



Regulation of Harmful
Content Online in
Bosnia and Herzegovina

BETWEEN FREEDOM OF EXPRESSION AND HARMS TO DEMOCRACY

Regulation of Harmful Content Online in
Bosnia and Herzegovina
**BETWEEN FREEDOM OF EXPRESSION
AND HARMS TO DEMOCRACY**

Anida Sokol and Maja Čalović

Sarajevo, 2022

TITLE:

Regulation of Harmful Content Online in Bosnia and Herzegovina:
Between Freedom of Expression and Harms to Democracy

PUBLISHER:

Mediacentar Sarajevo

ADDRESS:

Koševo 26, 71000 Sarajevo

AUTHORS:

Anida Sokol and Maja Čalović

REVIEWER:

Sanela Hodžić

PROOFREADER:

Gina Landor

DESIGN:

Emina Lagumdžija

YEAR OF PUBLICATION:

2022

This publication was produced with the financial support of the European Union and UNESCO. The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO or the European Union concerning the legal status of any country, territory, city or area or its authorities, or concerning the delimitation of its frontiers or boundaries.

The authors are responsible for the choice and the presentation of the facts contained in this publication and for the opinions expressed therein, which are not necessarily those of UNESCO or the European Union and do not commit the Organizations.

ISBN 978-9958-584-40-4

CIP zapis dostupan u COBISS sistemu
Nacionalne i univerzitetske biblioteke BiH
pod ID brojem 49929734

EXECUTIVE SUMMARY

The study Regulation of Harmful Content Online in Bosnia and Herzegovina: Between Freedom of Expression and Harms to Democracy is part of the UNESCO project Social Media 4 Peace, funded by the European Union, whose aim is to strengthen the resilience of societies to potentially harmful content spread online, while protecting freedom of expression and enhancing the promotion of peace through digital technologies, notably social media.

The study gives an overview of types of harmful content and their main target groups, and presents the legislative, regulatory and self-regulatory frameworks against harmful content online in Bosnia and Herzegovina. The study also looks at the obstacles in implementing the available frameworks in Bosnia and Herzegovina and their alignment with international standards, and provides recommendations for their improvement.

There are no unanimous definitions of what is considered harmful since the concept of harm is subjective and depends greatly on the context and culture, and can be different among states. Illegal harmful content is content that is not compliant with legislation, while potentially harmful content might not be prohibited, but can harm groups and individuals. This study focuses on illegal content and potentially harmful content, taking into consideration the delicate political situation, the difficult war heritage and the existence of divisive ethno-nationalistic rhetoric in Bosnia and Herzegovina. The study focuses only on harmful content that can thwart democratic processes, hamper citizens' ability to take informed actions and impede reconciliation.

The study specifically focuses on five types of harmful content: a) hate speech and hate narratives; b) denials of war crimes and glorification of war criminals; c) ethno-nationally and/or politically biased media reporting; d) disinformation; and e) attacks, threats and smear campaigns against individuals.

After giving overviews of the five types of harmful content, their targets and consequences, the following chapters are dedicated to the legislative, regulatory and self-regulatory frameworks for the five types of harmful content, how effectively they are used online, what the major obstacles are in their implementation and to what extent they are aligned with international standards. The study also addresses the practices of the courts, the Communications Regulatory Agency of Bosnia and Herzegovina, the Press Council of Bosnia and Herzegovina and other relevant actors in countering harmful content. The final parts of the study are dedicated to the community guidelines of social networks and examples of frameworks in other countries.

The scope of harmful content online in Bosnia and Herzegovina is worrying and calls for a comprehensive response. The study emphasizes the need to safeguard freedom of expression and to find responses and practices that are aligned with international human rights law and that do not chill or censor online speech or discourage the flow of diverse sources of information and opinions.

CONTENT

10 1. Introduction	14 2. Terminology and definitions	20 3. Methodology	24 4. The information system in BiH – an overview	47 6.2. Prohibition of condoning, denial and justification of crimes of genocide, crimes against humanity and war crimes	47 6.2.1. Legislation	49 6.2.2. International standards and recommendations	50 6.3. Regulation of disinformation
30 5. Harmful content: types, consequences and targets	32 5.1. Hate speech and hate narratives	33 5.2. Denials of war crimes and glorification of war criminals	34 5.3. Ethno-nationally and politically biased media reporting	50 6.3.1. The regulatory framework	51 6.3.2. Defamation laws	52 6.3.3. Laws on public order	53 6.3.4. International standards and recommendations
35 5.4. Disinformation	36 5.5. Attacks, threats and smear campaigns against individuals	38 6. Regulation of harmful content	40 6.1. Regulation of hate speech	54 6.4. Regulation of ethno-nationally and/or politically biased reporting	55 6.4.1. The regulatory framework	55 6.4.2. The Election Law	56 6.4.3. International standards
40 6.1.1. Criminal codes	43 6.1.2. Sanctioning language that incites to violence during election periods	44 6.1.3. Prohibition of hate speech on television and radio stations	45 6.1.4. International standards on hate speech and recommendations	57 6.5. Regulation of threats, attacks and smear campaigns	59 International standards and recommendations	59 6.6. Other mechanisms – the Institution of the Human Rights Ombudsman	62 7. Self-regulation and monitoring of harmful content

63 7.1. The Press Council of Bosnia and Herzegovina	64 7.2. Civil society and international organizations	64 7.2.1. Fact-checking and media watchdogs	65 7.2.2. Monitoring and reporting mechanisms	84 9.3. The Digital Services Act	85 9.4. The EU Code of conduct on countering illegal hate speech online	86 10. Additional initiatives to influence social media content moderation	87 10.1. The Santa Clara Principles
67 7.3. Online media	67 7.4. Recommendations	70 8. Community guidelines of social networks against harmful content	73 8.1.1. Overview of community guidelines of social networks	87 10.2. The Christchurch Call	88 10.3. The Manila Principles on Intermediary Liability	88 10.4. Social media councils	90 11. Conclusion and recommendations
73 8.1.2. Facebook	76 8.1.3. Instagram	77 8.1.4. YouTube	79 8.1.5. Twitter				
80 8.1.6. TikTok	82 9. Examples of regulations of harmful content from EU countries	83 9.1. Network Enforcement Act	84 9.2. Communications Platform Act				

1. INTRODUCTION

The information system in BiH has changed drastically over the past decade, giving rise to numerous online media, accounts and groups on social media and non-professional content producers. The internet and social networks have provided unprecedented space for people to exercise their right to freedom of opinion and the freedom of expression. However, the internet and social networks have also posed new challenges to freedom of expression, including content that can harm the rights and reputation of others, national security, public order and public health.

Twenty-seven years after the end of the war, Bosnia and Herzegovina is facing a difficult political crisis. The continued divisive ethno-nationalistic rhetoric and the warmongering language of political elites disseminated by the media intensified in the second half of 2021, inciting harmful content online, including hate speech and disinformation. Such content is problematic in a country that faces a difficult war heritage, mutually exclusive memories of the past and constant disagreements between ethno-national political elites that halt the country's progress to a fully functioning democratic society.

The media and public communication play a key role in maintaining ethno-national divisions. Through a selective covering of current and historical events and a lack of criticism towards inflammatory and divisive ethno-national politics, the media continue to perpetuate ethno-national divisions and contribute to the mutually exclusive memories of the past. The divisive and inflammatory rhetoric of political elites polarizes public debates and incites and legitimizes hateful and harmful content on the internet, on social networks and in user-generated content.

The internet and social networks are used by diverse actors for spreading hate speech, denials of war crimes and glorification of war criminals, disinformation, politically and ethno-nationally biased media reporting and smear campaigns against individuals for diverse motives that can undermine freedom of expression and democratic processes. Hate speech is rarely present in the content of the media, but social media platforms such as Facebook, YouTube and Twitter and comments sections of online media contain hate speech and hateful narratives most often targeting ethno-national groups, minorities and women. Disinformation, on the other hand, has been spread by both mainstream media and anonymous online portals and groups on social networks for political and/or financial aims or unintentionally, and especially proliferated during the pandemic of COVID-19, endangering public health. Ethno-nationally and/or politically biased media reporting has been part of the content of online media and anonymous portals influencing public opinion and voters' behavior, and online media generally lack more constructive and in-depth media reporting. The internet and social networks are also used as platforms for smear campaigns, attacks and threats against individuals, journalists, civil society representatives and political opponents with the aim to discredit and intimidate them.

Despite the amendments to the criminal law that prohibit public condoning, denial, gross diminution or attempts to justify the crimes of genocide, crimes against humanity, war crimes and glorifications of war criminals that were introduced in 2021, such instances even intensified in the beginning of 2022, fueled by the inflammatory and nationalistic rhetoric of political elites. They might increase before the general elections in 2022, as war crime denials and glorification of war

criminals are narratives that attract voters, and failure to prevent and sanction them is an obstacle to democracy and reconciliation.

All of this calls for better tools and mechanisms for the regulation of harmful content online, taking into consideration the specific local context and protection of freedom of expression.

The aim of the research is to map the legislative, regulatory and self-regulatory frameworks in place in Bosnia and Herzegovina that address harmful content online. The research gives an overview of the types of harmful content and its main target groups, presents the legislative, regulatory and self-regulatory frameworks against these types of harmful content and social media community guidelines, and provides recommendations for their improvement. The study points out that measures against harmful content should be compatible with international human rights law and not used to discourage freedom of expression and exchange of information, or to chill and censor online speech.

This research study is part of the UNESCO project Social Media 4 Peace, whose aim is to strengthen the resilience of societies to potentially harmful content spread online, in particular hate speech inciting violence, while protecting freedom of expression and enhancing the promotion of peace through digital technologies, notably social media. The project, funded by the European Union, is aligned with UNESCO's overall strategy to combat disinformation by fostering information as a public good and strengthening the transparency of the internet ecosystem. It also aims to contribute to the UN Plan of Action on Hate Speech launched by UN Secretary General Antonio Gutierrez to combat the disturbing online groundswell of xenophobia, racism and intolerance.

2. TERMINOLOGY AND DEFINITIONS

The terms illegal and harmful content and the need to differentiate the two were discussed for the first time in the document “Illegal and Harmful Content on the Internet”, by the Commission of the European Communities in 1996. The document argues that the two categories of content pose radically different issues of principle and call for very different legal and technological responses (Commission of the European Communities 1996, 10). In the document illegal content is discussed as content that falls under different rules that prohibit its use and distribution and could pertain to civil law, such as breaches of copyright, or to criminal law, such as child pornography, trafficking in human beings and dissemination of racist material (Commission of the European Communities 1996, 10). Harmful content, on the other hand, is defined as content that offends the values and feelings of other persons, such as content expressing political opinions, religious beliefs or views on racial matters (Commission of the European Communities 1996, 11).

The European Commission’s Action Plan for a Safer Internet 1994-2004 further specified that illegal content must be dealt with at source by the police and the judicial authorities, but assistance for restricting the circulation of illegal content can be provided by means of self-regulation, such as codes of conduct with consumer backing. On the other hand, harmful content, which includes both content that is authorized but has restricted circulation, and content that could be offensive to some users even if publication is not restricted because of freedom of speech, should be dealt with by filtering tools and rating mechanisms and by the promotion of awareness and fostering self-regulation (Action Plan for a Safer Internet 1994-2004, Summary).

