and Political Rights, a legally binding treaty, acknowledges that the exercise of freedom of expression may be limited, if it is necessary, “to protect the rights and reputations of others,” while limitations on expression are actually needed, when the speech constitutes “advocacy of national, racial, or religious hatred that constitutes incitement to discrimination, hostility or violence” ([Rosenblat, 2025](#)). As pointed out, human rights law makes a distinction between hate speech and merely offensive speech, while Meta omits this distinction, especially when relating to topics such as transgender rights. Hence, statements like “democracy is a worthless system” or “all religion is garbage” are arguably allowed under human rights doctrine, while threats against people based on their sexual orientation (e.g., if you are gay, you do not deserve to live) or dehumanizing slurs meant to incite hatred

against racial group (e.g., let us deport all Jews to concentration camps”) are not.

Meta can be regarded as eliding the distinction, as it is stated in a Guardian coverage by [Booth \(2025\)](#) of the change of policies on hateful content: “Meta’s global policies on hateful content now include allowing users to call transgender people “it,” with the guidelines stating: We do allow allegations of mental illness or abnormality when based on gender or sexual orientation.” Wired piece by [Knibbs \(2025\)](#), also mentions this point: “Some of the most striking changes were made to Meta’s “Hateful Conduct Policy,” which covers discussions on immigration and gender. In a notable shift, the company now allows, “Allegations of mental illness or abnormality when based on gender or sexual orientation, given political and religious discourse about “transgenderism” and homosexuality and common non-serious usage of words like “weird.”” It is further claimed in a [DW \(2025\)](#) that, Meta platforms would “simplify” their content policies “and get rid of a bunch of restrictions on topics like immigration and gender because they are just out of touch with mainstream discourse.” Furthermore, it is pointed out in a TechPolicy Press piece by senior director of the GLAAD social media safety program Jenni Olson that ([Olson, 2025](#)), “These new policy and product changes presage (and very obviously mirror) the relentless deluge of anti-trans and anti-DEI efforts now emanating from the White House. And they send a clear message that Meta and its leadership may *hold (and espouse)* bigoted, homophobic, and transphobic beliefs about LGBTQ people.” What’s more, Olson’s further comments in the same piece are striking: “The purpose of maintaining hate speech policies is to keep all users safe, not to put a target on the backs of one historically marginalized group. Meta enshrining a statement that LGBTQ people are “abnormal” and “mentally ill,” and intentionally employing overtly bigoted anti-LGBTQ dog whistle language, deliberately *expressing* animus towards one group of people based on their protected characteristics in a hate speech policy—is unprecedented” ([Olson, 2025](#)).

While the news coverage touches on some aspects of the impact of policy changes on women and LGBTIQ+ identities, it does not actually explain in full detail the changes of these hate speech policies, and there is the necessity of examining especially the actual policy document of Meta (self-regulative community guidelines), such as the “Hateful Conduct Policy” document to perceive the actual changes as well as assessing their potential impact on women and LGBTIQ+ identities. Not only the allowance of LGBTIQ+ people to be called mentally ill or abnormal, but also a range of other changes were made to the Hateful Conduct Policy document starting with Meta’s definition of hate speech and slurs as well as a wide range of changes—permittances and allowances of problematic content to circulate online that are detrimental to human rights and LGBTIQ rights from an intersectional feminist standpoint, strong objectivity and group oppression theory. According to GLAAD (the largest lesbian, gay, bisexual, transgender, and queer (LGBTQ) media advocacy organization and network) ([GLAAD, 2025](#)), Meta first started its efforts to combat hate speech in the 2010s, adding improvements over the last decade. Meta is considered to formalize its hate speech policy in 2018, clarifying definitions of hate speech including attacks based on race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender identity, and disability. Meta was observed to make changes to its hate speech policy on 7th January 2025 following the announcements of its CEO Mark Zuckerberg. So, when the tab 7th January 2025 on the official website is exploratively

observed and analyzed as well as the official statement of GLAAD (2025) published on 7 January, the new “Hateful Conduct Policies,” “Community Guidelines” of Meta can be found. In the official tab published on 7th January 2025, the removed/ changed policy statements are crossed over with red and new additions are observed to be made. According to this mentioned online document with crossed-over changes and the official statement of GLAAD, following are the observed changes:

a. Adjustment for Meta’s definition of hate speech.

- Hatespeech was altered to Hateful Conduct: “We do not allow Hateful Conduct on Facebook, Instagram and Threads.”
- The clause “it creates an environment of intimidation and exclusion, and in some cases, may promote offline violence” is deleted.
- The clause “we also remove statements of inferiority, expressions or of contempt or disgust; cursing; and calls for exclusion or segregation” is deleted.
- The clause “we define a hate speech attack as dehumanizing speech” is changed to “we remove dehumanizing speech.”

b. Adjustment for Meta’s definition of slurs.

- The clause, “we also prohibit the use of slurs that are used to attack people on the basis of their protected characteristics” is deleted.

c. Eliminating the prohibition against the dehumanization of women and non-binary people.

- The clause “Do Not Post,” certain objects (women as household objects or property or objects in general; black people as farm equipment; transgender or non-binary people as “it”) is deleted.”

d. Eliminating the prohibition against statements denying existence.

- The clause, “protected characteristics or quasi protected characteristics do not exist,” “no such thing as protected characteristics or quasi protected characteristics or protected characteristics or quasi protected characteristics should not exist.”

e. Permittance of attacks on LGBTIQ+ people, women and immigrants:

2025 Hateful Conduct Policy states “People sometimes use sex-or gender exclusive language when discussing access to spaces often limited by sex or gender, such as access to bathrooms, specific schools,

specific military, law enforcement, or teaching roles, and health or support groups. Other times, they call for exclusion, or use insulting language in the context of discussing political or religious topics, such as when discussing transgender rights, immigration or homosexuality. Finally sometimes, people curse at a gender in the context of a break-up. Our policies are designed to allow room for these types of speech.”

f. Permittance for economic discrimination of women and transgender people:

2025 Hateful Conduct Policy states: “Economic exclusion which means denying access to economic entitlements and limiting participation in the labor market. We do allow content arguing for gender-based limitations of military, law enforcement, and teaching jobs. We also allow the same content based on sexual orientation, when the content is based on religious beliefs.”

g. Permittance for social exclusion of women and LGBTIQ+ people:

2025 Hateful Conduct Policy states: “It now allows content about support for denying access to spaces (physical and online), and social services, except for sex or gender-based exclusion from spaces commonly limited by sex or gender, such as restrooms, sports and sports leagues, health and support groups and specific schools.”

h. Permittance for statements that LGBTIQ+ people to be called “mentally ill” that is also widely covered by the news articles:

2025 Hateful Conduct Policy states: “We do allow allegations of mental illness or abnormality when based on gender or sexual orientation, given political and religious discourse about transgenderism [sic] and homosexuality and common non-serious usage of words like ‘weird.’” (*The current usage of “transgenderism” arises from anti-trans extremists who seek to delegitimize and dehumanize trans people by implying that being trans is an ideology rather than an identity, GLAAD, 2025*).

i. The clause “we have additional restrictions for paid content” is deleted.

