

Hate Speech in Electronic Media and Challenges for the Criminal Justice System of the Republic Of Serbia

Prof. dr. Aleksandar R. Ivanović¹

Abstract

The author deals with the analysis of hate speech in electronic media in the Republic of Serbia, with a special focus on criminally legal repression. In this regard, the author seeks to analyze the current state of hate speech in electronic media and the provisions of criminal substantive law governing this matter, with the aim to point out the challenges that this form of criminal manifestations imposes to the criminal justice system of Republic of Serbia, and to offer certain recommendations for improving the situation on the field of combating online hate speech with measures of criminal law.

Keywords: *online hate speech, criminal offences, detection, prosecution, criminal justice system*

I. Introduction

Digital technologies provide opportunities for learning and education, provide access to information, can provide economic opportunities by offering young people training opportunities and access to job search services, or by creating new types of jobs. However, there is also the dark side of the Internet and digital technologies, from sexual or other types of abuse, child trafficking and other illegal activities that harm children, hate speech, violent radicalization and extremization, and so on.

The basic idea of the creators of the Internet is that the Internet should be a free highway, a public space where everyone can say what they have in mind. This wonderful innovation of the unrestrained platform is unfortunately abused. The Internet is open to use and abuse. Internet abuse threatens public space and poses many challenges at all levels: to the individual, the community, the state and the international community. Public authorities are in the early stages of learning how to deal with and how to combat Internet abuse, and this is especially true in the fight against hate speech on the Internet².

The Internet is the safest and most powerful means of spreading hate speech. A virtual space that can mobilize a large number of like-minded people at high speed and poses a real threat to society if it moves to the real world, which often happens. The Internet enables the creation of global groups whose members are physically

¹ Associate professor of Criminal law, Department for law sciences, International University of Novi Pazar, Republic of Serbia, Contact: a.ivanovic@uninp.edu.rs.

² Ivanović, R.A. (2018) *Hate speech on internet through illegality to punishment*. Yearbook, no.1 – Human rights protection „From unlawfulness to legality“, Novi Sad: Provincial Protector of Citizens – Ombudsman and Institute of Criminological and Sociological Research Belgrade, p. 416.

distant but between whom there is a sense of closeness due to the interactivity of social networks. It is in the explanation of the behavior of the individual in the crowd that the explanation of the dangers of hate speech over the Internet lies. In a group, the individual loses his individuality. She enjoys the protection of the majority and the comfortable comfort of anonymity. Due to the difficulty of identifying the individual in the mass, the sense of responsibility for one's actions is lost, which becomes much more extreme and destructive. There is a mass spread of certain behaviors like viruses, because individuals in the crowd tend to accept and follow the ideas, opinions and emotions of the majority without question³.

Hate speech is certainly not a stranger on the Internet, or even in conventional media, but it has certainly found its full negative affirmation on social media. This enables him, first of all, the anonymity of the subject as well as the focus on as large a population as possible, but also the distance from the passive subject, i.e. the subject to whom the hate speech itself is addressed⁴.

The benefits of the Internet for disseminating views and opinions are used by those who promote democratic values and the rule of law, as well as by racist or similar hate groups. The Internet is filled with a multitude of diverse commercial and private users. Among the attitudes available in interactions on the Internet are those that humiliate people on the basis of their race, ethnicity, national origin, gender and sexual orientation, and the like. Hate groups exploit this relatively cheap medium for ideological distribution. In this way, they can distribute pamphlets, letters and pictures to groups of users who can anonymously participate in strategic racist meetings and planning committees. One of the ugly sides of the internet is that it provides a global forum for advocates of intolerance and inequality⁵.

Hate speech often appears online, especially on social networks. Hate speech is defined as an expression that insults, threatens or encourages attacks on individuals or social groups based on race, religion, national origin, sexual orientation, disability or other characteristics. As in the real world, so in the virtual world, when we talk about the problem of hate speech, it is impossible not to touch on the issue of the relationship between freedom of speech on the Internet and hate speech on the Internet. In this regard, one of the most complicated questions regarding the relationship between freedom of speech and hate speech is the question of where does free speech end and where "offensive words" begin? If we start looking for an answer to this dilemma among theorists from the United States, we can conclude that in that sense, the gray zone has never transformed into black or white in this regard. Namely, libertarians who claim that freedom of speech implies hate speech believe that placing restrictions on freedom of speech weakens all rights. They also firmly believe that the state government should not restrict hate speech, unless it directly threatens peace. Communitarists share the opposite view. They believe that freedom of speech takes the last position in terms of the security of the community and all its citizens. Their view is that hate speech should not be allowed when propagated by those who should prevent individuals or groups from being treated with dignity and

3

⁴ Miladinović, A., (2013) *Facebook and crime [Fejsbuk i kriminalitet]*, Banja Luka: International Association of Criminalists, p. 180.

⁵ Tsesis, A., (2001) *Hate in Cyberspace: Regulating Hate Speech On the Internet*, San Diego Law Review 817/01, p. 832.

respect. Libertarians, on the other hand, argue that such restrictions would severely restrict freedom of speech and would be subject to an interpretation of intent that would be impossible to determine⁶.

It is often very difficult to define precisely when speech actually hurts someone, that is, where the dividing line is between freedom of speech and hate speech. Freedom of speech is a concept according to which no one can be legally prosecuted for the things he said or stated. According to this concept, no one can stop someone from saying something that contains hatred or prejudice. It is his right to freedom of expression.