Another, more recent study commissioned by the European Parliament, “Reform of the EU liability regime for online intermediaries: Background on the forthcoming digital services act”, defines illegal content as a large variety of content categories that are not compliant with legislation, such as violations of copyright, counterfeiting, privacy, defamation, data protection, hate speech, and child sexual abuse material (Madiega 2020, 10). Potentially harmful content, on the other hand, is not strictly forbidden, but can have harmful effects, including bullying, disinformation, misinformation and malinformation, which could hamper the ability of citizens to take informed actions (Madiega 2020, 10). The study also stressed the difficulty in distinguishing what is illegal content online from content that is harmful but not illegal, since the concept of harmful is subjective and greatly depends on the context, and can vary considerably between states (Madiega 2020, 11). It points to the danger that introducing rules to address harmful online content into legislation would have great consequences for freedom of expression and other fundamental rights (Madiega 2020, 11).

In the proposal of the Digital Service Act (DSA), which was proposed by the European Commission in 2020 with the aim to create a safer digital space, illegal content is defined as “any information, which, in itself or by its reference to an activity, including the sale of products or provision of services is not in compliance with Union law or the law of a Member State” (Article 2, g). In the Explanatory Memorandum of the proposal, a general agreement among stakeholders was reached during consultations that harmful content should not be defined in the Digital Services Act and should not be subject to removal obligations because this is a delicate area with severe implications for the protection of freedom of expression (Explanatory Memorandum). The fact that harmful content is not defined by the

DSA poses difficulties however, since this would mean that intermediary service providers will decide on its interpretation and the task will be left to voluntary codes of conduct.

Discussions on what should be prohibited and definitions of certain types of illegal and harmful content can be followed through UN instruments and reports of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. A report of the Special Rapporteur submitted to the UN General Assembly in 2011 distinguishes three categories of content: (a) expression that constitutes an offence under international law and can be prosecuted criminally; (b) expression that is not criminally punishable but may justify a restriction and a civil suit; and (c) expression that does not give rise to criminal or civil sanctions, but still raises concerns in terms of tolerance, civility and respect for others (A/66/290, 7-12). The first category includes child pornography, direct and public incitement to commit genocide, advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence and incitement to terrorism. According to the report, content that does not fall into the first category should not be criminalized, while for content that raises concerns in terms of civility and respect for others, efforts should be focused on addressing the root causes and implementing strategies of prevention (A/66/290, 14).

Starting from international standards and discussions on illegal and harmful content, this study focuses on local concerns regarding regulation and moderation of such content. Taking into consideration the delicate political situation, the difficult war heritage and the existence of divisive ethno-nationalistic rhetoric in Bosnia and Herzegovina, this research focuses on illegal content and potentially harmful content that is specific to a post-conflict society. Even though different types of illegal and potentially harmful content exist, such as child sexual exploitation and breaches of copyright, the research concentrates on harmful content, both illegal and potentially harmful, that can thwart democratic processes, hamper citizens' ability to take informed actions and impede reconciliation processes in Bosnia and Herzegovina. It distinguishes five types of harmful content, which were selected due to their potential to impact democratic processes in the country and based on extensive desk research and interviews. These types of content include: a) hate speech and hate narratives; b) denials of war crimes and glorification of war criminals; c) ethno-nationally and/or politically biased manipulative content; d) disinformation; and e) attacks, threats and smear campaigns against individuals.

HATE SPEECH AND HATE NARRATIVES

There are no unanimous definitions of hate speech. The United Nations Action Plan on Hate Speech launched in 2019 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."

The Council of Europe defines hate speech as "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" (Recommendation No. R (97) 20).

The International Covenant on Civil and Political Rights requires states to prohibit "any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence" (Article 20) and the International Convention on the Elimination of All Forms of Racial Discrimination calls states to "declare an offence punishable by law all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin" (Article 4).

In addition to incitement to discrimination, hostility and violence, which is prohibited by international law, there are other types of hate speech that are not illegal but could be harmful. For example, hate narratives that construct other groups based on ethnicity, gender, sexual orientation, professional, political views, and other identity factors in a negative light. These types of content often fall into a gray area, can be harmful and require different mechanisms to address them.

DENIALS OF WAR CRIMES AND GLORIFICATION OF WAR CRIMINALS

Denials of war crimes and glorification of war criminals directed against a group of persons defined in reference to their race, colour, religion or other protected characteristic is also a form of hate speech. The Council Framework Decision 2008/913/JHA on combating certain forms and expressions of racism and xenophobia by means of criminal law requires member states to prohibit “publicly condoning, denying or grossly trivialising crimes of genocide, crimes against humanity and war crimes ... directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin” (Article 1, c). However, such content, as defined in the framework, should be punishable “when the conduct is carried out in a manner that is likely to incite to violence or hatred against such a group or a member of such a group” (Article 1, c). Such acts include any attempt to claim that the crime did not take place, to deny the established facts and to challenge that there was the intent to commit the crime. It refers not only to direct denials but also diverse forms that downplay the scope and character of crimes, their justifications and approval (Gačanica 2021; Gačanica and Finkeldey 2019).

DISINFORMATION

There is no universally accepted definition of disinformation. A report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (A/HRC/47/25, 4) states that disinformation is understood as false information that is disseminated intentionally to cause serious social harm, while misinformation is dissemination of false information unknowingly. The European Code of Practice on Disinformation defines disinformation as “verifiably false or misleading information” which, cumulatively, is “created, presented and disseminated for economic gain or to intentionally deceive the public”; and “may cause public harm”, intended as “threats to democratic political and policymaking processes as well as public goods such as the protection of EU citizens’ health, the environment or security”.

The lack of a unanimous definition and the difficulty in distinguishing the line between the presence and absence of the intention to cause harm pose difficulties in providing an adequate response (A/HRC/47/25, 3). Even though disinformation can be harmful, freedom of expression can be restricted only for the protection of one of the legitimate aims of respecting the rights and reputations of others, or for protecting national security, public order and health as set out in Article 10 of the European Convention on Human Rights and Article 19 of the International Covenant on Civil and Political Rights. The right to freedom of expression applies to all kinds of information, and ideas, and states should refrain from imposing restrictions that could chill online speech (A/HRC/47/25, 18).

ETHNO-NATIONALLY AND/OR POLITICALLY BIASED CONTENT

There are no specific mentions of biased content in international documents, but it is extensively discussed in relation to journalistic standards and in communication studies. Propaganda bias, according to Denis McQuail, involves media reporting with the deliberate intention of making the case for a political party or policy without explicitly stating it (Steel 2001, 20). According to the methodology used by the fact-checking platform Raskrinkavanje, biased media reporting favors facts, attitudes, and conclusions that correspond to a particular narrative and could include the selective presentation of facts, where facts that support a certain thesis are emphasized, while facts that do not confirm it are omitted. Biased media reporting includes positive reporting on a particular group and cause, leaving aside facts and attitudes that are critical of that group or cause.¹ This type of content can be factual, thus falling out of the scope of disinformation and is included in this report due to its large presence in traditional and online media and the harm it can cause in Bosnia and Herzegovina. In addition to media content, there are also other manipulative methods that are outside of the scope of regulatory and self-regulatory bodies for the media, such as hidden political advertisements and the use of bots and trolls on social networks for manipulative purposes, for which specific measures should be provided.

ATTACKS, THREATS AND SMEAR CAMPAIGNS AGAINST INDIVIDUALS

Attacks, threats and smear campaigns against individuals is content that targets individuals with the aim to intimidate them or discredit their reputation. It usually targets journalists, opinion makers, representatives of political groups and civil society organizations. It can include malinformation – “genuine information that is shared to cause harm, including revealing private information that is spread to harm a person or his/her reputation” (First Draft 2020). Attacks, threats and smear campaigns – so-called “character assassination” – pertain to both illegal and harmful content and fall under different legislative and self-regulatory mechanisms. Death threats are usually punishable by criminal law, while defamation laws prohibit the publication of false content with the aim to discredit the reputation of a person.

¹ See the methodology used by the fact-checking platform Raskrinkavanje, of the organization Why Not: <https://raskrinkavanje.ba/metodologija>

3. METHODOLOGY

The study was conducted from September 2021 to December 2021 and is based on extensive desk research, semi-structured interviews with stakeholders and consultative meetings with representatives of civil society organizations. Additional desk review was done in March 2022. The aim of the study was to address the following research questions:

- 1) What are the main types and targets of harmful online content in Bosnia and Herzegovina that can thwart democratic processes, hamper citizens' ability to take informed actions and impede reconciliation?
- 2) What legislative, regulatory and self-regulatory frameworks against these types of harmful content exist in the country and what is the effectiveness of their implementation?
- 3) To what extent are the legislative, regulatory and self-regulatory frameworks in Bosnia and Herzegovina against these types of harmful content aligned with international standards?
- 4) How do community guidelines of the most popular social networks in Bosnia and Herzegovina address harmful content?

The desk research included reviews of:

- a) the legislative and regulatory frameworks in Bosnia and Herzegovina to address harmful content, including crime codes, laws against defamation, laws on the protection of public order, codes of the Communications Regulatory Agency of Bosnia and Herzegovina;
- b) available reports of the Communications Regulatory Agency of Bosnia and Herzegovina, the Press Council of Bosnia and Herzegovina, the Institution of the Human Rights Ombudsman of Bosnia and Herzegovina, civil society organizations and social networks;
- c) available court judgements on specific case studies;
- d) self-regulatory tools developed by the online media and the Press Council of Bosnia and Herzegovina;
- e) community guidelines of the most popular social networks in Bosnia and Herzegovina;
- f) international standards and examples of legislative frameworks for combating harmful content in EU countries;
- g) secondary literature on the media sector in BiH, including on media consumption habits and on harmful content.

Fifteen in-depth semi-structured interviews were conducted, with the representatives of the Communications Regulatory Agency of Bosnia and Herzegovina, the Institution of the Human Rights Ombudsman of Bosnia and Herzegovina, the Ministry of Interior of Republika Srpska, the OSCE Mission to Bosnia and Herzegovina, the Interreligious Council in Bosnia and Herzegovina,

and the Personal Data Protection Agency in Bosnia and Herzegovina, and with the representatives of the following civil society organizations: Citizens' Association Why Not, Network for Building Peace, Association for Democratic Initiatives, Helsinki Committee for Human Rights Bijeljina, Adopt Srebrenica, forumZFD, Balkan Investigative Reporting Network BiH, Share Foundation, and Direkt portal, as well as with an independent researcher. The aim of the interviews was to better understand the main types and targets of harmful online content in Bosnia and Herzegovina and the implementation of the legislative, regulatory and self-regulatory frameworks for countering harmful content online.

In addition, three consultative meetings with fifteen civil society representatives were held in: Sarajevo, November 24, 2021, Banja Luka, November 19, 2021 and Mostar, December 13, 2021. The consultative meetings were held with the aim to map the work of civil society organizations in countering harmful content, and to understand their experiences in the use of social networks and their recommendations for strengthening the resilience of societies to harmful content spread online, while protecting freedom of expression.