Going beyond just regarding these changes as any other attack on LGBTIQ+ rights and human rights, I argue that these policy changes need to be perceived as part of the broader anti-gender backlash. Judith Butler asks, “can we say how many contemporary fears gather at the site of gender?” (Butler, 2024, p. 6). When the word “gender” absorbs an array of fears and becomes a catch-all phantasm for the contemporary Right, the various conditions that actually give rise to those fears lose their names. It is pointed out that gender both collects and incites those fears, keeping us away from what there is actually to fear, and how the current grim state of the world came about in the first place. Circulating and disseminating the phantasm of gender is also one way for existing powers—states, churches and political movements—to frighten people to come back to their ranks, and to externalize their fear and hatred onto vulnerable communities. As argued, the weaponization of the fearsome phantasm of “gender” is authoritarian at its core as rolling back progressive legislation is surely fuelled by backlash. However, more than this, the project of restoring

the world to a time before “gender” promises a return to a patriarchal dream order that may never have existed but that occupies the place of “history” or “nature”—an order that only a strong state can restore. The targeting of sexual and gender minorities as dangers to society, as pointing towards the most destructive forces in the world, in order to strip them off their fundamental rights, protections and freedoms, refers to the anti-gender ideology in fascism (Butler, 2024, p. 6–8).

The existing research reveals the existence of several discursive axes of hate and conspiracy thinking, with interconnected patterns of communication on social media platforms (Santos Fernández, 2024). Among the narratives identified, the pathologisation of affective-sexual diversity, biological reductionism, heterosexism and Catholic morality stand out, with ideas associating LGBTIQ+ people with paedophilia, indoctrination and genital mutilation of minors, the erasure of cisgender women and the destruction of social order and cultural models (Santos Fernández, 2024).

Needless to say, in general, it is arguable that the complex balancing between the need to guarantee people the right to freely express themselves and to advocate their ideas on one side, and the defence of other people’s right to be free from verbal abuse and to be protected as equal members of a society on the other, is not easy (Gorenc, 2022). However, from an intersectional feminist standpoint, strong objectivity and group oppression theory, it can be articulated that these new changes has right-wing authoritarian tendencies and can potentially pose novel and heightened risks for women and LGBTIQ+ people as it increases their chances of encounter with elevated levels of misogyny, homophobia, and transphobia in online and beyond the online spaces. Existing research points out that dangerous anti-LGBTIQ+ discourses emanating from manipulated news feeds and subsequently intertwined together across social networking sites may progressively lead to the construction of an LGBTIQ+ conspiracy theory and wreak damage upon a minority community (Balirano and Hughes, 2023). Understood in this way, it is not viable to talk about these policy changes as leading towards online content and discourses that can be considered as legitimate forms of “free speech.” Furthermore, in the context of trans rights, TERFs, for example, oppose transgender people’s basic claims of self-determination, freedom and autonomy, rights to be protected from violence, and rights of access to public space, and to health care without discrimination, all of which are they, as feminists, fight for and depend on otherwise (Butler, 2024, p. 149). As the policy clause i. “we have additional restrictions for paid content” is deleted as well, so with these changes, Meta will be continuing to potentially normalise anti-women and anti-LGBTIQ hatred for profit—at the expense of its users and true freedom of speech and expression.

Helle Thorning-Schmidt, co-chair of Meta’s oversight board and former Prime Minister of Denmark, told the BBC she welcomed aspects of the shake-up, which will see users decide about the accuracy of posts via X-style “community notes” (Fraser, 2025). However, speaking on BBC Radio 4’s today programme, she said, there were “huge problems” with what had been announced, including the potential impact on the LGBTIQ+ community, as well as gender and trans rights. She said “We are seeing many instances where hate speech can lead to real-life harm, so we will be watching that space very carefully.” In an open letter shared with the Tech Brief ahead of its publication Thursday, the Global Coalition for Tech Justice says “Meta has abandoned any pretense of oversight.” The coalition, which says it represents more than 250 organizations and experts from 55 countries,

criticizes the board for not publicly pushing back on changes that it says will foster lies, degrade discourse and fuel attacks on women and LGBTIQ+ people. From an intersectional feminist standpoint theory and strong objectivity (Harding, 1986; Hartsock, 1983), although the discursive statements and comprehensions in news coverage are understood to touch on the importance of the ramifications of the changes of policies on women and LGBTIQ+ identities, they still remain to be cursory, vague and discursive and they do not really go in-depth to understand and explain the specifics of these changes and how they can have an impact on these identities structurally while not commenting on the existence of the anti-gender backlash.

3.3 Frightening developments for the rise of misinformation/disinformation

Moreover, in the news coverage, these moves by Meta are comprehended to be “frightening” and the breeding ground for disinformation/misinformation. In a EuroNews article (2025), Carlos Hernández-Echevarría, associate director of the independent Spanish fact-checking platform, Maldita.es says “You have to see the celebration on the channels of well-known disinformation actors to know this is bad news for Meta users, while he added “fact-checkers did not censor anyone, and never ever asked Meta to remove anything legal” (Chadwick, 2025). In an article on Politico (Körömi et al., 2025), the changes are considered as Mark Zuckerberg aligning with Elon Musk: “Zuck goes full Musk, dumps Facebook fact-checking program.” Moreover, the Nobel prize winning American-Filipino journalist Maria Ressa predicts in a Guardian article (Milmo, 2025), “extremely dangerous times” as she said “Meta is going to “allow lies, anger, fear and hate to infect every single person on the platform.” While it is important to acknowledge that disinformation can influence different segments of society and “all” the people, from the approach of intersectional feminist standpoint theory and strong objectivity, disinformation, misogyny, homophobia, transphobia can have especially adverse effects on women and LGBTIQ+ people and hence, should be taken into account accordingly.

