Review Paper

HATE SPEECH ON THE INTERNET

UDC 343.41:004.738.4
342.727:004.738.4

Miomira Kostić¹, Vida Vilić²

¹Faculty of Law, University of Niš, Republic of Serbia
²Legal Affairs Office, the Dentistry Clinic in Niš, Republic of Serbia

Abstract. In addition to the benefits offered by the Internet and social networks, there has been an increase of cyberspace abuses related to certain groups of people belonging to particular age, gender, sex, national or religious groups. By virtue of social networks, these individuals and groups are exposed to different forms of harassment, insults, calls for lynching (etc.), all of which amount to hate speech. The Universal Declaration of Human Rights recognizes the freedom of expression as a human right but the freedom of expression on the Internet is subject to certain limitations because the Internet is not void of different forms of deviant behavior (such as hate speech), which incite discrimination, hatred or violence against groups or individuals on the basis of their personal characteristics. In Serbia, the relevant legal provisions for combating hate speech are embodied in the Serbian Constitution and a number of legislative acts: the Anti-Discrimination Act, the Public Information Act, the Act on the Prevention of Discrimination against Persons with Disabilities, and the Criminal Code of the Republic of Serbia. The most important factors in preventive action are: a detailed normative regulation of this subject matter, detection of all forms of Internet abuse and relevant sanctions for these offences, detection of contents published on social networks which are most frequently used to incite and disseminate hate speech, raising public awareness, ensuring constant monitoring and control by network administrators and users.

Key words: Internet, freedom of expression, hate speech, prevention.

1. INTRODUCTION

The use of information and communication technologies has spread to all areas of human life, work, daily entertainment and many other private and business activities, as a result of which almost everything in today's society may be done online, starting from signing contracts to committing different criminal offences. The use of the Internet and
other related technologies has transformed three specific areas of our society: privacy, freedom of expression and the freedom of information exchange. Social changes related to information and communication technologies in a modern society have influenced the perception of a new phenomenon: the information society. Information becomes an important element of freedom of expression and freedom of information exchange, which to a large extent depends on the legitimacy and possibility of database collection management (Catanese, De Meo, Ferrara, Fiumara, Provett, 2011).

In the past, the virtual space was full of interesting and useful information, but there was very little possibility for interactive communication and active data creation. The Internet has become especially popular amongst the younger generations who have been brought up "computer literate” and are therefore not discouraged by any technicalities associated with it (Nazaryan, Gridchin, 2006). In addition to the benefits offered by the Internet and social networks, there has been an increase in abuse-related acts in the virtual space, especially in the sphere of the users’ right to privacy. Certain groups of people are particularly susceptible to computer-related abuse of privacy (based on age, gender, sexual orientation, nationality or religious belief) and exposed to different forms of harassment, insult, calls for lynching (etc), all of which constitute hate speech.

2. FREEDOM OF SPEECH AND HATE SPEECH

The Internet and social networks are defined as media exempted from any influence of the state and traditional norms. Equal opportunities of access to the Internet and social networks as well as the freedom of speech apply to all users, which means that the users themselves create content using an unlimited number of sources available. Thus, the Internet becomes a platform for communication and exchange of ideas, social and cultural values, political beliefs, expressing opinions and attitudes, and promoting civil activism.

Freedom of expression, which existed before the advent of the Internet, is recognized as a human right in Article 19 of the Universal Declaration of Human Rights. It is proportional to the level of democracy in one society and closely linked to freedom of the media (including the Internet). The basic postulate of a democratic society is to establish a balance between the right to freedom of expression and other guaranteed rights, such as the freedom of thought, conscience and religion or the anti-discrimination right. Consequently, the freedom of expression on the Internet has certain limitations because the Internet is not void of different forms of deviant behavior (including hate speech), which incite discrimination, hatred or violence against groups or individuals on the basis of their personal characteristics. In this way, the society strengthens its defence mechanisms against phenomena that can jeopardise an individual, certain group or a community in global (Veljanovski, 2012).