However, hate speech is only someone's opinion or attitude, but the possibility that he could actually harm someone with that speech. And that is forbidden under international law. It is very ungrateful to try to point out the boundary between freedom of speech and hate speech. In this regard, in the next part, we will try to point out through examples where the limit of freedom of speech ends, and hate speech begins.

For example, someone may tell a particular person that they do not love him/her because of their ethnicity or cultural heritage. That is somewhat wrong, but it is his right and he cannot be arrested, prosecuted or punished by law for that. However, if someone threatens to kill someone or publicly calls on others to attack someone because of ethnicity or cultural heritage etc., it crosses the line of freedom of speech. This has gone from freedom of speech to a real threat that must be taken seriously. Therefore, hate speech has all the characteristics of freedom of speech plus some additional characteristics that distinguish it from it. What are the characteristics that distinguish freedom of speech from hate speech, are seen differently in the eyes of observers of that speech, from individual to individual, from one social group to another, and therefore we have a problem to say clearly in some situations this is hate speech and not freedom of speech, and vice versa.

When we considering hate speech on the Internet, we must take into account the existence of two types of media, the classic ones, which with the advent of the Internet got their digital, i.e. online version and the so-called social media. So we can divide the media into those that serve us ready-made information (newspapers, radio, and television) and those that allow us to participate in the creation, assessment, exchange and distribution of the same. The former are usually call as classic medias and the latter social media⁷.

While these first long exists at social life, and with the appearance of the Internet are just get their digital or online version, the latter appeared only with the development of modern information and communication technologies, especially the Internet, which now for the first time enables interactivity between the sender – recipient of message. Interactivity allows the establishment of a feedback of the participants in the communication process⁸.

Social media is most often defined as a set of new sources of online information that is created, initiated, circulated and used by consumers with the intention of

⁶ Gitlin, M. (2018) *When Is Free Speech Hate Speech?* New York: Greenhaven Publishing, 2018, p. 8.

⁷ Miladinović, S., (2015) *The growing importance of social media in a dynamic society* [Rastući značaj socijalni medija u dinamičnom društvu], Conference proceedings: Statehood, democratization and a culture of peace, Faculty of Philosophy, University of Nis, Nis, p. 342.

⁸ *Ibid.*

exchanging information about products, brands, services, features and specifics. They include numerous and diverse media, such as blogs by company sponsors for discussion, chat rooms, websites and consumer forums for ranking products and services, online discussion forums, blogs (sites containing digital, audio, images, movies or photos), social networking websites etc. The most common form of social medias are social networks. In this regard, in the next part of the paper we will present the forms of expression of hate speech through classical media, which have received their digital or online form through the Internet, such as the so-called electronic newspapers and through social media, among which from the aspect of hate speech the most current are social networks. However, before embarking on the analysis of the forms of expression of hate speech on the Internet, it is necessary to point out the current state regarding the issue of online hate speech and criminal repression in the Republic of Serbia. With regard to legal norms and sanctions prescribed for them, we can divide all hate speech into those that are allowed by law, those that are prohibited by law and represent a criminal offense and those that are sanctioned by some other norms that most often belong to self-regulation (codes of ethics, professional standards and conditions of use of communication platforms on the Internet etc.). It should be noted here that although precise data covering all content in internet communication is not yet available, according to existing research, hate speech that is considered a crime, extreme speech that calls for violence and other crimes, is very rare. In the analysis of hate speech, which is not illegal in Serbia, it was noticed that this type of speech makes up about 20% of the entire internet content reported for violating the rules of the Code of Journalists on various grounds⁹.

In the Final Report on Serbia of the European Commission against Racism and Intolerance (ECRI), which was adopted on March 22, 2017, ECRI expressed great concern over the increasing hate speech in public discourse in Serbia, which is further encouraged by media coverage. Politicians and the media use incendiary, pejorative and nationalist language and there has been a significant increase in regional tensions in the former Yugoslavia. The current public discourse is reminiscent of hate speech used before the recent wars in the region, and polls indicate the existence of great social distance between different segments of the population. Hate speech is spreading more and more online¹⁰. ECRI also states that it is necessary to improve the recording, investigation and punishment of hate speech and hate crimes and to ban racist, homophobic and transphobic hooligan groups (ECRI report 2017:5). According to the statistics of the prosecutor's offices, between January 1, 2011 and May 30, 2016, criminal charges were filed for hate speech against 216 people. Of these, 211 were filed under Article 317 of the CC (Instigating of national, racial and religious hatred and intolerance), three under Article 387 of the CC (Racial and other discrimination) and two under Article 174 of the CC (Infringement of reputation due to racial, religious, national or other affiliation) racial, religious, national or other affiliations). In 207 cases, hate speech referred to the victim's national or ethnic origin, in five cases to her religion, in one case to citizenship and in one to sexual orientation. Most crimes are committed against Roma and lesbians, gays, bisexuals or transgender (LGBT) people. 106 people were indicted and 41 people were convicted. 138 reports were

⁹ Nikolić, P., (2018), *Internet hate speech in Serbia [Govor mržnje u internet komunikaciji u Srbiji]*, Doctoral Dissertation, Faculty of political sciences, Belgrade, p. 4.

¹⁰ ECRI report on Serbia (fifth monitoring cycle) (2017), p. 4.

filed due to racial (20) and homophobic, i.e. transphobic (118) cybercrime. The outcome of these cases is convictions against 20 persons under Article 138 of the CC (Endangering Security), against three persons under Article 317 of the CC (Instigating of national, racial and religious hatred and intolerance) and against one person under Article 387. CC (Racial and other discrimination)¹¹.