The following chapters give a background to the information system in Bosnia and Herzegovina and media consumption habits of citizens; outline the scope, types and targets of the five types of harmful content online; and present the legislative, regulatory and self-regulatory frameworks for the five types of harmful content in Bosnia and Herzegovina. There are also segments on how effectively these frameworks are used online, what the major obstacles are in their implementation and to what extent they are aligned with international standards. A specific segment is dedicated to the community guidelines of the most popular social networks in BiH and to examples of frameworks from EU countries. At the end the study provides a conclusion and recommendations for better regulation of harmful content online, taking into consideration international standards and other means to counter harmful content, including raising media and information literacy skills of citizens and strengthening self-regulatory measures.

4. THE INFORMATION SYSTEM IN BIH – AN OVERVIEW

The information system in Bosnia and Herzegovina has changed drastically in the past ten years, following global trends. The ways in which information is produced, consumed and circulated have changed with digital technologies and national legislation has often not been able to keep pace with them (Sokol and Jukić-Mujkić 2021). Most of the radio, print and television stations in the country today have their own online media and accounts on social networks, while some also had to shut down their print versions and have moved completely online (Petković and Hodžić 2019, 7).

Digital transformation also brought the rise of non-professional content producers on the internet, including anonymous websites and diverse YouTube channels. A recent study identified 615 informative online media in BiH, and only 27% of them have impressums on their webpages (Osmančević 2021, 11). Another study showed a rise in the number of YouTubers from Bosnia and Herzegovina that are popular and followed by young people, producing content on beauty, fashion, and lifestyle (Sokol and Alibegović 2021, 13).

Digital transformation changed the media consumption habits of citizens. Internet penetration and the number of social media users in Bosnia and Herzegovina have been on a constant rise in recent years. According to the Communications Regulatory Agency of Bosnia and Herzegovina, in 2020 internet penetration was 94.95% which is a drastic increase compared to, for example, 2011, when it was 55% (CRA 2021b, 21). The use of the internet and social media platforms as news sources has been on the rise, particularly among youth and young adults and social networks and web portals are the main source of information for young people (Sokol and Alibegović 2021, 9-10). Facebook has been the most popular social media platform, followed by Instagram – in BiH there are over 1.8 million Facebook users, even though Instagram is becoming more popular among younger generations.² Online media also rely on social media platforms to reach their audiences and journalists increasingly use social media for the dissemination of the content that they produce, but also to search for information and contact potential interviewees (Đelilović 2021, 8-9).

A recent study on the media consumption habits of adults in BiH showed that nine out of ten respondents use the internet (87%) (Hasanagić et al 2021, 5).³ They are most likely to access the internet via a smartphone (93%) and spend 19 hours a week online, which is on average 2 hours and 42 minutes per day. Younger respondents aged 18-24 tend to spend on average 4 hours and 2 minutes per day online (Hasanagić et al 2021, 5). Another study on media and youth demonstrated that young people spend most of their free time on the internet, most often scrolling and consuming information that they find interesting (Hodžić and Sokol 2019, 50-51).

² Data obtained from [statista.com](https://www.statista.com).

³ The study was conducted based on 1400 respondents.

Despite the rise of non-professional content producers, the number of traditional media in the country still remains significant⁴ and television is still the most important and the most trusted source of information among the general public (Sokol 2021b, 6). However, digital transformation and the consequences it has had on media consumption habits and the advertising market have altered the production of media and news content. Many local and regional advertisers have moved their ads from traditional media to both local and global digital platforms – also due to lower prices – impacting the already modest advertising market in BiH, whose revenues have been falling on a constant basis (Petković and Hodžić 2019, 9).⁵ Market pressure and altered media consumption habits demand fast news production, the consequences of which are most often copy-pasted and sensationalistic news content with the aim to attract readers' attention and thus more revenue, and often lacking sources and verified information.

In addition to low quality media content, particularly online, and lack of in-depth constructive media reporting, political influence on the media remains significant. Overall, most of the media lack sufficient revenues and capacities for independent and professional journalism.⁶ Through non-transparent and arbitrary financing patterns from public budgets and ownership, political and business groups meddle in editorial affairs, which is visible in ethno-nationally and politically biased media content (Hodžić and Sokol 2017).⁷ The media are fragmented along ethno-national and political party lines, which is manifested in the different interpretations of controversial issues, and stark differences in the presentation and selection of topics, collocutors and standpoints. Ethno-national and political bias in media content is primarily visible in selective reporting, including a focus on and a lack of criticism of the leading parties of a particular ethnic group and an overall emphasis on the issues involved in the agenda of that particular ethnic group and political party (Sokol 2019, 34). Fragmentation along ethno-national and political party lines is also manifested in media trust and the attitudes audiences have towards different media. A study on media trust showed that audiences follow and trust the media that show their group – ethno-national or political – in a positive light. The study also showed that the main reasons for distrust are related to the perception that most media outlets are under political influence and that they spread disinformation – overgeneralizations that can harm the work of professional and independent media (Sokol 2021, 23–25).

⁴ In the country there are around 107 television stations, 157 radio stations, 7 news agencies, 8 dailies, and 181 different publications and magazines. There are 65 public radio and 15 public TV stations, which function as public enterprises and are funded directly from the public budgets of cantons, towns and municipalities (information retrieved from the registers of the Communications Regulatory Agency and the Press Council BiH).

⁵ The overall value of the advertising market, based on estimates from marketing agencies, was around €22.9 million (\$27 million) in 2020 (Sokol and Jukić-Mujkić 2021, 5).

⁶ The financial consequences of the pandemic impacted the media sector and advertising revenues sharply decreased, especially during the first three months of the pandemic in 2020 (Sokol and Jukić-Mujkić 2021, 5).

⁷ For example, municipalities, towns, cantons and entities give subsidies and grants to local media outlets on a yearly basis, but such allocations lack precise criteria, expert commissions and protection mechanisms from political interference in editorial policies. These government agencies, as well as public companies, sign commercial agreements with private media to follow their activities, but the legitimacy of such agreements is questionable (Sokol and Hodžić 2017).

Despite the increase in the use of digital technologies, research shows that the level of media and information literacy among citizens is still considered poor. Only a quarter of BiH respondents (24%) in the study on media consumption habits of adults believe in their ability to identify false information in the media and information-communications services, and over half who use the internet (56%) do not perform any fact-checking of information they encounter online (Hasanagić et al 2021, 7–8). Another study showed that young people in BiH are aware of problematic media content but were mostly not able to identify problematic points in articles containing propaganda, biased reporting and disinformation (Hodžić and Sokol 2019, 71–74). A study that included research on young people's perceptions on media freedoms demonstrated that their attitudes towards media freedoms are mostly related to the courage of journalists to speak the truth despite political and ethno-national pressure, and they show a tendency to support punitive measures for hate speech and other problematic content in user generated comments (Sokol and Alibegović 2021, 25–28). Such analyses point to the need of media and information literacy education, which would also include education on international standards regarding freedom of expression.

Regardless of the needs posed by the changes in the information system in Bosnia and Herzegovina, policies and legislation have not been keeping pace with these trends. Media and information literacy is still not included as a subject of the school curricula throughout the country or at various levels of education, although the subject Digital World was introduced in the second grade in primary schools in Republika Srpska (RS) in 2021 and the RS government plans to introduce the subject Media Literacy in some secondary schools in 2022 (Stokić 2021b). The Communications Regulatory Agency, non-governmental and international organizations continue to work on promoting and advancing media and information literacy, but systematic solutions are needed for the educational system and for the inclusion of both urban and rural areas and different age groups.

Bosnia and Herzegovina has a functioning regulatory and self-regulatory system for the media. The Communications Regulatory Agency of Bosnia and Herzegovina can issue fines and warnings to television and radio stations for breaches of its codes and rules such as for breaching the rules on the principles of fairness and impartiality, right to reply, and privacy and hate speech, but it mostly acts on citizens' reports and does not conduct regular monitoring. Moreover, increasing concerns over its politicization have been affecting its work and lowering the trust of the media and the public (Bubonjić 2022a). The Press Council of Bosnia and Herzegovina, on the other hand, self-regulates the work of online and print media but its activities are restricted to mediation and non-binding decisions about media content that violates the standards of the Press Code of Bosnia and Herzegovina, including, for example, hate speech. In 2021 the Press Code of Bosnia and Herzegovina was amended, among other things, to expand editorial responsibility to the overall content in online media, including user generated comments, and introduce, among other things, provisions on disinformation and the use of information technologies (Išerić 2022). The number of user complaints to the Press Council of Bosnia and Herzegovina has been rising in recent years and a large number of complaints have been resolved by mediation, but the effect of the self-regulatory system is particularly limited regarding anonymous portals.

A number of laws that pertain to the work of the media, freedom of expression and public communication exist in the country, but some of their provisions are not entirely aligned with international standards and their implementation remains partial.

Freedom of expression is guaranteed by the Constitution of Bosnia and Herzegovina, the Constitution of the Federation of Bosnia and Herzegovina and the Constitution of Republika Srpska. The Constitution of Bosnia and Herzegovina stipulates that Bosnia and Herzegovina and both entities shall ensure the highest level of internationally recognized human rights and fundamental freedoms and that the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Protocols shall apply directly and have priority over other law (Article 2, II).

The criminal codes of the state and the entities criminalize public incitement to national and racial hatred, discord and hostility, and laws of the entities and the Brčko District on protection against defamation regulate civil liability for harm caused to the reputation of a natural or legal person by the making or disseminating of false facts. Both laws, as shall be demonstrated, are not entirely aligned with international standards and the law against defamation has been used as a way to intimidate journalists (Radević 2021, 11-12). The right to free access to information is protected by the state and entities' laws on the right to free access to information, but some of their provisions do not meet international standards, particularly in terms of proactive transparency, and citizens and journalists still face difficulties in obtaining information. In 2021 the Ministry of Justice of Bosnia and Herzegovina prepared a draft of a new law on access to information but it was criticized by civil society organizations for containing a large number of exceptions (Hasečić 2021; Učanbarlić 2021).

The lack of media ownership transparency has been of huge concern and the country still has not adopted a law to tackle this issue. In its yearly progress reports for Bosnia and Herzegovina, including 2021, The European Commission has been emphasizing that Bosnia and Herzegovina should adopt legislation on media ownership transparency and criteria on public advertising (European Commission 2021, 27). The reports also point to the need of ensuring the protection of journalists and establishing a systematic institutional follow-up on threats and violence against them, as well as ensuring the financial sustainability and political independence of public broadcasters (European Commission 2021, 27). However, in addition to the "old issues" faced by traditional media and professional journalists, such as political pressure, lack of media transparency, financial instability and attacks on journalists and different violations of their rights, there is a whole new set of emerging issues posed by the internet and social networks that could curb freedom of expression and even thwart democracy, but which have so far not been researched or recognized by civil society organizations, the media and authorities. In addition to disinformation and hate speech on the internet, over which civil society organizations and the self-regulatory body have been active in recent years, hidden political advertisements, algorithmic bias, the use of bots and trolls for manipulative purposes, harmful content, and non-transparent and non-legitimate content removal on social networks are issues that also pose threats to freedom of expression. However, so far, based on the conclusions of the consultative meetings, they are only vaguely understood – due also to the lack of transparency of these

processes – and there is a lack of expertise and awareness among the authorities and civil society organizations. The research of Why Not, for example, demonstrated that stakeholders, including government representatives and civil society, have low awareness about targeted disinformation campaigns on the internet (Cvjetičanin et al 2019, 7). This points to the need for sharing experiences and raising knowledge and expertise about digital rights and harmful content online in Bosnia and Herzegovina.