Hate speech can be defined as any form of expression that incites, promotes or justifies intolerance, discrimination and hostility against a member of another race, nationality, religion, gender, sexual orientation, ethnic origin and some other personal characteristic of an individual or a group.1 The goal of using hate speech on the Internet is to offend another person or a particular group by referring to characteristics of an individual or group

---

members, and it is aimed at causing some negative consequences and marginalizing the specific person or group.

Hate speech may be demonstrated by means of generating contempt for a particular person/group; creating negative stereotypes about a particular person/group; inciting discrimination and hostility; expressing social condemnation or disapproval of a particular person/group; causing feelings of insecurity and fear in a particular person or group members; inflicting physical and psychological pain on an individual/group member; threatening a specific person/group; encouraging and inciting violence against a person/group; creating a public opinion that specific behavior is socially undesirable, unjustified and thus intolerable.2

Hate speech is frequently just an “introduction” to some other form of violence or for some other criminal offence. For example, humiliating gender-based comments, sexual remarks, or dirty jokes published on the Internet and social networks can be considered as a form of hate speech and can lead to sexual harassment and stalking of the targeted victim (Vilić, 2013).

3. WHICH CONDUCTS MAY BE QUALIFIED AS A HATE SPEECH?

Although the term "hate speech" came into wide use in the late 1980s and during 1990s, there is no universally accepted definition of hate speech. In criminological literature, "hate speech" is designated to include statements that intimidate, insult or harass individuals/groups, and/or statements that incite violence, hatred or discrimination against individuals/groups.

Given the absence of a precise definition of hate speech, we may refer to the judicial practice (case law) of the European Court of Human Rights, where the term "hate speech" is used to describe forms of expression which spread, incite, promote or justify hatred based on intolerance, including religious intolerance. Hate speech, thus, involves incitement to racial hatred and hatred based on religion (including hatred based on the distinction between believers and non-believers), as well as incitement to other forms of hatred that generate intolerance, aggressive nationalism and ethnocentrism. Homophobic speech also falls into the scope of hate speech. Hate speech and gender intolerance (mostly based on sexual orientation, intolerance towards different political opinions and national or social status) have widely expanded through the Internet and social networks.

The second widely used definition of hate speech was given by the Committee of Ministers of the Council of Europe in 1997. According to this definition, "hate speech includes all forms of expression which spread, incite, promote or justify interracial hatred, xenophobia, anti-Semitism or hatred based on intolerance, including intolerance expressed by aggressive nationalism and ethnocentrism, discrimination or animosity towards minorities, immigrants or people of immigrant origin".3 Hate speech can be manifested in the public expression of discriminatory attitudes by writing graffiti, posting messages or symbols of discriminatory content, attending public assemblies, sporting events and some other massive public events.

Hate speech implies the use of words, phrases and sentences including offensive and discriminatory content about an specific individual or people who belong to a particular

---

2 Ibid.

3 Newspaper article: Sloboda na internetu i govor mržnje online: medijska politika i internet u BIH (Freedom on the Internet and hate speech online: the policy of media and the Internet in Bosnia and Herzegovina), Sarajevo, 2014, retrieved from: internews.ba/.../govor%20mrznje%20/2012/02.12.2015.
race, nationality, religion, ideology, sexual orientation (etc.). Hate speech is aimed at stigmatizing, labelling, humiliating, defaming, degrading, harming or ridiculing a particular individual, social group or group member. In defining hate speech, a distinction has to be made between three types of hate speech: a) **direct** hate speech (spreading false allegations, statements and accusations, the use of profanity and vulgar language); b) **indirect** hate speech (spreading false conclusion with reference to specific sources and authorities, false-friends’ advice given in a quasi-friendly manner, statements imitating journalists’ reporting style, the use of slang/jargon and anti-proverbs); and c) **latent hate speech** (the use of non-explicit statements with a hidden message or ambiguous meaning, witty retorts and puns).