Despite a number of positive developments, ECRI notes with regret that the criminal justice system still fails to effectively deal with hate speech. First, ECRI was informed that non-reporting is very common: many victims do not report hate speech to the police and other authorities. To address this issue, there need to be police officers and prosecutors specialized in investigating hate speech and hate crimes. As hate speech spreads more and more online, these officers must have the specialist knowledge and technical means to conduct investigations effectively; they must also work closely with the MUP's High-Tech Crime Department. Additional training on dealing with hate speech and hate crimes should be concentrated on these individuals¹².

Second, civil society groups state that the police are not always open to receiving complaints about hate speech cases, especially in relation to Roma and LGBT people. In addition, members of vulnerable groups often do not know where and how to complain about these incidents. More than 2 years have passed from the adoption of this report to the moment of writing this paper, and the Republic of Serbia has taken numerous measures to improve the situation in terms of criminally legal repression of online hate speech. Also, when the WHO declared a pandemic of the COVID-19 virus, at Republic of Serbia was introduced emergency state during which hate speech towards members of minority groups has been increased. The state of emergency and the restriction of freedom of movement as well as bigger use of social networks and Internet platforms', together with fear of infection lead to increasing of online hate speech against migrants and foreigners. Also, when it comes to online hate speech during the crisis with the virus, it should be emphasized that in the Republic of Serbia, in the electronic media and on social networks, there has been an expansion of sexist and misogynistic statements and attacks mainly on women performing public functions (teachers, doctors, members of various crisis groups etc.).

For the above reasons, in this paper we want to address the following issues: whether the Republic of Serbia has an adequate criminal framework for combating hate speech in electronic media, especially in conditions imposed by the crisis with the COVID-19 virus, as well as, what are challenges on combating hate speech on the Internet before the criminal justice system of the Republic of Serbia.

II. Hate Speech in Online Versions of Classic Media

The increasing use of the Internet has led to the classic media developing their online versions. So today we have the sites of many television channels and radio channels, where we can relatively quickly and easily find an archive of video and audio content that is broadcast on a given television or radio. Also, many television and radio channels have on their sites, the option of watching their program live over the Internet (so-called live streaming). Also, when it comes to print media, i.e.

¹¹ *Ibid*, p. 11.

¹² *Ibid*, p. 17.

newspapers, they also have their own websites, where journalistic articles and news, together with archives of printed content can be found very easily and quickly. What is a common feature of the online versions of the aforementioned media is that not only did they increase the potential audience in this way, that is, they accelerated the process of publishing information and provided the so-called "Digital eternity"¹³ of the same, they have already created the possibility to receive information from their audience, i.e. to have two-way communication with the audience. Namely, almost all of these media have the option of commenting, i.e. leaving comments below the published video, audio or textual content. Which, as in the case of social networks, which will be discussed later, gives users of these media the opportunity to influence the creation of content, or information that is disseminated through online versions of these media.

So, online versions of these media have two-way communication, the first that goes through the published video, audio or textual content by the editor of that media, and the second that appears in the form of comments from the audience, on the published content on that media.

The expansion of online media in this format represents a significant step in the democratization of reporting and the rise of the so-called. "Citizens journalism", which was further encouraged by social networks, primarily Facebook, Twitter and YouTube. Video broadcasts of protests, speeches by state officials, reporting on communal and similar problems, and analysis of social phenomena – all this and much more information is available at any time, on any device that can connect to the Internet. The emergence of online media is particularly characteristic of the local level. Namely, the expansion of online media and journalistic reporting through them at the level of local communities occurs primarily because the national media generally do not pay enough attention to local problems and opportunities in smaller communities.

When it comes to hate speech, it can appear on these media in classic forms, texts, videos, audio recordings that contain a message of hatred and a call for discriminatory or offensive activities. This is especially a problem with online media of this type that are of the local type. Namely, it is very common at the level of local communities at the position of journalists or editors in those medias are appointed persons who do not have a formal education in the field of journalism. Also, it is not so rare that these media are funded, i.e. they are owned by persons who are associated with certain political parties or interest groups operating in the area of that local community. In such situations, it often happens that these media are used as a means of a certain interest group to place content with which they want to discredit their political opponents, i.e. to propagate hatred and intolerance towards certain social groups, political options and their members or supporters. As a consequence, we have tendentious texts at this type of media, i.e. other types of content that directly or indirectly disparage individuals or social groups due to some personal characteristics and call for their exclusion from society, their persecution, and even open physical attack. Given that these are online media of the local type, such content generally does

¹³ Ivanović, R. A., Dečković, S., (2016) *The right to be forgotten as a fundamental principle of personal data protection in the digital age* [Pravo da budeš zaboravljen (Right to be forgotten) kao temeljni princip zaštite ličnih podataka u digitalnoj eri], Conference proceedings – Volume 2 / Scientific conference with international participation: European integration: justice, freedom and security, Tara, 24-26. May 2016, Criminalistic-Police Academy, Hanns Seidel Foundation, Belgrade, p. 478.

not attract public attention at the national level, but at the local level can have very negative consequences, because the creation and transmission of such content can be an initial trigger for discriminatory or offensive activities against social groups targeted by published hate speech.