5. HARMFUL CONTENT: TYPES, CONSEQUENCES AND TARGETS

Hate speech and hate narratives, denials of war crimes and glorification of war criminals, ethno-nationally and/or politically biased reporting, disinformation, attacks, threats and smear campaigns against individuals are types of harmful content that can hinder democratic processes, reconciliation and processes of facing the past, and can hurt and raise hatred towards groups and individuals and even in some cases retraumatize citizens, particularly victims of war or families of war victims. According to the interviewees and available studies and reports of civil society organizations that have reporting mechanisms or monitor the presence of certain types of content such as hate speech, mostly in user generated comments (the Press Council of BiH), denials of genocide (Srebrenica Memorial Center), disinformation (fact-checking platform Raskrinkavanje), and attacks and threats against journalists (Association BH Journalists), their presence remains significant. The sources of harmful content on the internet, depending on the type, are manifold:

- Online media and anonymous websites for financial and political gains;
- Political and state actors, who utilize the public space and the internet to advance their political agendas;
- Online media and traditional media that advance certain ethno-national or political agendas, especially in relation to ethno-national and politically biased media reporting;
- Individuals, groups and channels on social networks for commercial and political gains or advancing certain ideologies;
- Radical websites and groups on social networks that spread hatred towards others.

Interviewees agree that the scope of such content online is worrying. A recent study demonstrated that 35% of respondents said that they have been exposed to potentially harmful content in the media and information-communication environment, including hate speech (48%) and discriminatory content (32%) and only a fifth (19%) reported being exposed to inappropriate content (Hasanagić et al 2021, 7-8). Interviewees said that harmful content online in Bosnia and Herzegovina involves diverse topics but often pertains to mutually exclusive memories of the past and conflicting ethno-national narratives.⁸

The consequences of harmful content, according to the interviewees, are extensive and include retraumatization of the public, political apathy, and decisions not to abide by, for example, epidemiological measures.⁹ Interviewees state that such content targets diverse groups, such as minority groups, migrants and refugees,

⁸ Mirjana Ćuskić, Helsinki Committee for Human Rights, interview; Sunita Dautbegović – Bošnjaković, forumZFD, interview, Bekir Halilović, Adopt Srebrenica, interview.

⁹ Tijana Cvjetičanin, fact-checking platform, Raskrinkavanje, interview; Maida Zagorac, Network for Building Peace, interview.

returnees and women¹⁰ and can particularly influence young people, who are online the most,¹¹ and those who lack media and information literacy skills. Studies on media reporting on migrants and refugees point out that stereotypical and often xenophobic media reporting on migrants and refugees in BiH can raise hatred towards people on the move in BiH (Sokol 2021a, 7-11; Adilagić 2019). A recent study of the association Why Not showed that misconceptions about the Covid-19 pandemic, which were very popular on the internet, have gained significant foothold in society and are more prevalent than other conspiracy theories (Karadžuz et al 2022, 5). A publication of forumZFD states that through ethno-nationally biased content the media perpetuate divisive rhetoric, and do not allow society any respite from, or a chance to reflect on their traumatic pasts (ZFD 2019). This all points to the importance of monitoring and analyzing such content and finding adequate ways to counter it, deal with its root causes and devise prevention mechanisms.

5.1. HATE SPEECH AND HATE NARRATIVES

Hate speech and hate narratives are present on the internet in user-generated comments and social networks. There are a number of websites of far-right extremist groups and specialized webpages that propagate the supremacy of their own group, promote hatred toward other groups and/or glorify war crimes and war criminals, some of which are even present on social networks.¹² The 2018 study of Aktis and Mediacentar that was based on the automated scraping of comments on different platforms and websites in BiH demonstrated that hate speech in the online space is prolific and widespread. The researchers also concluded that YouTube comments sections particularly contained violent instances of hate (Aktis 2018, 35). The researchers stated that public forums such as *StormFront* and *HercegBosna* provide “elaborate hate speech that is generally mixed with right-wing extremism and drawn-out historical elaborations” (Aktis 2018, 35). They also concluded that hate speech is generally spread by a wide range of ‘ordinary’ members of the public, as well as some more specific groups. From the data that they collected, most of such users seem to be male, and a significant proportion are also located in the diaspora. Interviewees mention, that, even though related to the war, hate speech is also being spread by young people who were born after it.¹³

¹⁰ Dajana Čelebić, Association for Democratic Initiatives, interview.

¹¹ Mirjana Ćuskić, Helsinki Committee for Human Rights, interview; Sunita Dautbegović – Bošnjaković, forumZFD, interview.

¹² Such groups were identified by the Balkan Investigative Network BiH. They include: the Movement of Serbian Chetniks of Ravna Gora (*Ravnogorski četnički pokret*) and the Bosnian Movement of National Pride (*Bosanski pokret nacionalnog ponosa*). Others include the webpage and social media accounts of Despotovina.info and the webpage of Antimigrant.ba. A report from 2018 listed a number of such groups, some of which are still active: Naš Haber and Prosječna Komunjara (Aktis 2018).

¹³ Maida Zagorac, Network for Building Peace, interview.

While such content can function as a mechanism for public venting, hate speech harms the culture of public discussion, normalizes negative sentiments towards the “other” and can lead to real harm to individuals belonging to that group. The targets of hate speech in Bosnia and Herzegovina are mostly ethno-national groups, migrants and refugees, journalists, minorities and women. The European Commission’s Report 2021 points out that minority returnees are among the most common targets of ethnically driven hate speech and hate crimes (2021, 32) and that hate speech on religious grounds continues to occur (2021, 29).

A research study on antisemitic discourse in the Western Balkans showed instances of antisemitism in comments sections on Facebook and online media under different articles related to Israel, Holocaust remembrance or foreign affairs (Sokol and Memišević 2020, 58).¹⁴ The Sarajevo Open Centre registered 83 instances of hate speech against the LGBTIQ community in 2018-2019, mostly in connection to the first Sarajevo Pride in 2019, which were found in user generated content on internet portals (Blažević 2019, 20).

Even though illegal hate speech is mostly present in user generated comments, the mainstream media, through selective reporting and also publishing of inflammatory statements of political leaders, are also sources of hate narratives. For example, media reporting in recent years has been stigmatizing migrants and refugees, portraying them as perpetrators and dangerous persons (Adilagić 2018). Such portrayals, which also come from politicians and public figures, can increase intolerance and hate towards migrants and refugees and increase the amount of illegal hate speech in user-generated content (Sokol 2020a, 7-11). What remains highly problematic, and not dealt with, are hate narratives that come from political figures, mostly against other ethno-national groups, and which could have high impact, especially due to their influence in the society.

5.2. DENIALS OF WAR CRIMES AND GLORIFICATION OF WAR CRIMINALS

Another type of harmful content that highly impacts the process of facing the past and reconciliation are instances of condoning, denying, gross diminution or attempts to justify the crimes of genocide, crimes against humanity and war crimes established by final judgments. Such instances have been present not only in user generated comments, but also in the mainstream media, and come from political figures, public institutions, and government representatives (Gačanica and Finkeldey 2019, 12-13). Gačanica and Finkeldey state that denials or minimizations are visible in various places of political activity and as a result are imprinted in many areas of everyday life: decorations of war criminals; rehabilitation of crimes, establishment of associations and organizations bearing the names of war criminals; the existence and the legal activities of fascist organizations, all of which are visible

¹⁴ In this study around 1400 articles and comments related to antisemitism in BiH were examined in selected time periods during 2020. Out of these there were around 101 instances of antisemitism, most of which were on Facebook.

on the internet. The 2021 Srebrenica Genocide Denial Report demonstrated that genocide denial, including the glorification of war crimes and criminals, remains widespread in BiH and the region (Cvjetičanin et al 2021). Researchers identified 234 instances of genocide denial online in Serbia (142), BiH (60), and Montenegro (19). The three most common rhetorical tactics used in genocide denial are: disputing the number and identity of victims, conspiracy theories which challenge the rulings and integrity of international courts, and nationalist historical revisionism. The report also concluded that online portals accounted for a great deal of denial – through publishing original content or copy-pasting it from other sources (Cvjetičanin et al 2021, 4-5).

The monitoring by the Srebrenica Memorial Center has shown that since the amendments to the Criminal Code of BiH that prohibit condoning, denying, gross diminution or attempts to justify the crimes of genocide, crimes against humanity and war crimes established by final judgments entered into force in July 2021, the denial of genocide has been carried out less explicitly and directly in BiH and that the number of such acts has decreased. However, it was also stated that politicians continued to deny genocide, mostly in the media in Serbia, whose statements were then transmitted by the media in BiH (Srebrenica Memorial Centre 2021). The database Mapping Hate, of the Balkan Investigative Reporting Network, shows that high-ranking politicians continue to deny war crimes and their statements are published by the mainstream media without any critical stance.¹⁵ Such content, as pointed out by the interviewees, can hurt and retraumatize those who survived the war and the families of war victims, and also impact reconciliation processes.¹⁶

5.3. ETHNO-NATIONALLY AND POLITICALLY BIASED MEDIA REPORTING

The third type of harmful content online identified in the research is ethno-nationally and/or politically biased media reporting, particularly in relation to past events, ethno-national discourse and politics. As already mentioned, the media sector in Bosnia and Herzegovina is mainly divided along ethno-national and political party lines, and influenced by political party groups through financing patterns and ownership (Hodžić and Sokol 2017). Such divisions are reflected in the breaches of professional journalistic standards and biased reporting – presenting one side of the story and neglecting others. The most pronounced is selective reporting on current affairs, including negligence of and/or a critical stance towards other ethno-national groups and their representatives, and favoritism towards the interests of one's own ethno-national group and towards its political leaders. Biased reporting is present in both traditional media and online media, and the largest number of complaints that the Communications Regulatory Agency of BiH receives on a yearly basis, and the largest number of breaches it confirms,

¹⁵ See the map: <https://mapiranjemrznje.detektor.ba/>

¹⁶ Maida Zagorac, Network for Buidling Peace; Mirjana Ćuskić, Helsinki Committee for Human Rights, interview; Sunita Dautbegović – Bošnjaković, forumZFD, interview.

pertain to fair reporting and impartiality.¹⁷ Biased, propagandistic reporting is highly present online and especially in election years – when mainly anonymous portals mushroom in pre-election periods and publish favorable news about certain political parties (Blagovčanin 2019; Turčilo 2020). Even though this reporting is omnipresent in the media in BiH there is a lack of a monitoring mechanism or a comprehensive in-detail research study that would show the overall scope, unravel its sub-narratives and explain its nuances and main sources. Research studies have so far focused on certain case studies or only pointed to the fact that even with a superficial look one can see opposingly different interpretations of certain issues in the media in BiH along ethno-national and political party lines. Articles published by the watchdog media platforms Media.ba and Analiziraj show instances of such reporting, particularly in regard to the three public service broadcasters at the state and entity levels (BHRT, FTV, RTRS) in BiH and the most popular online media. One research study pointed out that ethno-national bias in the media content in the Birač region is primarily visible in the selective coverage of the commemoration of war victims, sometimes involving the glorification of the armies and heroes of one ethno-national group, and partial and/or one-sided interpretations of war events (Sokol 2019, 34). Another article published by Mediacentar showed that biased content is also present in media reporting on international news, such as on the war in Ukraine, in which the media in Republika Srpska relativized the role of Russia and its responsibility (Bubonjić 2022b). Such media reporting polarizes the public, creating groups of audiences that follow and trust the media that show their group in a mainly positive light (Petković, Bašić Hrvatinić and Hodžić 2021b).