Especially some changes related to policy clauses have adverse effects on LGBTIQ+ rights and human rights as articulated before in the article, while having the potential to create misinformation/disinformation which are not covered in the news coverage accordingly—disinformation and misinformation tend to affect just “everyone” in the news coverage and not especially specific groups of people like women and LGBTIQ+. Based on the policy changes; potentially, the changes mentioned above; c. eliminating the prohibition against the dehumanization of women and non-binary people, d. eliminating the prohibition against statements denying existence, e. permittance of attacks on LGBTIQ+ people, women and immigrants, g. permittance for social exclusion of women and LGBTIQ+ people, h. permittance for statements that LGBTIQ+ people to be called “mentally ill” would potentially allow for the further production, circulation and dissemination of misinformation/disinformation in the online spaces with regards to women and LGBTIQ+ people. Existing research already demonstrates that disinformation/misinformation, misogyny, homophobia, transphobia is flourishing in online platforms.

It is important to also take into account this anti-gender backlash in the production and further circulation of disinformation/

misinformation, misogyny, homophobia and transphobia. As pointed out before, there is a clash between TERFs (trans-exclusionary radical feminists) who comply with right-wing actors on many levels (in our case US president Donald Trump, his government actors, supporters, related conservative civil society and religious actors and technology CEOs like Mark Zuckerberg) and trans-inclusive intersectional feminists considering the rights and existence of transgender people. There is existing research that point towards the claim that especially gender-critical discourse can constitute a coordinated disinformation campaign that is part of a broader political strategy to oppose transgender rights (Billard, 2023). Various anti-trans actors—including TERF organizations LGB Alliance and Fair Play for Women in the United Kingdom and conservative political advocacy organizations like the Alliance Defending Freedom and Concerned Women for America in the United States—paid money to Meta to ensure the wider spread of disinformation on their platforms with a clear intention of justifying anti-transgender policies (Billard, 2022). From a feminist standpoint theory, strong objectivity and group oppression theory, I agree that these disinformation campaigns have material political effects that go beyond rhetorical claims of harm; they have succeeded in shaping policy in ways that curtail the rights and damage the life chances of trans people at various levels of governance (Billard, 2022; Hughes et al., 2021; Johnson, 2022; Lepore et al., 2022; Park et al., 2021). The disinformation Billard describes involves othering narratives that alienate trans people, essentializing narratives that craft negative generalizing tropes about trans populations, and authenticating narratives that undermine (in particular) trans women's claims to womanhood (see also Bassi and LaFleur, 2022; Bettcher, 2017; Koyama, 2020; Pape, 2022). From a feminist standpoint theory, strong objectivity and group oppression theory, I also agree that disinformation not only relies on epistemologically suspect claims but also entails “strategic attempts to undermine and weaponize people's social identities and group membership for political gains ... in accord with racial and other structures that underlie social differentiation” (Reddi et al., 2021, p. 13). As gender-critical trans-exclusionary feminists align with right wing actors and technology CEOs on their views regarding LGBTIQ+ people, this has the potential to be detrimental to the material well-being of women and LGBTIQ+ considering the changes to the hateful conduct policy document of Meta. This is the case because, the allowance and permissance of more derogatory speech on social media platforms are not only on discursive and rhetorical levels, but they can potentially have other real and material consequences on the lives of women and LGBTIQ+ people including real life violence and harm but also potential change of policies in other areas that can significantly impact their well-being.

In response to the anti-gender articulations where the analysed hateful conduct policy changes can be considered to be part of, Judith Butler, the voice of trans-inclusive feminism reminds us that, feminism, has always been a struggle for justice, or is, at its best, precisely such a struggle, formed in alliance and affirming difference. Trans-exclusionary feminism is not feminism, or rather, it should not be (Butler, 2024, p.168). There is no sense of replicating the transphobia of the Right in the name of feminism, to further the phantasm, since what is required is an alliance that knows and strengthens the interdependencies without which we cannot live. Can feminism join in an alliance against the forces of destruction rather than become a destructive force allied with other such forces Butler asks (Butler, 2024, p. 169). This is an affirmative question given how

central to new fascism are the vicious attacks on women, trans people, gay and lesbian people, Black and brown people, who belong to all these categories, and in whom all these categories also live. The categories need to open for so many to live, to find life livable, even as the categories are important to seize for those who have not yet been recognized within their terms (Butler, 2024, p. 169). The category of “woman” does not say in advance, how many people can participate in the reality it describes, nor does it limit in advance, the forms that reality can take. In fact, feminism has always insisted that what a woman is, is an open-ended question, a premise that has allowed women to pursue possibilities that were traditionally denied to their sex (Butler, 2024, p. 145).

3.4 Zuckerberg as “re-setting relations with Europe, ignoring digital services act of EU and the need to debunk ‘institutionalization of censorship’ claims”

Furthermore, it is comprehended that the European digital regulation comes under attack from Trump, Musk and Zuckerberg and the recent actions potentially go against EU's Digital Services Act. In a LeMonde piece (Malingre, 2025), “The European Union seems powerless to counter the assault waged by the US president-elect and Big Tech, against laws designed to regulate the EU's digital space.” Moreover, Zuckerberg is understood to be re-negotiating and re-calibrating the relations of Europe and US where Zuckerberg is also going against what he calls “the institutionalization of censorship.” In a Reuters article by George Hay and Kwok (2025), they say “Mark Zuckerberg is, among other things, re-setting his relations with Europe...while Zuckerberg criticized European technology regulations and pledged to work with President-elect Trump to “push back on governments going after American companies and pushing to censor more.” Oliver Marsh, a former Downing Street adviser and head of technology research at the Berlin and Zurich non-profit Algorithwatch, moreover, said in Guardian article. “If these policy changes mean that you can spread lies, that end with attacks on groups then there is a case Meta would be going against the EU's digital services act.” Moreover, it is mentioned in a PoliticoEurope coverage that (Gkritsi and Griera, 2025), “Freedom of speech is a hot-button issue for Trump and his followers, who see EU efforts to curb online disinformation as excessive.” These remarks in the news media, although important, remain to be too vague, taking into account the ramifications of the changes for women and LGBTIQ+ especially from the perspective of EU's revised Code of Conduct on countering illegal hate speech.