In addition to the previously described form of hate speech in these media, it also occurs in the form of comments, i.e. two-way communication between the editors / journalists of these media and the audience. Namely, the publication of some content on these media, which has to do with interethnic, interfaith, inter-party and other issues, generally accompanies the possibility of commenting on the given content by the audience (users). This is where the space for propagating hate speech appears. People who comment on such content do so under a pseudonym, i.e. a nickname that they assign to themselves. In this way, they provide anonymity and in their comments they utter insulting, disparaging and other similar words with which they propagate hatred towards members of certain social groups. Of course, the members of such media should take care of such comments and remove them if they represent hate speech. However, many editorial boards of such medias do not have administrators in charge of supervising such comments, or they have them but not in sufficient numbers to be able to follow all the comments published on their portals. In addition, a good part of these types of media lack an efficient comment pre-moderation procedure. Namely, the rule on what is allowed in the comments and what does not have to be clearly stated in order for users to be familiar with it. That is, to know on what basis their comment was not published or was removed. When it comes to online versions of classic media at the local level, it often happens that the pre-moderation of comments is not performed, i.e. that comments that represent hate speech appear on such media. Since the possibility of commenting creates space for public discussion, then a comment that contains hate speech through other comments that appear as a comment on the original comment is only multiplied. This phenomenon can be explained by the so-called spiral of conflict. Namely, the conflict in these comments arises when two or more people or parties express disagreement on a topic. Conflict can be over different things or sides (commentators) may simply have different opinions or views on a given issue. In the second phase of the conflict, the parties begin to discuss and present arguments defending their views. Each side is convinced that its opinion is right, and cares to explain its views to others. There is no real dialogue, because neither one side does not listen to the other. Both sides experience a situation that leads nowhere, and that makes them frustrated. In a situation where the other side is not listening and there is no real dialogue, each side will repeat their arguments, but with heightened emotions. At this stage, the participants in the conflict usually start to get angry and feel that their interests are not taken into account, and that they are being treated unfairly. At this stage, things start to be mentioned that, in essence, have nothing to do with the conflict. For example, instead of dealing with the current subject of discussion, that is, discussion, they begin a discussion about the characteristics of the other. Generalizations and presentation of negative characteristics of another person, i.e. the group to which that person belongs, are typical in these situations. The inclusion of such irrelevant elements is in fact a suitable ground for the appearance of hate speech in these comments. Usually it starts with insulting words at the expense of some social group and continues with the same or similar response from the other side. This is followed by the involvement of a larger number of people in mutual insults, which usually leads to the classic form of hate

speech in which it publicly invokes discriminatory and offensive activities against members of another social group. Here we should mention the so-called. mass influence. Namely, in such situations, individuals in the mass of similar comments identify with the opinion of the group, that is, the masses are encouraged by support, which often leads to the extremization of their attitudes and the intensification of hate speech. At this stage, comments are already beginning to be made, which in terms of content have no essential connection with the news, that is, the information around which the discussion started. In the worst case, such commenting on online media can result in the transfer of conflict from the virtual environment to the real one, that is, scheduling the place and time of physical reckoning in the real world. So, the point is that the bigger the spiral of conflict, that is, the mutual conflict in the comments, the harder it is to put it under control. Therefore, it is very important to punish those persons who publishing in the form of comments all provoke further commenting. Namely, it is completely natural for people to feel uncomfortable due to reading comments in which offensive words are uttered at the expense of the social group to which they belong, whether they are national, religious, political or some other social groups. Which is especially dangerous when it comes to young people, who still do not have a firmly formed awareness and attitudes on many issues, where such comments can cause prejudice, stereotypes and intolerance towards members of other social groups¹⁴.

III. Hate Speech on Social Media

Social networks on the Internet today represent content that is very popular among Internet users, and especially among young people. In addition to representing a kind of trend, which contributes to their popularity, they also bring new opportunities in communication, as well as easier contact making, as well as the use and exchange of various content, which makes them attractive to use. Social networking platforms differ from other websites because they are based on individual activism. Another name for this type of connection is "web 2.0. applications", where these platforms are considered "hosts", while third parties are allowed to post content on them¹⁵.

There are three key actors in online speech on social media, and they are:

- a) the sender, i.e. the one who sends the information;
- b) the recipient, i.e. the one who receives the information;
- c) an intermediary or service provider acting as an intermediary between the speaker and the listener.

Social networking platforms act as intermediaries and ensure communication between the sender and the recipient of information. Social Networks are services or, better said, a functionally interconnected set (synergy) of services on the Internet that combines both technology (observed through services) and social interactivity in order to communicate between members of a particular social network. In addition to enabling communication between users, these networks are also a form of online socializing ("new sociability"), but they also allow (which is also important) a form of self-promotion for

¹⁴ Ivanović, R. A., Randelović, D., Totić, M., (2019) *Hate speech in electronic media and on social networks [Govor mržnje u elektronskim medijima i a društvenim mrežama]*, Novi Pazar: Monitor, p. 85.

¹⁵ Citron, D. K., & Norton, H. (2011) *Intermediaries and Hate Speech: Fostering Digital Citizenship for Our Information Age*. Boston University Law Review, (4), p. 1437.

various purposes, then the ability to post videos, photos, blogs, playing various games and other useful (or less useful) activities through these networks¹⁶.

Formal and one of the most well-known academic definitions of social networking sites, which contributes to further thinking and research of social online networking, defines online social networks as web-based services that allow individuals to create public or semi-public profiles within a closed centralized system, which have common connections (connections) and a selective list of connections which they call – friends (confidential circle of friends). These are usually a list of friends, as well as an overview of communication connections and practices that take place on the walls, notebooks, and personal pages of users¹⁷.