5.4. DISINFORMATION

Disinformation has proliferated in recent years, especially in relation to political propaganda, COVID-19 and the war in Ukraine. A study of the association Why Not from 2019 showed that false or misleading media reporting is highly present in BiH and the region, most often in the form of intentionally fabricated false information, and 60% was political in nature (Cvjetičanin et al 2019, 7). The research pointed to two sources of disinformation: “opportunistic disinformers, who operate mostly through anonymous websites and social media, with financial gain as their primary motive; and political and state actors, who utilize both public and commercial media outlets to spread disinformation to advance their political agenda” (Cvjetičanin et al 2019, 7). Disinformation, as pointed out by the study, comes from both local and foreign actors, and much also through regional disinformation hubs and media from neighboring countries that are both sources and redistributors of disinformation (Cvjetičanin et al 2019, 7). What is particularly worrying is that in addition to anonymous websites, which are the type of media that appear as the main source and redistributor of disinformation, public media are the largest individual sources of disinformation in BiH (Cvjetičanin et al 2019, 8).

¹⁷ Azra Maslo, CRA, interview.

The fact-checking platform Raskrinkavanje debunked hundreds of articles containing disinformation about the pandemic of COVID-19, which circulated on the internet and social networks (Krupalija et al 2021). Disinformation coupled with the poor communications of the health authorities and the lack of an official government campaign on the importance of immunization further undermined the public's health response. In November 2021, despite the availability of vaccines, only 20% of the population of Bosnia and Herzegovina had been vaccinated (Augustinović and Milojević 2021). Another study pointed to wrong beliefs about the Covid-19 pandemic and called for a multisector approach in countering such narratives, including building trust and education (Karađuz et al 2022). Studies also demonstrated that disinformation targets minority groups, including migrants and refugees, women and the LGBTIQ community (Krupalija et al 2021). What is particularly problematic is that journalists lack the capacities to fact-check content, and they find it especially difficult to fact-check the statements of officials and official sources (Sokol and Jukić-Mujkić 2021, 4).

5.5. ATTACKS, THREATS AND SMEAR CAMPAIGNS AGAINST INDIVIDUALS

The final type of harmful content identified in this study pertains to attacks, threats and smear campaigns mostly targeting individuals on the internet – professionals, journalists, media workers and representatives of civil society – with the aim to intimidate and/or discredit them. Terms that are also used for such harmful content are cyber/online violence and online harassment, which is characterized by anonymity and concealment of the identity of the perpetrator, the large availability of the target, access to the victim via the internet, and disinhibition, which includes a lack of restraint and self-control due to anonymity which would often not happen in the offline environment (Gačanica and Arnautović 2018, 11). Data of the Association of BH Journalists show that increasing numbers of attacks and threats against journalists are made online (Boračić-Mršo 2022). For example, in 2021 the Association registered 70 attacks, threats and other forms of violations of rights of journalists and 70% of them happened online (Boračić-Mršo 2022). Targets of such harmful content are especially journalists critical of the ruling regimes and ethno-national narratives, those who work with migrants and refugees, representatives of civil society and of political opposition, female journalists and lately also public figures who advocate immunization in relation to COVID-19. Such campaigns include user generated insults and threats, derogatory language and defaming of individuals, misogynist insults in relation to women, pictures and videos of the individuals shared on the internet and insults and threats from users through private messages. Gender is often used as a basis for discrimination and attacks – in the online space it ranges from insults based on physical appearance to explicit threats made publicly (comments, statuses) and privately (e-mails, social media messages) based on gender (Gačanica and Arnautović 2018, 17-18). However, apart from the registry of attacks and violations of journalists' rights of the Association BH Journalists, there is no overall data about registered attacks, threats and smear campaigns against individuals, particularly, civil society representatives, and many of these cases do not get registered.

6. REGULATION OF HARMFUL CONTENT

Countries around the world have been trying to counter harmful content online in ways that would not curb freedom of expression. Article 19 of the Universal Declaration of Human Rights states that “everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers”. According to Article 19 of the International Covenant of Civil and Political Rights, freedom of expression carries special duties and responsibilities and may be subject to certain restrictions, but only as provided by law and that are necessary for respect of the rights or reputations of others and for the protection of national security or of public order or of public health or morals. Similarly, Article 10 of the European Convention on Human Rights states that freedom of expression may be subject to formalities, conditions, restrictions or penalties that are necessary in democratic societies, “in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary”.

The report of the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression on online content regulation (A/HRC/38/35, 4) states that limitations of freedom of expression must meet the following conditions:

Legality: they must be adopted by regular legal processes and limit government discretion in a manner that distinguishes between lawful and unlawful expression with “sufficient precision”.

Necessity and proportionality: States must demonstrate that the restriction imposes the least burden on the exercise of the right and actually protects, or is likely to protect, the legitimate State interest at issue.

Legality: any restriction, to be lawful, must protect only those interests enumerated in article 19 (3): the rights or reputations of others, national security or public order, or public health or morals.

Freedom of expression is guaranteed by the Constitution of Bosnia and Herzegovina (Articles 2, 3, h), the Constitution of the Federation of Bosnia and Herzegovina (II, Article 2, I), and the Constitution of Republika Srpska (II, 26, 32, 34). The Constitution of Bosnia and Herzegovina stipulates that Bosnia and Herzegovina and both entities shall ensure the highest level of internationally recognized human rights and fundamental freedoms (Article 4, I) and that the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Protocols shall apply directly and have priority over other law (Article 4, II).

In Bosnia and Herzegovina there are legal, regulatory and self-regulatory frameworks for countering harmful content: hate speech and hate narratives, denials of war crimes and glorification of war criminals, ethno-nationally and politically biased reporting, disinformation, and threats, attacks and smear campaigns against individuals. However, most of them are fragmented across different laws, regulations and self-regulatory mechanisms, are not entirely aligned with international standards and their implementation remains inconsistent and particularly problematic for the online sector. The following chapters are

dedicated to legislative, regulatory and self-regulatory frameworks for the five types of identified harmful content. The sections also include segments on how effectively these frameworks are used online, what the major obstacles are in their implementation and to what extent they are aligned with international standards.

6.1. REGULATION OF HATE SPEECH

Hate speech is only partially regulated in Bosnia and Herzegovina and the country lacks legislation on hate speech that is aligned with international standards. Particular forms of hate speech are prohibited by criminal laws, and the Election Law prohibits political candidates from using language which could provoke someone to violence or spreading of hatred during pre-election campaigns. Hate speech is also prohibited by the codes of the Communications Regulatory Agency, which, however, apply only to television and radio stations. The main obstacles in regulating hate speech according to international standards is the lack of proper legislation and the lack of knowledge of international standards among judges and prosecutors. Not all cases of hate speech are illegal and a large part of hate narratives, even in political discourse, remain in gray zones for which there are no adequate counter measures.

6.1.1. Criminal codes

Public incitement to national, racial and religious hatred, discord and hostility is criminalized in Bosnia and Herzegovina. The Criminal Code of Bosnia and Herzegovina (145a, 1), the Federation of Bosnia and Herzegovina (163,1) and the Brčko District (160, 1) prohibit public incitement and inflaming to national, racial and religious hatred, discord or hostility among the constituent peoples and others living in Bosnia and Herzegovina, the Federation and the Brčko District, respectively.

The three criminal codes are limited, however, to the prohibition of inciting and inflaming hatred on national, racial and religious grounds, but do not include other protected categories including skin colour, gender, sexual orientation, disability and other grounds which are part of international standards (Blažević 2019, 10). These legal provisions only sanction public incitement and inflaming of national, racial and religious hatred, discord or hostility towards persons living or residing in BiH (Criminal Code of BiH) and towards persons living in the Federation and the Brčko District, respectively (Criminal Code of the Federation of BiH and the Criminal Code of the Brčko District). According to Filipović (2019, 6) if public incitement of national, racial or religious hatred, discord or hostility refers to persons who do not live or reside in BiH at the time the act was committed in BiH, the perpetrator cannot be punished.

The Criminal Code of Republika Srpska includes other protected categories, such as skin colour, sex, sexual orientation and disability and lists the means by which the offence can be committed, including a computer system and a social network, which is not specified in the other three criminal codes (Omerović and Hrustić 2020, 37). The Criminal Code of RS prescribes a “punishment for whoever, by using the press,

radio, television, a computer system or a social network, at a public gathering or at a public area or otherwise publicly calls for, incites or inflames or makes available to the public leaflets, images or any other materials that call for violence or hatred against a certain person or group on account of their national, racial or religious or ethnic affiliation, skin colour, sex, sexual orientation, disability, gender identity, origin or other properties” (Criminal Code of RS, Article 359, 1).¹⁸

Overall, cases of hate speech processing in Bosnia and Herzegovina are rare, particularly for hate speech on the internet, and court decisions are inconsistent. Interviewees state that there is a misunderstanding about the meaning of hate speech and freedom of expression even among practitioners and particular difficulties exist in defining what is considered public space.¹⁹ According to one study, from 2004 to 2019, 27 judgements were reached in BiH for criminal offences of inciting national, racial and religious hatred, discord and intolerance against 26 persons, of which in 12 cases the person was found guilty (Ferhatović and Trlin 2019, 144-145).²⁰ Out of these, only three criminal offences took place via the internet - two via Facebook profiles and one via a website. The number of initiated cases is rather small in relation to the large presence of hate speech on the internet and the large number of acquittals (Ferhatović and Trlin 2019, 144-145). Cases of the processing of hate speech have so far mostly been initiated against ordinary citizens, while there have been no cases in which those in power were sanctioned.²¹

¹⁸ In addition to these provisions, there are other provisions in criminal law that can be used to sanction hate speech but they do not refer to the internet. The criminal laws of the Federation and the Brčko District prescribe a fine or imprisonment (Criminal Code of FBiH, 363, Criminal Code of the Brčko District, 357) for whoever, by severely violating the standards of professional conduct of media and journalists, uses inciting or hateful language that obviously calls for or incites violence or national or ethnical conflicts, and thereby brings public peace and order into danger, but only via television and radio stations (Filipović 2019). The Code of Republika Srpska prescribes imprisonment for anyone who behaves in such a way or shouts slogans or carries placards at a sporting event or public gathering as to provoke national, racial, religious or some other kind of hatred or intolerance based on some discriminatory grounds, which resulted in violence or a physical altercation with participants (Criminal Code of RS, 363, 1).