In response to Mark Zuckerberg saying that “EU is institutionalizing censorship (Zuckerberg, 2025), European Commission Vice-President Henna Virkkunen told Politico (Vax and Haeck, 2025) that “We know that it's not true. In Europe, freedom of speech is one of our fundamental values and it's also respected and protected in our Digital Services Act. So it's very misleading also to say that.” From an intersectional feminist standpoint and strong objectivity, as European Union does not censor any speech that is legal, on 20 January 2025, the revised Code of Conduct on countering illegal hate speech online was integrated into the regulatory framework of the Digital Services Act seeking to prevent and address the spread of illegal hate speech online (European Commission, 2025). The Code of Conduct+ which builds on the code of conduct

in 2016, strengthens the way online platforms deal with content deemed illegal hate speech according to EU law and Member States' laws (ibid.). However, it is observable that the definition of hate speech in mentioned code of conduct is very vague and ambiguous, not specifying the different characteristics of disinformation, misogyny, homophobia, transphobia especially experienced by women and LGBTIQ+ identities that can be detrimental to their existence and well-being.

Moreover, although Mark Zuckerberg's claims about the EU as institutionalizing censorship are too far-fetched, it could be beneficial to think that curating and moderating online content require well-thought principles and steps from the concrete and nuanced definition of what hate speech constitutes, taking into account context specific sensitivities to the involvement of various stakeholders while preparing and drafting such a policy document to protect the freedom of speech for the most marginalized groups of women and LGBTIQ+. However, this approach might be too idealizing as many positive conceptions of freedom, or notions of positive communication rights, can also be criticized for trying to develop a pre-determined lists of abstract, universal preconditions, or specific ends that "authentic" communicative freedom would involve. In line with [Berlin's \(1969\)](#) famous discussion, positive freedom can be criticized for paternalism or essentialism, trying to define the communicative needs and rights of citizens from above ([Karppinen, 2019](#)).

It goes without saying that, while the endorsement and allowance of right-wing authoritarian hateful rhetoric should be unacceptable, some scholars like [Strossen \(2018, p. 164\)](#) points out, "it is essential for the well-being of both individuals and society that we encourage and facilitate (...) counterspeech rather than adopting the disempowering, anti-democratic censorial approach." [Benesch et al. \(2016, p.17\)](#) identify eight counter-speech strategies to combat hate speech that have proven effective in reducing hate speech online and on social media in particular, these are: (1) presentation of facts to correct mis-statements or misperceptions, (2) pointing out hypocrisy or contradictions, (3) warning of possible offline and online consequences of speech, (4) identification with original speaker or target group, (5) denouncing speech as hateful or dangerous, (6) use of visual media, (7) use of humor, and (8) use of a particular tone. As it is pointed out, evidence points out that counter-speech therefore stands as an effective alternative to silence, censorship or counterproductive "cancelling." In the spirit of this form of counterspeech, intersectional feminist scholars such as bell Hooks, argue that feminists need to "speak up," "talk back" at right-wing discourses and authoritarianism. For [hooks \(1989\)](#), talking back refers to the practice of adjusting with your speaking partner, rising to their speaking position, and "speaking as an equal to an authority figure." According to her, it is through this act of speaking that voices become liberated ([Sorce and Thomas, 2025](#)). When performed in a public setting, like online spaces, speaking up holds performative power as it encourages co-witnessing, and sharing in dissent through solidarity. [Hooks \(1989, p. 9\)](#) comprehends this experience as empowering: "Moving from silence into speech is for the oppressed, the colonized, the exploited, and those who stand and struggle side by side a gesture of defiance that heals, that makes new life and new growth possible."

When we look at non-ideal approaches to freedom of speech, the central claim of the more radical or agonistic theories of democracy, for instance, proposed by [Mouffe \(2000, 2005\)](#) is that, not only the

liberal model of the marketplace of ideas, but also, ideal conceptions of a rational and deliberative public sphere fail to sufficiently address the inevitable nature of power and existing forms of exclusion ([Karppinen, 2019](#)). Due to focusing on the ineradicable nature of hegemonic power relations, the purpose of promoting freedom of communication from the radical democratic perspective is not the complete elimination of power relations but their continuing contestation. According to [Brown \(1995, p. 6\)](#), freedom is neither a philosophical absolute nor a tangible entity, but a relational and contextual practice, that takes shape in opposition to whatever is locally and ideologically conceived as unfreedom." It is pointed out that such radical anti-foundationalism may not at first sight seem like a very sustainable basis for defending or promoting communicative freedom.

Postmodern or poststructuralist approaches can be easy to criticize for extreme relativism and subjectivism that abandons all principles, allowing the notion of freedom to be used for any purpose by anyone. On the other hand, it is pointed out that ([Karppinen, 2019](#)), in contrast to the absolute or universalizing rhetoric, a position that recognizes the contextual and limited nature of freedom can also be seen as a more tenable and empowering basis for expanding and reimagining communicative freedom as a normative value under present conditions. However, in response to these approaches, I would argue that a position of feminist standpoint theory, strong objectivity and group oppression theory ([Harding, 1986; Hartsock, 1983; Young, 1990](#)), would be better suited and attuned to the real and material conditions of women, LGBTIQ+ identities and minorities and would be able to better comprehend, elaborate and articulate the characteristics of the communicative freedom these groups necessitate beyond the online and offline division. This can circumvent the more top-down approaches adopted by the EU in hate speech policymaking procedure and open the space for grassroots, bottom-up, collaborative and multi-group approaches. Therefore, while it is important to be able to "speak up" and "talk back" (1989) at hate speech as a feminist approach, it is equally important to be able to develop bottom-up approaches/ ways of regulating and moderating hate speech starting from the experiences of the most marginalized, and this procedure can inform EU policy making in this area as well.