Use, i.e. communication via social networks takes place in such a way that it is first necessary for the user to register online, i.e. create his profile. Creating a profile involves entering basic information about yourself and possibly uploading a profile photo, followed by verification of identity by clicking on the link received from the social networking platform by email, which the user specified when creating his personal profile. After confirming the identity, i.e. valorization of the user, he starts using the given social network. The first steps are usually reflected in the search for friends or acquaintances that already have profiles on that network, or in sending invitations to friends and acquaintances in order to establish communication with them on a given network, or to establish a connection. After establishing connections with acquaintances, i.e. friends, communication follows, which takes place in the form of textual, video, audio or graphic expression and is mainly manifested in the form of dialogue or monologue. In addition to creating and posting their own content on social networks, they also offer other communication options, such as the so-called likes, i.e. visual marking that a given user likes some published content, or so-called hate, i.e. visual indication that a given user does not like some published content. Then, there is the possibility of commenting on the same, as well as the possibility of sharing, i.e. further forwarding of a given content through its connections on social networks.

When it comes to hate speech on social networks, we single out the following forms as the most common. The first is to create and publish various content (text, video, audio or graphics) on your own profile. However, the profile is real, not fake, i.e. the profile really belongs to the person listed as the profile owner. Such phenomena of hate speech are mostly found in persons who try to propagate the rhetoric of hate by abusing their freedom of speech, on the one hand, and on the other hand, to give importance to the declared hate speech with their authority. So, in such manifestations of hate speech on the Internet, we have when a person does not want to hide his identity, that is he does not want to be anonymous, but that those to whom the content was sent (the public, the audience, or the Internet community within a given social network), as well as those to whom hate speech was addressed, know that she stated it. This form of hate speech is mostly resorted to, the so-called public figures, such as politicians, representatives of religious communities, media workers, journalists, representatives of various associations or organizations, leaders of fan groups etc. These are people who enjoy respect in certain social communities, which is why in this way they want to impose their views on those who respect their character and work, or at least get their support. Here, however, it should be emphasized that

¹⁶ Miladinović, A., (2013), *op. cit.*, p. 36.

¹⁷ *Ibid*, p. 37.

this type of manifestation of hate speech on social networks is not only characteristic of public figures. Namely, it is often used by persons who are unknown to the general public, and who in this way want to draw attention to themselves, i.e. to be accepted by some groups because of their "courage" to publicly express the rhetoric of hatred.

The second model is the anonymous publication of content that contains hate speech. This model is usually resorted to by people who do not want the public, as well as members of the social group to whom hate speech is addressed, to know its true identity. The reason for the anonymous, i.e. more specific pseudonymous statement of hate rhetoric is reflected primarily in the intention to avoid responsibility and bear the legal consequences for the words spoken. This is usually done by creating a fake profile, with fake personal information, and then using it to propagate hate speech.

The third model is sharing of content that contains hate speech, and which someone else posted on social networks. This is actually a concept of promoting hate speech, which is characteristic primarily for social networks, because they provide that opportunity. Unlike previous models, where the person who posts content that contains hate speech is also the author of that content, in this model the person who shares or shares that material is not the author of the same, but primarily an intermediary or transmitter between the real creator of content and further audience.

The fourth model of hate speech on social networks is commenting, i.e. giving comments on published or shared content. In this regard, everything that is stated when commenting on content published on online versions of classic media applies.

Finally, it should be noted that sharing and commenting can be done both from real profiles and from fake ones, i.e. under a pseudonym.

IV. Characteristics of the Internet that Favor the Spread of Hate Speech

Various forms of angry, threatening, racist, misogynistic epithets are viral on the internet. The reasons why this happens lie in the factors related to the internet. Namely, the number of Internet users is increasing every year. People prefer to communicate through social networks than in classic, direct or direct communication. Also, another possibility offered by the internet is anonymity and virality. This means that many things can be said or posted online anonymously and that overnight content can be viral online. Because of all this, it is necessary to point out the specific characteristics of the Internet that affect the spread of hate speech. So the features of the Internet that favor the spread of hate speech in the online world are as follows:

Global availability. Information can reach a billion people in an instant. This is a positive characteristic of the Internet, however, it can also be negative at the same time because it enables the mobilization of like-minded people, the spread of hate speech, especially in situations when it appears as a companion to criminal attacks. The Internet is available to almost everyone today. It is no longer necessary to have expensive computers to have internet access. You can also access it via "smart" phones that are available to everyone. Through them, global accessibility gained a wider, temporal dimension. Content on the Internet can be accessed at any time and in any place. You don't even have to be constantly active on networks. They signal to you when the desired content arrives.

Anonymity. The term anonymity comes from a Greek word "anonymia" (ἀνωνυμία), which means anonymous, unidentified or unknown name, and usually refers to the appearance of a person in public. Consequently, anonymity occurs if a

person's identity is involved in a non-transparent / unpublished process in an indefinable sense, because the actor remains unknown to other actors acting or does not appear to other participants in the interaction without recognizing his true identity¹⁸. However, anonymity does not necessarily require complete ignorance of a person's identity or lack of knowledge of his or her name. In order to distinguish anonymity from the impossibility of identifying a person, it is necessary for one party to vaguely know about the existence of the other party without knowing his or her complete identity. It is also necessary to point out the difference between anonymity and pseudonym, which is characterized by the use of a false name, although this practice can also lead to anonymity. In fact, we can freely say that pseudonymity can be described as a common variant of anonymity in which a different name from the real author of a content is shown¹⁹. Which is especially pronounced when it comes to anonymity on the internet. You can be completely anonymous online. You do not need to sign with your personal name when leaving comments on content that is published on the Internet. It is not necessary to leave personal data. You can open fake accounts on social networks, post fake pictures, present yourself as a fake name. You can have multiple e-mail addresses and if a portal requires registration, you do not have to leave your private or business email. Because of this characteristic of the Internet, the sense of personal responsibility for actions that become much more extreme and destructive than they would be if they were done in the real world is lost. Terminological distinction and division into real and virtual worlds can lead to the dangerous misconception that the virtual world is not real, that what happens in it does not create consequences.