¹⁹ Halisa Skopljak, interview, OSCE Mission to BiH

²⁰ After 2019 there were a number of other cases where persons were sanctioned for hate speech online. For example, the Municipal Court in Ljubuški issued a final verdict against a member of the Armed Forces of BiH, for a message sent via Facebook to the writer Martina Mlinarević, the current ambassador of Bosnia and Herzegovina in the Czech Republic. He called her a “traitor” of her people and an “enemy of the Croats” (BH Journalists 2021).

²¹ Halisa Skopljak, OSCE mission to BiH, interview.

Case 1

The author of the Antimigrant.ba portal was acquitted by a first instance verdict before the Court of Bosnia and Herzegovina on charges of spreading national, racial and religious hatred and intolerance.²² The court found that the statements were addressed directly to migrants, and that the constituent peoples of BiH were only mentioned in connection with the statements, meaning that the author did not publicly incite national, racial and religious hatred, discord or hostility among the constituent peoples and others living in Bosnia and Herzegovina (Coalition 2021). The verdict also stated that the analysis of the evidence showed that the statements about migrants did not refer to any particular nation, race, religion or other specific group and that the statements were within the framework of a political, journalistic, free narrative, protected under freedom of speech (Coalition 2021). The Coalition for the Fight Against Hate Speech and Hate Crimes of BiH²³ reacted expressing public disagreement with the court, stating that it shows the urgency for changes in criminal legislation in Bosnia and Herzegovina, which would include recognizing and sanctioning hate speech and hate crimes against all vulnerable groups in society, but also the necessary education of judges and prosecutors to recognize and adequately sanction hate speech in accordance with the standards of the Council of Europe (Coalition 2021).

Case 2

In 2012, the Court of the Brčko District issued a one-year prison sentence, suspended for a period of three years, for inciting national, racial and religious hatred, discord or intolerance due to a series of posts on a publicly accessible online forum, in which the accused made statements on the manner in which Bosniak citizens of the Brčko District of BiH should act in the event of war and secession of Republika Srpska (European Court of Human Rights, 2016, 2). The accused complained that the forum could not be considered a public place, and that he did not incite hatred but only expressed his opinion, and that the case became known only after the media reported on it. In 2012, the Court of Appeals of the Brčko District upheld the Basic Court's ruling, arguing that the content of the forum was accessible to everyone and that the statements were not an expression of free opinion, but a very inappropriate form of dialogue advocating a strategy of behaviour towards one ethnic group (European Court of Human Rights, 2016, 3). The Constitutional Court and the European Court of Human Rights dismissed the appellant's appeals.

Case 3

The BiH Constitutional Court overturned a ruling by the Livno Cantonal Court dismissing an appeal against a basic court ruling finding a person guilty of inciting

²² Fatmir Alisphaić was charged with continuously publishing texts and videos that abound in statements that encourage and spread hatred towards migrants and the migrant population, as well as between the constituent peoples of BiH on several occasions during 2019 and 2020, through the online portal Antimigrant.ba and on social networks (MCOOnline 2021).

²³ The Coalition consists of 13 civil society organizations from BiH.

ethnic, racial and religious hatred by posting a photo on Facebook. On the photo, which contains the flag of Bosnia and Herzegovina, two dragons are flying towards the statue of Jesus Christ in Rio de Janeiro with the text "Let's tear it up". The Court in Livno concluded that the person publicly provoked religious hatred and intolerance towards Croats, members of the Roman Catholic faith, who live in BiH, by exposing the destruction of religious symbols, even though the appellant said that he published the photograph on the day the football team of Bosnia and Herzegovina was playing a qualifications game for the World Football Championship in Brazil. Among other things, the Constitutional Court cited the case law of the European Court of Human Rights that the expression of an individual to a small and limited group of Facebook users does not have the same weight as a statement published on a mainstream website (BiH Constitutional Court 2016). The Court also stated that only when the photo was published by local online media with a very suggestive title from the editor (Who is protecting us, a policeman is spreading hatred on Facebook), did the photograph become available to a large audience.

6.1.2. Sanctioning language that incites to violence during election periods

Through the Election Law it is possible to fine a political candidate who "uses language which could provoke or incite someone to violence or spreading of hatred, or publishes or uses pictures, symbols, audio and video recordings, SMS text messages, internet messages or other materials that can have such effect" (Article 19.9, paragraph j). This law also prohibits the conduct of an election campaign by way of electronic or printed media "where the contents are stereotypical and offensive against men and/or women or which encourages any stereotype and offensive behavior on the grounds of gender or any humiliating attitude against the members of different genders" (Article 16.14, paragraph 3). These provisions are an important tool since hate speech has been particularly present during election periods with the aim of attracting voters, and they were used to sanction political subjects for spreading hate speech on the internet and social networks. (Dragičević 2019, 10). For example, during the 2020 local elections, the Coalition Pod lupom recorded 53 cases of inappropriate speech, or content or speech that could incite religious or national intolerance, or intolerance on other grounds. According to the report, most cases relate to extremely inappropriate speech or intolerance based on political affiliation, especially on social networks (Coalition Pod lupom 2021, 37). What is problematic, however, is that these provisions do not provide specific definitions of hate speech, due to which free political speech could be infringed, and they apply only if such acts occur within 30 days before the date of the elections. Research showed that political parties and candidates start promoting their candidacies before the election campaign period, especially online and on social networks, and hate speech should be sanctioned even before the official start of the campaign (Sokol and Hasečić 2020).

Case 1

The Central Election Commission fined one political candidate in 2018 for using language on his Facebook profile that could lead or incite someone to violence or the spreading of hatred. The Court of BiH rejected the appellant's appeal, not accepting the allegations of the appeal that the views expressed on the candidate's private profile were not intended for the general public (Court of BiH 2018, 2).

Case 2

The Central Election Commission annulled the candidacy of one political entity in 2020 local elections for broadcasting a video on YouTube in which members of the constituent peoples and minorities in BiH are portrayed in a stereotypical and insulting manner (Coalition Pod lupom 2021, 37). The Court of BiH annulled this decision because the video was broadcast before the start of the official campaign (Coalition Pod lupom 2021, 37).

6.1.3. Prohibition of hate speech on television and radio stations

Hate speech is prohibited by the Code on Audio-Visual Media Services and the Code on Radio Media Services of the Communications Regulatory Agency of Bosnia and Herzegovina. Article 4 of the Code states that it is prohibited to humiliate, intimidate or incite hatred, violence or discrimination against a person or group on the grounds of sex, race, ethnicity, nationality, religion or belief, disability, special needs, age, sexual orientation, social origin or on the basis of any other circumstance, which has the purpose or consequence of preventing or endangering any person's recognition, enjoyment or exercise on an equal basis of their rights and freedoms (Article 4). The CRA Commercial Communications Code also does not allow hate speech according to the same definition given above (Article 3, 4b).

The regulatory tools apply only to television and radio stations and the CRA cannot sanction content online. During 2021, the CRA prepared draft amendments to the rules for audio-visual media services and radio media services, according to which the responsibility of licensees for television and radio broadcasting could be expanded to content published on the websites under their logo. The provisions of the Code that would apply to this content pertain to programme standards relating to the prohibition of the spread and incitement to hatred, violence or discrimination, the transmission of a clear and immediate risk of adverse consequences for human safety and health, the protection of minors, including the protection of their privacy, and the right of reply. The draft is in the public consultation phase, but the opinions of media representatives are divided, especially considering that these changes would not include other online media outlets that are not the websites of public radio and television broadcasters (Tomić 2021a).

Cases of sanctioning television and radio broadcasters for spreading hate speech in recent years are rare. In the last five years (2017-2021), the Agency received 25

complaints on hate speech, but most of them were dismissed.²⁴ In 2021, the Agency also received complaints on hate speech but concluded that in these cases there were no breaches of the code and that they were in line with freedom of expression (CRA 2021c). The Agency, however, does not conduct monitoring and mostly reacts based on citizens' complaints. In addition, there are concerns over its politicization, which would also impact the Agency's decision-making over breaches of its codes (Sokol and Jukić-Mujkić 2021, 8).

Case 1

In 2020, the Agency issued written warnings to TV Igman and TV Podrinje for violating Article 3, which prohibits content that includes any discrimination or prejudice based on gender, race, ethnicity, religion or beliefs, disability, special needs, age, or sexual orientation. The warnings were issued for the broadcasting of the programme *Defte Hefter sa Fatmirom Alispahićem*, due to the host's comments on the differences between Bosniaks and Serbs (in the case of TV Igman), as well as the host's comments on the threat to Bosniaks resulting from the arrival of migrants (in the case of TV Igman and TV Podrinje) (CRA 2021a, 7-8).

Case 2

In 2019, the Agency ruled a violation of Article 4 on hate speech by HTV Oscar C, which rebroadcast the programme *Bujica* of Z1 television, Zagreb, in which the host of the show made xenophobic and racist statements against migrants, and, in addition to the host, the guests also expressed extreme attitudes towards migrants. The Agency issued HTV Oscar C a fine of BAM 6,000 (CRA 2020a, 10).

6.1.4. International standards on hate speech and recommendations

In its analytical report following the EU Opinion, the European Commission recommended that the country's legislation on hate speech needs to be fully aligned with the European Court on Human Rights case-law and the EU acquis, as it currently only provides for a criminal offence of incitement to religious and national hatred and intolerance (European Commission 2019, 47).

The International Convention on the Elimination of All Forms of Racial Discrimination (ICERD), which BiH ratified, requires states to incriminate any spread of ideas based on racial superiority or hatred, any incitement to racial discrimination, any act of violence or incitement to such violence directed against any race or group of persons of another colour or other ethnic origin, or the provision of assistance to racist activities, including their financing, and to declare as illegal and to ban organizations as well as organized and all other propaganda activities that

²⁴ Azra Maslo, interview.

encourage racial discrimination as well as declare participation in those activities illegal (ICERD 1965; Filipović 2019, 3).

The European Commission against Racism and Intolerance (ECRI) recommended that BiH should, among other things, and according to international standards, add other protected categories to the enumerated prohibited grounds of inciting and inflaming; criminalize racist insults, defamation and threats, and the public expression, with a racist aim, of an ideology which claims the superiority of, or which depreciates or denigrates, a grouping of persons on the grounds of race, colour, language, religion, nationality, or national or ethnic origin; sanction the creation or leadership of a group which promotes racism, as well as the support for such a group and the participation in its activities; prohibit the public dissemination or public distribution, or the production or storage aimed at public dissemination or public distribution, with a racist aim, of written, pictorial or other material with racist content (ECRI 2017, 35).