4 Concluding remarks

Following the announcements of Meta CEO [Zuckerberg \(2025\)](#), this article examined (i) how are the proposed policy changes of Meta discursively constructed in major European online news outlets; (ii) what are the actual policy changes, taking into account women and LGBTIQ+ identities? (iii) what should/ could be the European level response to this situation considering policy and civil society level discussions? Based on these research questions and by adopting an exploratory approach, I reflexively analysed articles from major European news sources published online between 7 January and 31 January 2025 related to the announcements of Mark Zuckerberg as well as responses to these announcements and Meta's official "Hateful Conduct" policy document. Adopting an intersectional feminist standpoint theory, strong objectivity and group oppression theory, the emerging themes of this research are (i) retreat on human rights, LGBTIQ+

rights and freedom of expression, (ii) frightening developments for the rise of misinformation/disinformation, (iii) Zuckerberg as “re-setting relations with Europe, ignoring Digital Services Act of EU and the need to debunk ‘institutionalization of censorship claims.’” As there is a lack of research from media and communication studies perspective about the news coverage and discourses about the existing/emerging policies as well as, to what extent existing hate speech policies and changes to existing policies can have an impact on women and LGBTIQ+ identities, this article attempted to make a novel contribution to academic scholarship by studying the actual policy changes in tandem and intertwined with the discourses about the policy changes in online news media with a feminist exploratory approach and from a feminist standpoint. This also allowed extending the legal category of hate speech and hate speech as mundane everyday practice, where the article proposed the lens of hateful speech: comprehended as discourse that without openly inciting to hatred, this form of speech articulates symbolic and material exclusion through the normalization of degrading or pathologizing language and disinformation, and therefore offers a conceptually pertinent lens for examining discursive practices such as the depiction of women and LGBTIQ+ identities as mentally ill or unworthy of access to public services.

In this way, as well as revealing the actual implications of policy changes to material realities of women and LGBTIQ+, this study could also uncover the rising anti-gender backlash and the existing authoritarian power dynamics camouflaged and implied in these policy moves by Meta actors and the right-wing conservatives and how gender-critical feminist voices align with right-wing actors in their endeavors. Going beyond the postmodern or poststructuralist approaches, I also argued that a position of feminist standpoint theory, strong objectivity and group oppression theory (Harding, 1986; Hartsock, 1983; Young, 1990), in addition to being suited and attuned to the real and material conditions of women, LGBTIQ+ identities and minorities, would also be able to better comprehend, elaborate and articulate the characteristics of the communicative freedom these groups necessitate beyond the online and offline division. This can circumvent the more top-down approaches adopted by the EU in hate speech policymaking procedure and open the space for grassroots, bottom-up, collaborative and multi-group approaches in the face of this rising authoritarianism. Furthermore, this research revealed the necessity of making more nuanced definitions of hate speech that would take into account speech and discourses that violate the rights of women and LGBTIQ+, misogynist, homophobic and transphobic discourses that can transform into coordinated disinformation campaigns and have other real and material consequences on the lives of women and LGBTIQ+ people including real life violence and harm but also this form of speech leading up to a change of policies in other areas that can significantly impact their well-being. Therefore, while it is comprehended to be prominent to be able to “speak up” and “talk back” (1989) at hate speech as a feminist approach in the manner of bell Hooks, it is equally important to be able to develop bottom-up approaches/ways of regulating and moderating hate speech starting from the experiences of the most marginalized, and I would argue that this procedure can inform EU policy making in this area for the future as well.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Author contributions

DA: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

The author would like to thank Marlen Komorowski for sending an invitation to submit her work to this edited volume.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Correction note

This article has been corrected with minor changes. These changes do not impact the scientific content of the article.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2025.1640230/full#supplementary-material>