Interactivity. The Internet allows interaction between its users. Interaction that is not limited by geographical boundaries. This characteristic enables like-minded people to mobilize and self-organize. It also allows users to create their own content that is placed on the Internet with their comments. Users themselves become distributors of content on the Internet by transmitting it.

Instantness. The Internet allows you to transfer information in an instant, with one click. Event information becomes available to billions of Internet users almost instantly when it occurs.

The decentralized structure prevents uniform application of regulations because the placement of content on servers located in other countries prevents the application of national regulations on them.

V. Criminal Law Sanctioning of Hate Speech on the Internet in the Republic of Serbia

Criminal Code of the Republic of Serbia²⁰ Article 317 prescribes the criminal offense of Instigating National, Racial and Religious Hatred and Intolerance. The

¹⁸ Weber, R., Heinrich, U., (2012) *Anonymization*, Springer: London, p. 1.

¹⁹ Ivanović, R. A., (2017) *The right to privacy and anonymity within the scope of freedom and security*, Conference proceedings: International Scientific Conference „Freedom, security: the right to privacy (ed. Zoran Pavlović), Provincial Protector of Citizens – Ombudsman and Institute for Criminological and Sociological Research, Novi Sad, p. 157.

²⁰ Criminal Code of Republic of Serbia ("OG", no. 85/2005, 88/2005, 107/2005, 72/2009, 111/2009, 121/2012, 104/2013, 108/2014, 94/2016 and 35/2019).

protective object of this crime is the national, religious and racial equality of the citizens of the Republic of Serbia. Equality and unity of citizens in the Republic of Serbia is built on mutual respect and trust of all who live in it. This unity is preserved by constant nurturing, respect for all the characteristics of national, religious, racial, cultural and other values that characterize certain peoples and ethnic communities in the Republic of Serbia²¹. The Code prescribes the basic, qualified and special form of this criminal offense. Intent is the only form of guilt for this criminal offence.

The act of committing the basic form of this crime is set alternatively as instigating or exacerbating national, racial or religious hatred, or intolerance (paragraph 1). This further means that the act will exist both in a situation when there is no national, racial, religious hatred or intolerance between peoples or ethnic communities, so it is provoked in different ways (i.e. it creates, causes), and when it exists as latent, a state or as a state of lesser intensity, and then further inflamed²². Instigating implies the creation of hatred or intolerance on the mentioned grounds, in a situation where such phenomena did not exist, while exacerbating is an increase (deepening, i.e. expansion) of already existing hatred, intolerance and bigotry. Thus, Instigating or exacerbating national, racial or religious hatred or intolerance can be done in different ways and by different means: speech, inscription and other actions, which influences the creation or strengthening of certain feelings or attitudes towards a nation or ethnic community. In terms of the nature of this crime, hatred is understood as a negative feeling of high intensity which represents a very fertile ground for undertaking various excessive behaviors towards members of a certain people or ethnic community. When it comes to intolerance, it is understood in relation to hatred as a negative feeling of "milder" intensity (the so-called vestibule of feelings of hatred), which can also cause certain negative behaviors that may ultimately contribute to the commission of this crime. It should be noted here that in criminal law theory, the question arises, and there are conflicting opinions, as to whether contempt for a nation or ethnic community can be included in the notion of intolerance. Namely, unlike hatred and intolerance, contempt for a nation or ethnic community, is a negative attitude that is most often expressed in the form of ignoring and as such does not contain the potential to perform certain activities and actions against a nation or ethnic community, so that provoking and inciting contempt would still not be sufficient for the existence of this crime²³. The passive subject in this criminal offense is represented by peoples, national minorities and ethnic communities living in the Republic of Serbia, and in connection with this, committing actions under this criminal offense against national or ethnic groups not living in the Republic of Serbia cannot constitute this criminal offense.

The criminal offence has been done by undertaking an activity with the aim of instigating or exacerbating hatred or intolerance, i.e. it is not necessary that these activities have led to the provocation or incitement of hatred or intolerance, or that a consequence has occurred. Therefore, the very consequence of this act is endangering the equality of citizens regardless of national, racial or religious affiliation. This form of criminal offence is punishable by imprisonment from six months to five years.

²¹ Jović, M. (2011). *Criminal law – special part – script I* [Krivično pravo – posebni deo – skripta I], Novi Pazar: International university of Novi Pazar, p. 240.

²² Stojanović, Z. (2017). *Commentary on the Criminal Code: Sixth Amended Edition* [Komentar Krivičnog zakonika: šesto izmenjeno i dopunjeno izdanje], Belgrade: Službeni glasnik, p. 936.

²³ *Ibid.*

The qualified form of this criminal offense exists when some of the legally established qualifying circumstances (paragraph 2) have been met, which indicate the commission of the basic form of the criminal offense: a) coercion, b) abuse, c) endangering security, d) exposing national, ethnic or religious symbols, e) damage to other people's property, f) desecration of monuments, memorials or graves²⁴. For this form of criminal offence is punishable by one to eight years in prison.