4. INTERNATIONAL AND NATIONAL DOCUMENTS PROHIBITING HATE SPEECH

At the international level, the first international document on hate speech was the International Convention on the Elimination of All Forms of Racial Discrimination,\(^4\) adopted by the UN General Assembly in 1965. The International Covenant on Civil and Political Rights (ICCPR)\(^5\), which entered into force in 1976, envisages that any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law. The definition of hate speech is also provided in Recommendation No. R (97) 20 of the Committee of Ministers of the Member States (of 30.9.1997.): the term "hate speech" shall be understood as covering all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, including: intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants and people of immigrant origin.\(^6\)

At the regional level, the European Convention on Human Rights and the American Convention on Human Rights guarantee the freedom of expression but do not impose a specific prohibition of hate speech. The Optional Protocol to the CoE Convention on Cybercrime obliges member states to adopt domestic legislation and other relevant measures to prohibit online hate speech and classify it as a criminal offense.

In the Republic of Serbia, the legal provisions relevant for combating hate speech are embodied in the Serbian Constitution and in a number of other legislative acts: the Anti-Discrimination Act, the Public Information Act, the Act on the Prevention of Discrimination against Persons with Disabilities, and the Criminal Code of the Republic of Serbia.

The Serbian Constitution guarantees the freedom of expression and prescribes the conditions under which freedom of expression can be limited (Art. 46). The Constitution also guarantees the freedom of the media (Art. 50) and prescribes the conditions under which a competent court may prevent the dissemination of information and ideas through the mass media. The Constitution also prohibits any incitement of racial, ethnic, religious or other

---


\(^5\) Međunarodni pakt o građanskim i političkim pravima (International Covenant on Civil and Political Rights - ICCPR), Official Gazette SFRJ, 7/1971.

inequality, hatred and intolerance (Art. 49), as well as any other direct or indirect form of discrimination, based on race, sex, national origin, social origin, birth, religion, political or other opinion, property status, culture, language, age, mental or physical disability (Art. 21, §3).

The Serbian Anti-Discrimination Act defines hate speech as the expression of ideas, information and opinions that incite discrimination, hatred or violence against persons or group of persons based on their personal characteristics, in the media and other publications, at public gatherings and places accessible to the public, writing and displaying messages or symbols and otherwise (Art. 11). Hate speech in the media is prohibited by the Public Information Act (Art. 38), which explicitly bans publishing ideas, information and opinions that incite discrimination, hatred or violence against persons or a group of persons based on their race, religion, nation, ethnic group, gender or their sexual orientation, regardless of whether the act of publication constitutes a criminal offense. The Act on the Prevention of Discrimination against Persons with Disabilities stipulates that it is forbidden to publish and display in public places or in any other way disseminate messages and symbols calling for discriminatory treatment of persons with disabilities (Art. 10).

In the Criminal Code of the Republic of Serbia, hate speech on the Internet and social networks is not envisaged as a specific criminal offense. In Chapter XVII of the Criminal Code, dealing with “Criminal offenses against honor and reputation”, there is an offense of “Causing harm to one’s reputation on the basis of racial, religious, ethnic or other affiliation” (Art. 174); this offence has been designated as an act of public humiliation and ridicule of persons or groups based on race, color, religion, nationality, ethnic origin or other characteristics. Chapter XXVII of the Criminal Code, dealing with “Criminal offenses against the constitutional order and security of the Republic of Serbia”, contains a criminal offense of “Instigating national, racial and religious hatred and intolerance” (Art. 317), which provides sanctions for everyone who instigates national, racial or religious hatred or intolerance among nations and ethnic communities living in Serbia. The qualified (aggravated) form of this offense exists when the offense has been committed by coercion, abuse, compromising or endangering security, ridiculing national, ethnic or religious symbols, damaging other people’s property, desecrating monuments, memorials or graves. The most severe form of this offence exist when the act is committed by abusing one’s official position or authority, or in case where the commission of these offences has resulted in civil revolt, riots, violence or other grave consequences detrimental for the co-existence of nations, national minorities or ethnic groups in Serbia.