References

- Abaday, D. U. (2024). "How Turkish anti-gender actors interpret trump's win", the loop ECPR's political science blog, 23 December. Available at: <https://theloopecpr.eu/how-turkish-anti-gender-actors-interpret-trumps-win/> (Accessed May 26, 2025).
- Asardag, D., and Donders, K. (2021). C'est quoi post-truth, new era or time-immemorial? Conceptual reflections for democracy, power and free speech. *Int. Commun. Gaz.* 84, 136–156. doi: 10.1177/1748048520983731
- Awan, I. (2016). Islamophobia on social media: a qualitative analysis of the facebook's walls of hate. *Int. J. Cyber Criminol.* 10, 1–20. doi: 10.5281/zenodo.58517
- Balirano, G., and Hughes, B. (2023). "The rainbow conspiracy: a Corpus-based social media analysis of anti-LGBTIQ+ rhetoric in digital landscapes" in *The Routledge handbook of discourse and disinformation*. eds. S. M. Maci, M. Demata, M. McGlashan and P. Seargeant (London/New York: Routledge), 306–324.
- Banaji, S., and Bhat, R. (2022). *Social media and hate*: Routledge.
- Bassi, S., and LaFleur, G. (2022). TERFs, gender-critical movements, and postfascist feminisms. *Transgender Stud. Q.* 9, 311–333. doi: 10.1215/23289252-9836008
- Benesch, S., Ruths, D., Dillon, K. P., Saleem, H. M., and Wright, L. (2016). "Counterspeech on twitter: A field study", Public Safety Canada.
- Berlin, I. (1969). "Two concepts of liberty", In Berlin, I. *Four essays on liberty*. Oxford, UK: Oxford University Press.
- Bettcher, T. M. (2017). Trans feminism: recent philosophical developments. *Philos Compass* 12:e12438. doi: 10.1111/phc3.12438
- Billard, T. J. (2022). The politics of transgender health misinformation [manuscript submitted for publication]. Department of Communication Studies: Northwestern University, Evanston, Illinois.
- Billard, T. J. (2023). "Gender-critical" discourse as disinformation: unpacking TERF strategies of political communication. *Women's Stud. Commun.* 46, 235–243. doi: 10.1080/07491409.2023.2193545
- Blanco-Herrero, D., Trilling, D., and Arcila-Calderón, C. (2025). I'll share it when i believe it! An experimental study on the effects of hateful and false content on credibility and sharing intention. *Int. J. Commun.* 19, 974–997.
- Booth, R. (2025). "Revisions of 'hateful conduct': what users can now say on Meta platforms" in *Guardian* 8 January. Available at: <https://www.theguardian.com/technology/2025/jan/08/permitted-hateful-conduct-what-users-can-now-say-on-meta-platforms> (Accessed July 1, 2025).
- Brown, W. (1995). *States of injury: Power and freedom in late modernity*. Princeton, NJ: Princeton University Press.
- Brown, A. (2017a). What is hate speech? Part 1: the myth of hate. *Law Philos.* 36, 419–468. doi: 10.1007/s10982-017-9297-1
- Brown, A. (2017b). What is hate speech? Part 2: family resemblances. *Law Philos.* 36, 561–613. doi: 10.1007/s10982-017-9300-x
- Burnap, P., and 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, 223–242. doi: 10.1002/poi3.85
- Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. New York: Routledge.
- Butler, J. (2024). *Who's afraid of gender*. Dublin: Penguin Books.
- Carlson, C. R. (2021). *Hate Speech*. Cambridge, Massachusetts: The MIT Press.
- Casula, M., Rangarajan, N., and Shields, P. (2021). The potential of working hypotheses for deductive exploratory research. *Qual. Quant.* 55, 1703–1725. doi: 10.1007/s11135-020-01072-9
- Chadwick, L. (2025). "Meta to replace fact-checkers with 'community notes' system similar to X" in *EuroNews*, 8 January. Available at: <https://www.euronews.com/next/2025/01/07/meta-to-replace-fact-checkers-with-community-notes-system-similar-to-x> (Accessed May 26, 2025).
- Crenshaw, K. (1991). Mapping the margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Rev.* 43, 1241–1299. doi: 10.2307/1229039
- Delgado, R., and Stefancic, J. (1995). Ten arguments against hate-speech regulation: how valid symposium: political correctness in the 1990's and beyond. *N. Ky. L. Rev.* 23:475.
- DW (2025). "Meta ends fact-checking to 'reduce censorship'" in *DW*, 7 January. Available at: <https://www.dw.com/en/meta-ends-fact-checking-introduces-community-notes-v2/a-71238829> (Accessed May 26, 2025).
- European Commission (2016). The code of conduct illegal online hate speech questions and answers. Brussels: Directorate-General for Justice and Consumers.
- European Commission (2025). The code of conduct on countering illegal hate speech online + (revised). Brussels: Digital Services Act. Available at: <https://digital-strategy.ec.europa.eu/en/library/code-conduct-countering-illegal-hate-speech-online> (Accessed May 26, 2025).
- Farkas, J., Schou, J., and Neumayer, C. (2018). Cloaked Facebook pages: exploring fake Islamist propaganda in social media. *New Media Soc.* 20, 1850–1867. doi: 10.1177/1461444817707759
- Fraser, G. (2025). "Huge problems' with Instagram and Facebook changes, says oversight board" in *BBC*, 8 January. Available at: <https://www.bbc.com/news/articles/cjwlwlpwx7o> (Accessed May 26, 2025).
- Gilgun, J. F. (2005). Qualitative research and family psychology. *J. Fam. Psychol.* 19, 40–50. doi: 10.1037/0893-3200.19.1.40
- Gilgun, J. F. (2009). "Methods for enhancing theory and knowledge about problems, policies, and practice" In *The Sage Handbook of Social Work Research*. eds. K. Briar, J. Orme, R. Ruckdeschel and I. Shaw (Thousand Oaks, CA), 281–297.
- Gilgun, J. F. (2015). *Deductive qualitative analysis as middle ground: Theory-guided qualitative research*. Seattle, WA, USA: Amazon Digital Services LLC.
- Gkritsi, E., and Grier, M. (2025). "EU lawmakers try to counter tech bros' Washington lobbying" in *PoliticoEurope* 3 March. Available at: <https://www.politico.eu/article/eu-lawmakers-counter-tech-bro-washington-lobby-mark-zuckerberg-elon-musk/>
- GLAAD. (2025). GLAAD response: META and Mark Zuckerberg remove long-standing anti-lgbtq hate speech policies after announcing end of fact-checking program. Available at: <https://glaad.org/releases/meta-removes-anti-lgbtq-hate-speech-policies-ends-fact-checking/> (Accessed May 26, 2025).
- Glaser, B. G., and Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Gorenc, N. (2022). Hate speech or free speech: an ethical dilemma? *Int. Rev. Sociol.* 32, 413–425. doi: 10.1080/03906701.2022.2133406
- Govil, N., and Baishya, A. K. (2018). The bully in the pulpit: autocracy, digital social media and right-wing populist technoculture. *Commun. Cult. Critique* 11, 67–84. doi: 10.1093/ccc/tcx001
- Grasswick, H. (2016). "Strong objectivity" in *The Wiley Blackwell encyclopedia of gender and sexuality studies*. ed. N. A. Naples (New Jersey: John Wiley & Sons).
- Harding, S. (1986). *The science question in feminism*. Cornell: Cornell University Press.
- Harding, S. (1991). *Whose science? Whose knowledge? Thinking from women's lives*. Ithaca, NY: Cornell University Press.
- Harsin, J. (2018). *Post-truth and critical communication studies*. Oxford research encyclopedia of communication.
- Hartsock, N. C. M. (1983). "The feminist standpoint: developing the ground for a specifically feminist historical materialism" in *Discovering reality*. Synthese Library, vol 161. eds. S. Harding and M. B. Hintikka (Dordrecht: Springer).
- Hay, G., and Kwok, K. (2025). "Meta's Brussels bark may be worse than its bite" in *Reuters*, 8 January. Available at: <https://www.reuters.com/breakingviews/metass-brussels-bark-may-be-worse-than-its-bite-2025-01-08/> (Accessed May 26, 2025).
- Hendrix, J. (2025). "Transcript: mark Zuckerberg announces major changes to Meta's content moderation policies and operations", *Tech Policy Press*, 7 January. Available at: <https://www.techpolicy.press/transcript-mark-zuckerberg-announces-major-changes-to-metas-content-moderation-policies-and-operations/> (Accessed May 26, 2025).
- Herzog, C., Handke, C., and Hitters, E. (2019). "Analyzing Talk and Text II: Thematic Analysis" in *The Palgrave Handbook of Methods for Media Policy Research*. eds. Van den Bulck, M. Puppis, K. Donders and L. Audenhove. (Cham, Switzerland), 385–401.
- hooks, B. (1989). *Talking back: Thinking feminist, thinking black*. Boston: South End Press.
- Hughes, L. D., Kidd, K. M., Gamarel, K. E., Operario, D., and Dowshen, N. (2021). These laws will be devastating: provider perspectives on legislation banning gender-affirming care for transgender adolescents. *J. Adolesc. Health* 69, 976–982. doi: 10.1016/j.jadohealth.2021.08.020
- Hyde, K. F. (2000). Recognising deductive processes in qualitative research. *Qualit. Market Res.* 3, 82–90. doi: 10.1108/13522750010322089
- Johnson, A. H. (2022). Tipping points and shifting expectations: the promise of applied trans studies for building structural competency. *Bullet. Appl. Transg. Stud.* 1, 163–177. doi: 10.57814/98we-xm94
- Karppinen, K. E. (2019). Freedom without idealization: non-ideal approaches to freedom of communication. *Commun. Theory* 29, 66–85. doi: 10.1093/ct/qty013
- Klein, A. (2012). Slipping racism into the mainstream: a theory of information laundering. *Commun. Theory* 22, 427–448. doi: 10.1111/j.1468-2885.2012.01415.x
- Knibbs, K. (2025). "Meta Now Lets Users Say Gay and Trans People Have 'Mental Illness'" in *Wired*, 7 January. Available at: <https://www.wired.com/story/meta-immigration-gender-policies-change/> (Accessed May 26, 2025).
- Koyama, E. (2020). Whose feminism is it anyway? The unspoken racism of the trans inclusion debate. *Sociol. Rev.* 68, 735–744. doi: 10.1177/0038026120934685
- Körömi, C., Haec, P., and Cheslow, D. (2025). "Zuck goes full Musk, dumps Facebook fact-checking program" in *Politico*, 7 January. Available at: <https://www.politico.eu/article/mark-zuckerberg-full-elon-musk-dump-facebook-fact-checker/> (Accessed May 26, 2025).