In addition, Bosnia and Herzegovina ratified the Council of Europe Convention on Cybercrime and its additional protocol, but its recommendations still need to be implemented (European Commission 2021, 20; Filipović 2019, 4). The protocol requires states to adopt at national level legislative and other necessary measures to establish as criminal offences the following actions made through a computer system: distributing and making available, racist and xenophobic material, threats motivated by racism and xenophobia; insults motivated by racism and xenophobia (Filipovic 2019, 4).

Regarding the courts in BiH, the judiciary should follow the practice of the European Court of Human Rights, according to which only serious and extreme examples should be criminalized (Constitutional Court 2016, point 33). The Rabat Plan of Action suggests a high threshold for defining restrictions on freedom of expression, incitement to hatred, and for the application of Article 20 of the International Covenant on Civil and Political Rights. Interviewees suggest that the courts in BiH should take into consideration the Rabat six-part threshold test, taking into account: (1) the social and political context, (2) the status of the speaker, (3) the intent to incite the audience against a target group, (4) the content and form of the speech, (5) the extent of its dissemination and (6) the likelihood of harm, including imminence.

The prohibition of hate speech, the vagueness of its meaning and lack of precise definitions, can lead to infringements of lawful expressions. In this regard, the Election Law should be amended to include a more precise definition of hate speech according to the above-mentioned international standards in order to limit undue limitation of the freedom of expression and ensure consistent decisions. The ECRI recommended an extension of the Central Electoral Commission's mandate to monitor the use of hate speech during the entire duration of election campaigns, not just for thirty days before the date of the elections (ECRI 2018, 35).

Regarding the regulatory framework, the EU Directive 2018/1808 extends the scope of the regulation – in addition to television broadcasting and on-demand audiovisual media services – to video-sharing platforms. According to the Directive, the video-sharing platforms should be subject to appropriate and proportionate measures by the EU member states, preferably through co-regulation, in order to protect the general public from illegal content such as hate speech (Petković 2021, 14-15).

The EU Directive states that harmful content and hate speech provided on video-sharing platform services have increasingly given rise to concern, due to which it is necessary to set out proportionate rules on those matters. According to the representatives of the Communications Regulatory Agency, the platforms will be responsible for introducing certain measures for protecting the users, including different tools for checking the age of users and mechanisms for reporting content, and the role of the regulator will be to monitor whether the platforms are implementing these rules.⁴¹

The EU Directive 2019/1808 stipulates that video-sharing platform providers should be required to take appropriate measures to protect minors from content that may impair their physical, mental or moral development and to take appropriate measures to protect the general public from content that contains incitement to violence or hatred directed against a group or a member of a group on any of the grounds referred to in Article 21 of the Charter of Fundamental Rights of the European Union (Article 47), which include sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation.

6.2. PROHIBITION OF CONDONING, DENIAL AND JUSTIFICATION OF CRIMES OF GENOCIDE, CRIMES AGAINST HUMANITY AND WAR CRIMES

Condoning, denial and justification of crimes of genocide, crimes against humanity and war crimes are prohibited by the Criminal Code of Bosnia and Herzegovina and the Criminal Code of the Federation of Bosnia and Herzegovina. These provisions in the Criminal Code of Bosnia and Herzegovina were introduced in 2021 but provoked a political crisis since the Serb political representatives stated that they were anti-Serb. Many cases have been reported to the Prosecutor's Office of Bosnia and Herzegovina but there are still no court proceedings, and denials of war crimes and glorifications of war criminals are still present in the country.

6.2.1. Legislation

The High Representative in Bosnia and Herzegovina amended Article 145a of the Criminal Code of BiH with new provisions by the Decision Enacting the Law on Amendments to the Criminal Code of Bosnia and Herzegovina in July 2021.²⁵ These provisions prohibit public condoning, denial, gross diminution or attempts to justify the crimes of genocide, crimes against humanity and war crimes established by

²⁵ The Criminal Code of FBiH provides a sentence for inciting and inflaming national, racial and religious hatred, discord or intolerance by publicly denying or justifying genocide, crimes against humanity or war crimes established by a final decision of the International Court of Justice, International Criminal Tribunal for the former Yugoslavia or a domestic court (Articles 163, 5).

final judgments,²⁶ directed against a group of persons or a member of a group determined by race, colour, religion, background or national or ethnic origin, in a manner that could incite violence or hatred directed against such a group of persons or a member of such a group (Article 145a, 3). These amendments do not prohibit all condoning, denial, gross trivialization or attempts to justify crimes of genocide, crimes against humanity or war crimes, but only if it is directed against a group of persons or a member of a group defined by reference to race, colour, religion, descent or national or ethnic origin, when the conduct is carried out in a manner likely to incite to violence or hatred against such a group or a member of such a group (Gačanica 2021).

In addition, these amendments prohibit public incitement to violence or hatred directed against a group of persons or a member of a group determined by race, colour, religion, origin, or national or ethnic origin (Article 145a, 2), but not on other grounds such as sexual orientation, gender identity and disability. They also prohibit giving recognition, award, memorial, any kind of memento, or any privilege or similar to a person sentenced by a final judgement for genocide, crimes against humanity or a war crime, or naming a public object such as a street, square, park, bridge, institution, building, municipality or city or similar, or registering a brand, after or under a name of a person sentenced by a final judgement for genocide, crimes against humanity or a war crime (Article 145a, 6). This article does not have the element of conditionality, i.e. that this act incites violence and hatred, which means that this act can be prosecuted relatively quickly (Gačanica 2021).

However, when the amendments were introduced, in 2021, they provoked a political crisis.²⁷ As a response to the amendments, the RS National Assembly adopted the Law on Non-Application of the Decision of the High Representative Enacting the Law on Amendments to the Criminal Code of Bosnia and Herzegovina as well as Amendments to the Criminal Code of RS. The Law states that the Decision of the High Representative will not be applied on the territory of RS (Article 1) and that the competent authorities of RS will not cooperate with the competent authorities of BiH regarding the application of the Decision of the High Representative (Article 2.1), with the goal of protecting the citizens of RS (Article 2, 2). In addition, an amendment has been introduced to the Criminal Code of RS, stipulating prison sentences of up to three years for publicly exposing to ridicule, contempt or grossly disparaging Republika Srpska, its flag, coat of arms, emblem or anthem (Article 280a, 1).²⁸ Such provisions can be very dangerous for freedom of expression and could be used to

²⁶ By a final judgment in accordance with the Charter of the International Military Tribunal annexed to the London Agreement of 8 August 1945 or the International Criminal Tribunal for the Former Yugoslavia or the International Criminal Tribunal or the Court of Bosnia and Herzegovina (Article 145a, 3).

²⁷ Sunita Dautbegović Bošnjaković, forumZFD, interview.

²⁸ If this act is committed in a way that marks the RS as an aggressor or genocidal creation or its peoples as an aggressor or genocidal, the perpetrator shall be punished by imprisonment for a term between six months and five years (Article 280a, 2), and if the offence is committed by a responsible person or a person in government institutions or a body financed from the public budget they shall be punished by imprisonment for a term between two and ten years (Article 280a, 3). Three to fifteen years in prison are envisaged for the commission of these acts if their aim is to change the constitutional order of the RS, its territorial integrity or independence (Article 280a, 4).

sanction critical voices. The reputation of public authorities (except the courts) are not provided as possible restrictions in paragraph 2 of Article 10 of the European Convention for the Protection of Human Rights and do not constitute legitimate aims justifying restrictions on freedom of expression (Ivanović 2021, 45).

Despite the political crisis, many cases of denials have been already reported to the Prosecutor's Office. In addition, the monitoring by the Srebrenica Memorial Center has shown that, since the amendments entered into force, the denial of genocide has been carried out less explicitly and directly in BiH and that the number of such acts has decreased. However, it was also stated that politicians have continued to deny genocide, mostly in the media in Serbia, whose statements are then transmitted by the media in BiH (Srebrenica Memorial Center 2021). The database Mapping Hate of the Balkan Investigative Reporting Network BiH (BIRN BiH) shows that high-ranking politicians also continue to deny war crimes and their statements are published by the mainstream media without any critical stance.²⁹

6.2.2. International standards and recommendations

The amendments to the Criminal Code of BiH follow the standards of the Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law. The Framework prescribes that states should punish publicly condoning, denying or grossly trivializing crimes of genocide, crimes against humanity, directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin when the conduct is carried out in a manner likely to incite to violence or hatred against such a group or a member of such a group. These provisions can be used only if the crimes referred to in these paragraphs have been established by a final decision of a national court of the state and/or an international court (Article 1, 4). Such frameworks are present in most European countries (Spain, Portugal, Malta, Croatia, Montenegro, Serbia) and the ECRI also recommended that Bosnia and Herzegovina should criminalize the public denial, trivialization, justification or condoning, with a racist aim, of crimes of genocide, crimes against humanity or war crimes according to international standards (ECRI 2017, 35; Sali-Terzić 2019, 6).³⁰

However, it should also be noted that such provisions should be used only when the conduct is carried out in a manner likely to incite to violence or hatred against a group or a member of a group defined by reference to race, colour, religion, descent or national or ethnic origin. The UN Human Rights Committee noted that opinions that are "erroneous" and "an incorrect interpretation of past events" may not be subject to general prohibition, and any restrictions on the expression of such opinion "should not go beyond what is permitted" under Article 19 (3) or "required under article 20" of the International Covenant on Civil and Political Rights. The UN Special Rapporteur on the promotion and protection of the right to freedom of

²⁹ See the map: <https://mapiranjemrznje.detektor.ba/>

³⁰ Sunita Dautbegović Bošnjaković, forumZFD, interview.

opinion and expression stated that denial of the historical accuracy of atrocities should not be subject to criminal penalty or other restrictions without further evaluation under the definitions and context noted above and that the application of any such restriction under international human rights law should involve the evaluation of the six factors noted in the Rabat Plan of Action (A/74/486, 10).

6.3. REGULATION OF DISINFORMATION

There are no specific legal frameworks to regulate disinformation online and the emphasis is placed on self-regulating actors such as the Press Council of Bosnia and Herzegovina, the fact-checking platform and social networks. Dissemination of false content that can harm the reputation of a person is punishable by defamation laws but these have often been used to intimidate journalists (Radević 2021, 11-12). Laws on public order prescribe misdemeanor sanctions for spreading “fake news” or allegations that can cause harassment of citizens and endanger public order and peace, but they lack precise definitions of “fake news”, and they could therefore be misused for curbing free speech.

6.3.1. The regulatory framework

The Code of Audio-visual Media and Radio Media Services of the Communications Regulatory Agency of Bosnia and Herzegovina does not contain terms such as “disinformation” or “fake news”, but through a number of articles of the Code it is possible to sanction certain forms of disinformation on television and radio stations. Article 5 (Fairness and impartiality) states that radio and television programmes shall ensure the accuracy of the information presented, and that all observed errors shall be corrected in a timely manner. Article 7 (Fake or misleading programmes) states that audio-visual and radio programmes shall not offer any content that is known or can be found to be fake or misleading on the basis of common sense or routine verification, or for which there is a reasonable presumption that it is false or deceptive.⁴⁵ This article also states that if a programme subsequently turns out to be fraudulent or misleading, a correction will be published as soon as possible.