In the Republic of Slovenia, the Slovenian Parliament voted in January 2016 to amend the Media Act, in the provision related to editing and sanctioning Internet comments recognized as hate speech; the provision envisages that editors of different media and social networks shall be held criminally liable for publishing hate speech. Namely, if the editor does not promptly remove or delete comments that are considered, pursuant to the Constitution and relevant laws of the Republic of Slovenia, to be encouraging intolerance and hostility based on religious, racial, ethnic or other grounds, the specific medium or social network may be punished by a fine ranging from 500 to 5,000 Euro.7

---

In the Republic of Croatia, young people gathered around the Croatian Youth Network, launched an interesting initiative which was articulated in the Position Paper of the Croatian Youth Network "Hate speech on the Internet and other forms of online violence" (2016). This document deals with the problem of hate speech as a form of violence that occurs among young people when using the Internet. Hate speech on the Internet is defined as "the use of statements arising from stereotypes and prejudice in various places on the Internet, which are insulting and/or humiliating to a specific individual and/or group." In addition to defining the legal origin of hate speech by providing an overview of relevant international legal documents, the Position Paper highlights the consequences of using hate speech on the Internet (such as: emotional, psychological and social effect, identity crisis, etc.). Moreover, the document includes recommendations of the Croatian Youth Network for the prevention of hate speech on the Internet. The recommendations are aimed at: a) competent authorities/institutions (which have to create an adequate legal framework to define violence and hate speech on the Internet, to introduce civic education in schools and educate young people to promote non-violence; to raise media awareness about different forms of hate speech, to develop a victim support system in order to mitigate the consequences of Internet violence, etc.); b) educational institutions (which have to organize professional teacher trainings on Internet violence, non-violence programs and human rights, to consider introducing preventive measures among young people, to develop learning activities promoting anti-discrimination, etc.); and c) the civil society organizations (which have to engage in public policy-making processes in order to ensure a safer cyberspace, to raise awareness of the media and general public about the consequences of Internet violence and possible prevention, to develop cooperation with other organizations, etc.).

The Federation of Bosnia and Herzegovina is a signatory to all international conventions on human rights, which are also incorporated into Annex IV of the Dayton Agreement. In addition to the provisions of the Constitution of Bosnia and Herzegovina, hate speech is also regulated in Article 363 of the Criminal Code of the Federation of Bosnia and Herzegovina and in Article 357 of the Criminal Code of the Brčko District. These documents prescribe a fine or a term of imprisonment not exceeding 3 years for anyone who, in violation of the media standards and rules of professional journalism, uses hate speech or any speech that directly calls for or incites violence, ethnic or national confrontation that constitutes a threat to the public order and peace. This provision is indirectly linked to the use of hate speech on the Internet.

---

8 Pozicijski dokument Mreže mladih Hrvatske „Govor mržnje na Internetu i ostali oblici online nasilja“ (Position Paper of Croatian Youth Network "Hate speech on the Internet and other forms of online violence"), retrieved from http://www.mmh.hr/files/ckfinder/files/Pozicijski_GovorMr%C5%BEnjeNaInternetu.pdf, 04.02.2016.
9 Ibid.
10 Ibid.
11 Newspaper article: Sloboda na internetu i govor mržnje online: medijska politika i internet u BIH (Freedom on the Internet and hate speech online: the media policy and the Internet in Bosnia and Herzegovina), 2014; retrieved from http://internews.ba/sites/default/files/resursi/Govor%20mrznje%20na%20internetu.pdf, 01.02.2016.
12 Article 363, the Criminal Code of the Federation of Bosnia and Herzegovina, Official Gazette of Bosnia and Herzegovina, 3/03, 32/03, 37/03, 54/04, 61/04, 30/05, 53/06, 55/06, 32/07, 8/10, 47/14, 22/15, 40/15
13 Article 357, the Criminal Code of the Brčko District, Official Gazette of Bosnia and Herzegovina, 10/03, 6/05, 21/10, 47/11, 52/11, 9/13
5. POSSIBLE CONSEQUENCES OF HATE SPEECH ON THE INTERNET

Hate speech has certain negative effects which include causing fear, anxiety, embarrassment or anger. Specific consequences may be reflected in generating contempt, negative stereotypes against a particular person or group, inciting discrimination and hatred, condemnation, causing feelings of insecurity and fear among group members, creating a general feeling that such discriminatory behavior against individuals or members of specific groups is socially desirable and justified, or that it may be tolerated without being sanctioned.