The second qualified form will exist when the basic form or the first serious form of this criminal offense is committed by abuse of position or authority or if as a result of these acts there were riots, violence or other serious consequences for the common life of peoples, national minorities or ethnic groups living in Serbia (paragraph 3). For this form of offense, in the event that under the qualifying circumstances the offense referred to in paragraph 1 was committed, a sentence of imprisonment of one to eight years is prescribed, and if the offense referred to in paragraph 2 is committed, imprisonment of two to ten years.

In addition to the aforementioned criminal offense, Article 387 prescribes the criminal offense of Racial and Other Discrimination, paragraph 3 of which prescribes the conduct of this criminal offense, which is reflected in: a) spreading the idea of superiority of one race over another, b) propagating racial hatred or inciting racial discrimination. These are actually three forms of racist propaganda²⁵. For this form of crime, a prison sentence of three months to three years was listed. Also, paragraph 4 of this criminal offense prescribes the act of committing this offense, the form of which is reflected in the dissemination or otherwise making publicly available texts, images or any other presentation of ideas or theories that advocate or incite hatred, discrimination or violence against any persons or groups of persons based on race, color, religion, nationality, ethnic origin or any other personal characteristic. This form of crime is punishable by imprisonment from three months to three years.

Article 174 prescribes the criminal offense of Infringement of Reputation Due to Racial, Religious, National or Other Affiliation, which stipulates that whoever publicly exposes a person or group for belonging to a certain race, skin color, religion, nationality, ethnic origin or any other personal characteristic shall be punished with a fine or imprisonment for up to one year.

Article 138 prescribes the criminal offense of Endangering Security, Article 1 of which stipulates that whoever endangers the security of a person by threatening to attack the life or body of that person or a person close to him, shall be punished by a fine or imprisonment for a term not exceeding one year. The act of committing the basic form of a criminal offense (paragraph 1) is the use of the threat to attack the life or body of a person or a person close to him. So, it is not ordinary, but the so-called a qualified threat, given that it is necessary to threaten a specific evil – an attack on the life or body of a passive subject or a person close to him. It can be threatened either only by an attack on life, or only by an attack on bodily integrity (therefore, the threat of inflicting light bodily injuries is sufficient), or cumulatively by an attack on life and

²⁴ Simonović, D. (2010) *Criminal offenses in Serbian legislation [Krivična dela u srpskoj legislativi]*, Belgrade: Službeni glasnik, p. 624.

²⁵ Ivanović, R. A., Randelović, D., (2019) *Sanctioning hate speech on the internet, according to the national regulations of the Republic of Serbia [Sankcionisanje govora mržnje nainternetu prema nacionalnoj regulative Republike Srbije]*, Arhiv za pravne i društvene nauke vol. 1, no. 1, Law faculty Lukavica, p. 55.

an attack on the body. Under the notion of a close person in the sense of this incrimination should be considered blood relatives in the direct line, brother and sister, spouse and a person living in a permanent extramarital union, adoptive parent and adoptee, custodian and custodian. As the criminal offense of endangering security in Serbian criminal law is set as a consequential criminal offense, it is understood that the passive subject knows about the threat. Also, for the existence of the criminal offense is not affected by whether passive subject find out for threat directly from the perpetrator or from another person (so-called "indirect threat"). Paragraph 2 stipulates that whoever commits the act referred to in paragraph 1 of this Article against several persons, or if the act caused disturbance of citizens or other severe consequences, shall be punished by imprisonment for a term between three months and three years.

VI. Conclusion

Based on all the above-mentioned, we can conclude that in Serbia there is a relatively good criminal law framework for the suppression of hate speech that spreads through the electronic media. So-called online hate speech is covered by several acts prescribed in the Criminal Code. However, it can be said that the prescribed criminal offences are not fully in line with ECRI's General Policy Recommendation no. 7 on national legislation to combat racism and racial discrimination. This is because certain elements of these recommendations have still not been introduced into the criminal justice system of the Republic of Serbia. Namely, in the criminal offense of Instigating National, Racial and Religious Hatred and Intolerance from Article 317, paragraph 1 of the Criminal Code of Republic of Serbia, incitement to national, racial or religious hatred or intolerance among peoples or ethnic communities living in Serbia is prescribed as an act of committing the criminal offense. This legal wording does not meet the requirements contained in the ECRI Recommendations in several respects: no incitement to violence is mentioned, and skin color, language, citizenship, national origin, sexual orientation and gender identity are omitted as grounds for hatred or intolerance. In addition, despite the fact that incitement to hatred against individuals and groups living outside Serbia should be punishable, as it is pointed out in the paper, this provision only protects against incitement to hatred among "peoples or ethnic communities living in Serbia. Encouraging discrimination is punishable under the provisions of the criminal offense of Instigating National, Racial and Religious Hatred and Intolerance 387, paragraph 3 of the Criminal Code of Republic of Serbia, which also does not state all the prohibited reasons listed in ECRI Recommendations no. 7. Article 387, paragraph 3 of the Criminal Code of Republic of Serbia prohibits the spread of ideas about the superiority of one race over another. This is not fully in line with the ECRI Recommendation, which should also prescribe as a criminal offense and the expression of an ideology that belittles or defames a group of persons on the grounds listed. Also, in this crime, the presentation of ideas or theories that advocate or incite hatred, discrimination or violence against any person or group of persons based on race, color, religion, nationality, ethnic origin or any other personal characteristic. This is not fully in line with the ECRI Recommendation, as it does not cover the preparatory actions for the production or storage of such material, and because it only mentions incitement to hatred, but not other acts listed in the Recommendation.