- Kus, O. (2016). Understanding digital hate speech: analysis of user comments on BBC world service Facebook page via text mining in context of Syrian refugee crisis case. *Istanbul Üniv. İletişim Fak. Derg.* 2016, 97–121. doi: 10.17064/iuifd.289373
- Laaksonen, S. M., Haapoja, J., Kinnunen, T., Nelimarkka, M., and Pöyhtäri, R. (2020). The Datafication of hate: expectations and challenges in automated hate speech monitoring. *Front. Big Data* 3:3. doi: 10.3389/fdata.2020.00003
- Lepore, C., Alstott, A., and McNamara, M. (2022). Scientific misinformation is criminalizing the standard of care for transgender youth. *JAMA Pediatr.* 176, 965–966. doi: 10.1001/jamapediatrics.2022.2959
- Lillian, D. (2007). A thorn by any other name: sexist discourse as hate speech. *Discourse Soc.* 18, 719–740. doi: 10.1177/0957926507082193
- Magalhães, J. C. (2025). “Meta is not returning to its free speech origins – it’s preparing for an autocratic future”, Media@LSE blog, 10 January. Available at: <https://blogs.lse.ac.uk/media/2025/01/10/meta-is-not-returning-to-its-free-speech-origins-its-preparing-for-an-autocratic-future/> (Accessed May 26, 2025).
- Malingre, V. (2025). “European digital regulation comes under attack from Trump, Musk and Zuckerberg” in LeMonde, 13 January. Available at: https://www.lemonde.fr/en/economy/article/2025/01/13/european-digital-regulation-comes-under-attack-from-trump-musk-and-zuckerberg_6737001_19.html (Accessed May 26, 2025).
- Matamoros-Fernández, A., and Farkas, J. (2021). Racism, hate speech, and social media: a systematic review and critique. *Telev. New Media* 22, 205–224. doi: 10.1177/1527476420982230
- Milmo, D. (2025). “Meta is ushering in a ‘world without facts’, says Nobel peace prize winner” in Guardian, 8 January. Available at: <https://www.theguardian.com/world/2025/jan/08/facebook-end-factchecking-nobel-peace-prize-winner-maria-ressa> (Accessed May 26, 2025).
- Mouffe, C. (2000). *Democratic paradox*. London, UK: Verso.
- Mouffe, C. (2005). *On the political*. London, UK: Routledge.
- Mršević, Z. (2013). Homophobia in Serbia and LGBT rights. *Southeastern Europe* 37, 60–87. doi: 10.1163/18763332-03701004
- Murthy, D., and Sharma, S. (2019). Visualizing YouTube’s comment space: online hostility as a networked phenomena. *New Media Soc.* 21, 191–213. doi: 10.1177/1461444818792393
- Naples, N. A. (2015). “Strong objectivity” in *The Blackwell Encyclopedia of Sociology*, ed. G. Ritzer (New Jersey: John Wiley & Sons).
- Nikunen, K. (2023). “Affective temporalities of digital hate cultures” in *Affective formation of publics*, eds. M. Lünenborg and B. Röttger-Rössler (London: Routledge).
- Nikunen, K., Haara, P., Kosonen, H., Knuutila, A., Pöyhtäri, R., and Saresma, T. (2025). Contested meanings of hate speech and the post-truth condition on digital platforms. *Soc. Media Soc.* 11:1794. doi: 10.1177/20563051251341794
- Nikunen, K., Hokka, J., and Nelimarkka, M. (2021). Affective practice of soldiering: how sharing images is used to spread extremist and racist ethos on soldiers of Odin Facebook site. *Telev. New Media* 22, 166–185. doi: 10.1177/1527476420982235
- Olson, J. (2025). “This is Not Normal: Meta’s Anti-LGBTQ Makeover” in *TechPolicy Press*, 7 February. Available at: <https://www.techpolicy.press/this-is-not-normal-metas-anti-lgbtq-makeover/> (Accessed May 26, 2025).
- O’Reagan, C. (2018). Hate speech regulation on social media: (an intractable) contemporary challenge. *Curr. Leg. Probl.* 71, 403–429. doi: 10.1093/clp/cuy012
- O’Reagan, C., and Theil, S. (2020). Hate speech regulation on social media: an intractable contemporary challenge. *researchoutreach*, 26 February. Available at: <https://researchoutreach.org/articles/hate-speech-regulation-social-media-intractable-contemporary-challenge/> (Accessed May 26, 2025).
- Ott, B. (2017). The age of twitter: Donald J. Trump and the politics of debasement. *Crit. Stud. Media Commun.* 34, 59–68. doi: 10.1080/15295036.2016.1266686
- Pape, M. (2022). Feminism, trans justice, and speech rights: a comparative perspective. *Law Contemp. Soc. Probl.* 85, 215–240.
- Park, B. C., Das, R. K., and Drolet, B. C. (2021). Increasing criminalization of gender-affirming care for transgender youths—a politically motivated crisis. *JAMA Pediatr.* 175, 1205–1206. doi: 10.1001/jamapediatrics.2021.2969
- Patel, P., Wemyss, G., and Assiter, A. (2022). Why free speech? *Feminist Dissent* 6, 1–38. doi: 10.31273/fd.n6.2022.1261
- Paz, M. A., Montero-Díaz, J., and Moreno-Delgado, A. (2020). Hate speech: a systematized review. *SAGE Open* 10:3022. doi: 10.1177/2158244020973022
- Piquard, A. (2025). “Mark Zuckerberg wants more ‘masculine energy’ and less diversity policy”, LeMonde, 12 January. Available at: https://www.lemonde.fr/en/economy/article/2025/01/12/mark-zuckerberg-wants-more-masculine-energy-and-less-diversity-policy_6736961_19.html (Accessed May 26, 2025).