The Internet and social networks are very convenient media for inciting national, racial and religious hatred. Online insults, ridicule, humiliating and degrading comments, underestimation of national, racial or religious feelings and other forms of expression of hate speech on the Internet websites or blogs are rapidly disseminated to an unlimited number of people, which contributes to inciting and compounding hatred. The perpetrator may be an individual but, more commonly, these activities are performed by groups and organizations whose main goal is to incite and enhance national, racial and religious hatred and intolerance. One of the most popular Internet forums for inciting hate speech is "Stormfront", an internet forum of the white nationalists' community advocating the white supremacy theory (Nikolić-Komlen, Gvozdenović, Radulović, Milosavljević, Jerković, Živković, Živanović, Reljanović, Aleksić, 2010: 127). In 2012, the Higher Court in Belgrade rendered a judicial decision sentencing the defendant Sima Vladičić to 3 months' imprisonment (suspended for two years) for committing the criminal offence of endangering the safety of another, by posting threats in the Facebook group "500,000 Serbs against gay parade" in order to intimidate members of the LGBT population. This judgment is thought to be the first judicial decision in Serbia where a criminal sanction has been issued for hatred-motivated threats.

In 2015, the Center for New Media "Liber" conducted a survey on hate speech in order to provide for a better understanding of hate speech, particularly in terms of the online discussions among Tweeter users from Serbia during the so-called "Gun Affair"; the main actor and the driving force in the discussions was the Editor of the daily newspaper "Informer" Dragan Vučićević and the target was the Serbian Ombudsman Saša Janković. Considering the Tweeter users' language and attitude towards Saša Janković, the survey showed that the Ombudsman was a victim of organized cyber abuse. The survey, which included a poll of 6,034 original tweets, provided an opportunity to examine the public opinion and public activity on a specific subject, to understand the importance of action and reaction in a dialogue, to study the genesis and development of a digital public discourse and to analyse the role and impact of insult and direct/indirect hate speech in discussions on burning social issues, when the public is highly motivated to express opinions on a sensitive or controversial issue. In the period from May 1st to May

---

21st 2015, a total of 2,609 Twitter users sent their tweets on this specific topic. The prevailing form of hate speech that refers to Saša Janković is direct hate speech, which included making accusations and false allegations, presenting unsubstantiated or unverified statements as facts, presenting qualifications as the ultimate truth, using profanity and vulgar language, drawing quasi-conclusions by referring to a specific source or authority.

At this point, it seems to be worth noting that the German Ministry of Justice has made significant progress in countering online hate speech by making an agreement with Internet companies such as Facebook, Google and Twitter. According to this agreement, these companies are obliged to remove any form of hate speech observed on these social networks within a period of 24 hours. In the joint statement officially published by the Facebook, Google and Twitter managements, they stated that it would facilitate the identification and targeting of hate speech by users and anti-racist groups. The content of Internet portals and posts published on social networks will be closely monitored and examined by teams of experts (network administrators), who will be authorized to remove the contents including hate speech from their websites within 24 hours.17

5. CONCLUSION

Cybercrime is a new form of crime which has emerged as a result of huge technological progress in telecommunications. From the standpoint of social development, the increasing use of the Internet and social networks, as well as the use of computer technology in everyday life, bring substantial advancement in all aspects of social life. On the other hand, the use of computer technology, particularly on the Internet and social networks, exposes a large number of users to the lurking threat of daily victimization in case the data transmitted via the Internet and posted on social networks is abused or misused.

The wide-spread use of computers in everyday life by people of diverse social backgrounds as well as the liberal approach to global communication and information networks have contributed to the rise of cybercrime, which has become a serious and imminent threat in the contemporary society. It is a reality of today’s life rather than futuristic science fiction. Hate speech on the Internet is only one aspect of abusing and misusing the cyberspace where billions of people communicate in the infinite virtual world.