In addition to the problem of insufficient compliance of domestic criminal legislation with ECRI Recommendations on General Policy no. 7 on national legislation to combat racism and racial discrimination, there is also a problem that is reflected in the fact that the general public, but also many staff in state bodies, as well as in the judicial system, are not sufficiently familiar or versed in dealing with such criminal offences. Namely, despite the relatively good legal framework that can regulate hate speech, judging by the very small number of processed and even initiated cases dealing with hate speech in Serbia, it is obvious that there are certain obstacles to its effective suppression.

Of course, it is ungrateful to talk about the efficiency of the police and criminal justice bodies, without the existence of official statistics on the number of reported, initiated and completed proceedings for such acts committed online, but based on reports from many NGOs we can conclude that the situation has not changed much from ECRI report from 2017. In this regard, there is still a need to raise staff capacity whose dealing with combating online hate speech, especially at the level of local police stations and prosecutors offices.

Based on all the above, we can conclude that there are two biggest challenges in terms of criminal repression in relation to hate speech in the electronic media in the Republic of Serbia. The first is reflected in the harmonization of the provisions of substantive criminal law with the ECRI Recommendations on General Policy no. 7 on national legislation to combat racism and racial discrimination and second, which is reflected in the need for constant education of police and prosecutors from the field of methodology of detecting and investigating such crimes, with special focus on local context of appearance of this negative phenomena.

References

1. Citron, D.K., Norton, H. (2011) *Intermediaries and Hate Speech: Fostering Digital Citizenship for Our Information Age*. Boston University Law Review, (4), p. 1435-1484;
2. Criminal Code of Republic of Serbia ("OG", no. 85/2005, 88/2005, 107/2005, 72/2009, 111/2009, 121/2012, 104/2013, 108/2014, 94/2016 and 35/2019);
3. ECRI report on Serbia (fifth monitoring cycle) (2017);
4. Gitlin, M. (2018) *When Is Free Speech Hate Speech?* New York: Greenhaven Publishing, 2018;
5. Ivanović, R.A. (2018) *Hate speech on internet through illegality to punishment*. Yearbook, no.1 – Human rights protection „From unlawfulness to legality“, Novi Sad: Provincial Protector of Citizens – Ombudsman and Institute of Criminological and Sociological Research Belgrade, p. 405-417;
6. Ivanović, R. A., (2017) *The right to privacy and anonymity within the scope of freedom and security*, Conference proceedings: International Scientific Conference „Freedom, security: the right to privacy (ed. Zoran Pavlović), Provincial Protector of Citizens – Ombudsman and Institute for Criminological and Sociological Research, Novi Sad, p. 152-165;
7. Ivanović, R.A., Dečković, S., (2016) „The right to be forgotten as a fundamental principle of personal data protection in the digital age“ [*Pravo da budeš zaboravljen* „(Right to be forgotten) kao temeljni princip zaštite ličnih podataka u digitalnoj eri], Conference proceedings – Volume 2 / Scientific conference with international participation: European integration: justice, freedom and security, Tara, 24-26. May 2016., Criminalistic-Police Academy, Hanns Seidel Foundation, Belgrade, p. 477-490;

8. Ivanović, R.A., Randelović, D., (2019) *Sanctioning hate speech on the internet, according to the national regulations of the Republic of Serbia [Sankcionisanje govora mržnje nainternetu prema nacionalnoj regulative Republike Srbije]*, Arhiv za pravne i društvene nauke vol. 1, no. 1, Law faculty Lukavica, p. 49-61;
9. Ivanović, R.A., Randelović, D., Totić, M., (2019) *Hate speech in electronic media and on social networks [Govor mržnje u elektronskim medijima in a društvenim mrežama]*, Novi Pazar: Monitor;
10. Jović, M. (2011) *Criminal law – special part – script I [Krivično pravo – posebni deo – skripta I]*, Novi Pazar: International university of Novi Pazar;
11. Miladinović, A., (2013) *Facebook and crime [Fejsbuk i kriminalitet]*, Banja Luka: International Association of Criminalists;
12. Miladinović, S., (2015) *The growing importance of social media in a dynamic society [Rastući značaj socijalni medija u dinamičnom društvu]*, Conference proceedings: Statehood, democratization and a culture of peace, Faculty of Philosophy, University of Nis, Nis, p. 341-355;
13. Nikolić, P., (2018) *Internet hate speech in Serbia [Govor mržnje u internet komunikaciji u Srbiji]*, Doctoral Dissertation, Faculty of political sciences, Belgrade;
14. Simonović, D. (2010) *Criminal offenses in Serbian legislation [Krivična dela u srpskoj legislativi]*, Belgrade: Službeni glasnik;
15. Stojanović, Z. (2017) *Commentary on the Criminal Code: Sixth Amended Edition [Komentar Krivičnog zakonika: šesto izmenjeno i dopunjeno izdanje]*, Belgrade: Službeni glasnik, p. 936;
16. Tsesis, A., (2001) *Hate in Cyberspace: Regulating Hate Speech On the Internet*, San Diego Law Review 817/01, p. 817-874;
17. Weber, R., Heinrich, U., (2012) *Anonymization*, Springer: London.