- Pohjonen, M., and Udupa, S. (2017). Extreme speech online: an anthropological critique of hate speech debates. *Int. J. Commun.* 11, 1173–1191.
- Pöyhtäri, R., Nelimarkka, M., Nikunen, K., Ojala, M., Pantti, M., and Pääkkönen, J. (2021). Refugee debate and networked framing in the hybrid media environment. *Int. Commun. Gaz.* 83, 81–102. doi: 10.1177/1748048519883520
- Pradel, F. (2024). Influence of hate speech about refugees in search algorithms on political attitudes: an online experiment. *New Media Soc.* 2024:4735. doi: 10.1177/14614448241244735
- Pukallus, S., and Arthur, C. (2024). Combating hate speech on social media: applying targeted regulation, developing civil-communicative skills and utilising local evidence-based anti-hate speech interventions. *J. Med.* 5, 467–484. doi: 10.3390/journalmedia5020031
- Reddi, M., Kuo, R., and Kreiss, D. (2021). Identity propaganda: racial narratives and disinformation. *New Media Soc.* 25, 2201–2218. doi: 10.1177/14614448211029293
- Reich, R., Mehran, S., and Weinstein, J. (2023). *System error: Where big tech went wrong and how we can reboot*. London: Hodder Paperbacks.
- Rosenblat, M. O. (2025). “How Meta Turned Its Back on Human Rights” in *Tech Policy Press*, 27 January. Available at: <https://www.techpolicy.press/how-meta-turned-its-back-on-human-rights/> (Accessed May 26, 2025).
- Sánchez-Sánchez, A. M., Ruiz-Muñoz, D., and Sánchez-Sánchez, F. J. (2024). Mapping homophobia and transphobia on social media. *Sex. Res. Soc. Policy.* 21, 210–226. doi: 10.1007/s13178-023-00879-z
- Santos Fernández, F. J. (2024). Online homophobia: hate speech and conspiracy theories towards LGBTQI+ people on twitter in Spain. *Cult. Stud. Soc. CuSoc* 9, 39–56.
- Schmid, U. K., Kümpel, A. S., and Rieger, D. (2022). How social media users perceive different forms of online hate speech: a qualitative multi-method study. *New Media Soc.* 26, 2614–2632. doi: 10.1177/14614448221091185
- Siapera, E. (2019). Organised and ambient digital racism: multidirectional flows in the Irish digital sphere. *Open Libr. Human.* 5, 1–34. doi: 10.16995/olh.405
- Sorce, T., and Thomas, T. (2025). *New digital feminist interventions: Speaking up, talking Back*. London & New York: Routledge.
- Sponholz, L., and Christofolletti, R. (2018). From preachers to comedians: ideal types of hate speakers in Brazil. *Glob. Media Commun.* 15, 67–84. doi: 10.1177/1742766518818870
- Strossen, N. (2018). *Hate: Why we should resist it with free speech, not censorship*. Oxford: Oxford University Press.
- Swedberg, R. (2020). “Exploratory research” in *The production of knowledge: Enhancing Progress in social science*, eds. C. Elman, J. Gerring and J. Mahoney (Cambridge: Cambridge University Press), 17–41.
- Titley, G. (2019). *Racism and media*. London: Sage.
- Tesis, A. (2002). *Destructive messages: How hate speech paves the way for harmful social movements*. New York: New York University Press.
- Udupa, S., Gliagladone, I., and Hervik, P. (2021). *Digital hate: The global conjuncture of extreme speech*. Indiana: Indiana University Press.
- Udupa, S., and Pohjonen, M. (2019). Extreme speech and global digital cultures—introduction. *Int. J. Commun.* 13, 3049–3067.
- United Nations (2019). *United Nations strategy and plan of action on hate speech*. New York: United Nations Strategy and Plan of Action on Hate Speech.
- Vax, E., and Haeck, P. (2025). “Zuckerberg’s censorship claims were ‘misleading’ -EU tech chief” in Politico, 10 January. Available at: <https://www.politico.eu/article/mark-zuckerberg-meta-misleading-censorship-henna-virkunen/> (Accessed May 26, 2025).
- Wintour, P. (2025). “JD Vance stuns Munich conference with blistering attack on Europe’s leaders”, Guardian, 14 February. Available at: <https://www.theguardian.com/us-news/2025/feb/14/jd-vance-stuns-munich-conference-with-blistering-attack-on-europes-leaders> (Accessed May 26, 2025).
- Wood, J. T. (2009). “Feminist standpoint theory” in eds. Littlejohn, S. W. and Foss, K. A. *Encyclopedia of communication theory* (Thousand Oaks, CA: SAGE), 397–399.
- Yamaguchi, T. (2013). Xenophobia in action: Ultrationalism, hate speech, and the internet in Japan. *Radic. Hist. Rev.* 2013, 98–118. doi: 10.1215/01636545-2210617
- Yang, M. (2025). “Google and Microsoft donate \$1m each to trump’s inaugural fund”, Guardian, 9 January. Available at: <https://www.theguardian.com/technology/2025/jan/09/google-microsoft-donate-trump-inaugural-fund> (Accessed May 26, 2025).
- Young, I. M. (1990). *Justice and the politics of difference*. Princeton: Princeton University Press.
- Zuckerberg, M. (2025). “It’s time to get back to our roots around free expression. We’re replacing fact checkers with community notes, simplifying our policies and focusing on reducing mistakes. Looking forward to this next chapter” [Facebook video] 7 January 2025. Available at: <https://www.facebook.com/zuck/videos/its-time-to-get-back-to-our-roots-around-free-expression-were-replacing-fact-checkers/> (Accessed May 26, 2025).