In order to reduce this form of abuse/misuse on the Internet and social networks, and to prevent the violation of their users’ right to privacy, it is necessary to establish relevant normative framework and legal mechanisms for the detection and sanctioning of these socially unacceptable forms of criminal behavior. It is also highly important to report all cybercrime offences to competent authorities, which is essential in terms of reducing the “dark figure” of crime, accomplishing better results in the identification, prevention and monitoring of these illegal acts, and overcoming the topical problem of underreporting these criminal offences.

In the Republic of Serbia, the legal provisions relevant to combating hate speech are embodied in the Serbian Constitution and a number of legislative acts: the Anti-Discrimination Act, the Public Information Act, the Act on the Prevention of Discrimination of Persons with

---

Hate Speech on the Internet

Disabilities, and the Criminal Code of the Republic of Serbia. The most important factors of preventive action are: a detailed regulation of this subject matter; detection of all forms of Internet abuse and prescribing relevant sanctions for these offences; detection of contents published on social networks which are most frequently used to incite and disseminate hate speech; raising public awareness about these issues; ensuring constant monitoring and control by network administrators and users.

However, given the fact that hate speech on the Internet has become a transnational problem with consequences reaching far beyond the territory of a single state, it is clear that the mechanisms for combating this type of crime may not focus only on national criminal legislation but also on taking appropriate technical, structural and educational measures, implementing international technical and legal instruments, as well as raising awareness about the information that may generate potential risk and commission of cybercrime.

REFERENCES

9. Nemačka uklanjanje govor mržnje sa Interneta [Newspaper article: Germany removes hate speech from the Internet], Radio televizija Srbije, retrieved 22 December 2015, from www.rtss.rs/.../Nemacka+uklanja+govor+mrzwe+sa+intermeta.html
11. Pozicijski dokument Mreže mladih Hrvatske „Govor mržnje na Internetu i ostali oblici online nasilja“ [Position paper of Croatian Youth Network "Hate speech on the Internet and other forms of online violence"], retrieved 04 February, 2016, from http://www.mnh.hr/files/ckfinder/files/Pozicijski_Govor_Mr%C5%BEenjaNaInternetu.pdf
GOVOR MRŽNJE NA INTERNETU

Korišćenje informaciono komunikacionih tehnologija zahvatilo je sva područja života ljudi, njihovog rada, zabave i brojnih drugih privatnih i poslovnih aktivnosti, tako da je skoro sve u društvu postalo on line, počev od potpisivanja ugovora do vršenja krivičnih dela. Pored prednosti koje internet i društvene mreže pružaju, zabeležen je porast zloupotreba vezanih za virutelni prostor, posebno u sferi prava na privatnost. Kompjuterskoj zloupotrebi privatnosti naročito su izložene određene grupe ljudi koje pripadaju posebnim starosnim, polnim, seksualnim, nacionalnim ili verskim grupama i koje su putem društvenih mreža izložene šikaniranju, vređanju, pozivanju na linč – jednom rečju, govoru mržnje.

Sloboda izražavanja je Univerzalnom deklaracijom o ljudskim pravima priznata kao ljudsko pravo, ali na internetu ona ima određena ograničenja jer podrazumeva da je internet slobodan od svih devijantnih ponašanja, među kojima je govor mržnje, kojim se podstiče diskriminacija, mržnja ili nasilje protiv lica ili grupa lica zbog njihovih ličnih svojstava. U Republici Srbiji, propisi koji su relevantni za suzbijanje govora mržnje nalaze se u Ustavu i nizu zakona: Zakonu o zabrani diskriminacije, Zakonu o javnom informisanju kao i Krivičnom zakoniku Republike Srbije. Detaljno zakonsko regulisanje, otkrivanje i sankcionisanje svih oblika zloupotreba Interneta kao i objavljenih sadržaja na društvenim mrežama, preko kojih se najčešće širi i potpiruje govor mržnje, uz povećanu pažnju, stalno praćenje i kontrolu od strane administratora i korisnika samo su najznačajniji faktori preventivnog delovanja.

Ključne reči: Internet, sloboda izražavanja, govor mržnje, prevencija.