These provisions are rarely used to sanction disinformation in the content of radio and television stations. In 2020 and 2019, the regulator imposed several fines and warnings for violating Article 7 (False or misleading programmes, Article 7) because information on charges to citizens’ calls to participate in certain shows was misrepresented in certain television broadcasts, i.e. calls were charged by duration and not per call as is prescribed by the Agency (CRA 2021, 11).

As has already been pointed out, the rules apply to radio and television programmers and not online media even though the Communications Regulatory Agency of Bosnia and Herzegovina has prepared draft amendments to the rules, which would extend the responsibility of licensees for television and radio broadcasting to content published on the websites under their logo. The draft amendments, however, are limited to only some harmful content, such as hatred, violence or discrimination,

the transmission of a clear and immediate risk of adverse consequences for human safety and health, and not disinformation. If adopted, disinformation published on online media under the logo of television and radio broadcasters that has a clear and immediate risk of adverse consequences for human safety and health could be punishable.

Case

The Communications Regulatory Agency of BiH received complaints in 2020 about the FACE TV programme *Centralni dnevnik* in which the guest Semir Osmanagić presented conspiracy theories about the coronavirus Covid-19 and advised the use of alternative medicine. The CRA found that there was no violation of the provisions of the Code, but pointed to the need for a more responsible and professional approach to topics related to the coronavirus pandemic, in order to prevent the spread of disinformation that could adversely affect human behavior and safety (CRA 2020b, 15). In the decision of the CRA it was stated that the entire programme of FACE TV had relevant interlocutors in the field of economy, health and safety, thus providing timely information to viewers (RAK 2020b, 15).

6.3.2. Defamation laws

Laws on protection against defamation of the entities and the Brčko District regulate civil liability for harm caused to the reputation of a natural or legal person by making or disseminating false content identifying that legal or natural person to a third person.³¹ Their goals, as stipulated in the laws, are to attain a) the right to freedom of expression, which constitutes one of the essential foundations of democratic societies; b) the right to freedom of expression that is not only applicable to expressions that are received as favorable or inoffensive but also to those that might offend, shock or disturb; and c) the essential role of media in the democratic process as public watchdogs and transmitters of information to the public (Article 2).

Decriminalization of defamation was an important step in promoting freedom of expression, but inconsistent application of the laws has led to legal uncertainty and instrumentalized lawsuits against the media and journalists. Laws on protection against defamation have been used to intimidate journalists and politicians continue to use civil suits to intimidate journalists (Radević 2021, 7-11). Numerous defamation lawsuits against journalists have been filed by politicians and other public officials, such as judges and prosecutors (Sokol and Jukić-Mujkić 2021, 6).

³¹ All three laws defined the essential elements of defamation in the same way, only with differences in wording (Halilović and Džihana 2012, 132).

Case

In 2021, the Principal Court in Sarajevo passed a first instance verdict in a defamation lawsuit sentencing the online magazine Žurnal and its editors and journalists to pay a total amount of more than 170,000 BAM, without interest, which could jeopardize the work of the media outlet as the fine is considered very high.³²

6.3.3. Laws on public order

A number of laws on public order prescribe sanctions for spreading “fake news”. The Law on Public Order of the Brčko District and six out of ten cantonal laws on the protection of public order and peace in the Federation of Bosnia and Herzegovina prescribe misdemeanor sanctions for spreading false news or allegations that can cause harassment of citizens and endanger public order and peace. For example, Article 10 of the Law on Public Order and Peace of the Brčko District states that “whoever presents or transmits false news, rumours or allegations which may cause disturbance of citizens or endanger public order and peace, shall be fined in the amount of BAM 100 up to 500”.

Most of these laws do not provide precise definitions of “fake news” or distinctions between disinformation and misinformation – the intention or lack of intention to harm – which is problematic for the implementation of the law and could curb freedom of expression. The Law on Public Order and Peace of the Herzegovina-Neretva Canton defines “dissemination of fake news” as “intentionally presenting or transmitting fake news or allegations, thereby disturbing public order and peace and tranquility of citizens”, but does not specify sanctions (Article 3, e). The Law on Public Order and Peace of the Una-Sana Canton and the Law on Public Order and Peace of the Tuzla Canton, in addition to fines, prescribe the possibility of imposing a prison sentence of up to 30 days and up to 60 days respectively for making or spreading fake news or claims which can cause citizen unrest or endanger public order and peace. Prescribing prison sentences for spreading fake news is not in line with international standards and as such can impede freedom of expression (Ivanović 2021, 45). These laws define public space as streets, squares, public roads, parks, sports stadiums and playgrounds, and cinemas, but not the internet. There were initiatives in the Parliament of the Federation of Bosnia and Herzegovina in 2021 to amend the laws and include the internet as public space (Ašerić 2021). Vague laws that confer excessive discretion can lead to arbitrary decision-making, and such amendments could be problematic as provisions of the laws could be used for sanctioning individuals on the internet on arbitrary grounds.

³² See the case: <https://bhnovinari.ba/en/2021/07/13/bh-journalists-the-verdict-against-the-zurnal-magazine-directly-endangers-the-freedom-and-work-of-the-investigative-media/>

Case

During the pandemic of Covid 19 in 2020 the Government of Republika Srpska adopted an order banning the presentation or transmission of fake news or allegations that cause panic or seriously disturb public order or peace or disable or significantly interfere with the implementation of decisions and measures of state bodies and organizations exercising public authority. In a similar manner, the Crisis Committee of the Brčko District adopted an order for the prevention of publicly presenting and transmitting fake news and information related to the coronavirus epidemic and which may cause fear and panic or otherwise negatively affect the implementation of measures to protect the population from the epidemic. The order specified the means for the release of such information: media, social networks, as well as any other means of disseminating such information. In the period from March to April 2020, several orders and fines were issued for spreading disinformation and panic under this regulation in the RS, including social networks, such as Facebook and Instagram. One person was fined for claiming on Instagram that the government was hiding the number of infected people³³ and another person for insulting in a harsh manner the representatives of state organs in RS on Facebook.³⁴ After criticism from international and local organizations, the RS Government repealed this regulation (Government of RS 2020b).

6.3.4 International standards and recommendations

The prohibition of inaccurate information is not in itself a legitimate aim under international human rights law. Freedom of information may be restricted only if it is necessary to respect the rights and reputation of others and to protect national security, public order and public health and morals as stipulated by Article 19 of the International Covenant on Civil and Political Rights. According to these provisions, only disinformation that endangers the rights and reputation of others, national security, public order and public health and morals may be restricted (A/HRC/47/25). It should be noted that the right to freedom of expression includes different types of information, “including those that may shock, offend or disturb, and irrespective of the truth or falsehood of the content” (A/HRC/47/25, 8).

Regulation of disinformation should be taken with caution and involve different actors. Transparent government and independent journalism are, for example, strong antidotes to disinformation (A/HRC/47/25, 5). Vague laws that give excessive discretion can lead to arbitrary decision-making and such provisions should be repealed.

The representatives of the Communications Regulatory Agency state that in line with the practices in the EU, the fight against disinformation should be placed on self-regulation and co-regulation by different actors, including the media, governments, ICT companies, advertisers, individual producers of content, non-governmental

³³ See: <https://mup.vladars.net/lat/index.php?vijest=23056&vrsta=novosti>

³⁴ See: <https://mup.vladars.net/lat/index.php?vijest=23082&vrsta=novosti>

organizations, etc.³⁵ The European Commission has a number of initiatives to tackle disinformation such as: the Code of Practice, which sets out principles and commitments for online platforms, leading social networks, advertisers and the advertising industry to counter the spread of disinformation online in the European Union; the European Digital Media Observatory, a European hub for fact-checkers, academics and other relevant stakeholders to support policy-makers; the action plan on disinformation that aims to strengthen EU capability and cooperation in the fight against disinformation; and the COVID-19 monitoring and reporting programme.³⁶ The key instrument of the EU approach to tackling disinformation is the Code of Practice on online disinformation – self-regulation by the leading online platforms, advertisers and the advertising industry. The Commission presented a Guidance to strengthen the Code of Practice on disinformation in 2021, which aims at evolving the existing Code of Practice towards a co-regulatory instrument, foreseen under the Digital Services Act.³⁷

Regarding defamation, in its progress report 2021, the European Commission recommended that in Bosnia and Herzegovina efforts should be made to ensure an expedient processing of defamation cases and consistency of case law on damage awards, to prevent any chilling effect that would force journalists into self-censorship (European Commission 2021, 28).

In accordance with the case law of the European Court of Human Rights the defamation laws in Bosnia and Herzegovina should be aligned with international standards to incorporate different standards of acceptability when it comes to alleged defamation against public figures in relation to private persons, as public figures are required to tolerate a greater degree of criticism than private citizens. In a General Comment, no. 34, The Human Rights Committee stated that the mere fact that forms of expression are considered to be insulting to a public figure is not sufficient to justify the imposition of penalties and all public figures are legitimately subject to criticism and political opposition (CCPR/C/GC/34, 9). Consideration should be given to avoiding penalizing or otherwise rendering unlawful untrue statements that have been published in error but without malice, at least with regard to comments about public figures (CCPR/C/GC/34, 12).

6.4. REGULATION OF ETHNO-NATIONALLY AND/OR POLITICALLY BIASED REPORTING

Ethno-nationally and/or politically biased media reporting is not in accordance with journalistic professional standards and the Communications Regulatory Agency of Bosnia and Herzegovina can impose sanctions on television and radio stations

³⁵ Its Code also contains special articles for content dealing with paranormal and parapsychological phenomena (Article 12), alternative medicine (Article 13), and quackery (Article 14). However, insight into the measures imposed by the Agency shows that these provisions are rarely used, much less to sanction misinformation.⁴⁶ Azra Maslo, CRA, interview.

³⁶ See: <https://digital-strategy.ec.europa.eu/en/policies/online-disinformation>

³⁷ See: <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>

for not providing fair and impartial information. These, however, cannot be used for online and print media, and the politicization of the Regulatory Agency for Communication has been raising concerns over its decision-making. In addition, Article 16 of the Election Law stipulates that the media should cover election activities in a just, professional and competent manner, but many of its provisions are vague and the law does not have specific provisions for online media.

6.4.1. The regulatory framework

The Code of Audio-visual Media and Radio Media Services of the Communications Regulatory Agency has provisions about the obligation of radio and television broadcasters to produce content that is fair and impartial (Article 5). The code states that audiovisual and radio programmes will ensure that all programme content, especially information-political and current affairs programmes, will be fair, impartial and balanced. It states that fairness, impartiality and balance are achieved by publishing all opposing views, either in the same programme or in a series of other programmes that represent a whole in the treatment of a particular topic. It also stipulates that audiovisual and radio programmes shall not tendentiously promote the interests of one political entity or any group or individual, nor will it be allowed that one opinion or position prevails when it comes to controversial political, economic and other topics of public interest. However, these rules only apply to television and radio stations and the Agency does not have jurisdiction over online media. The highest number of both the complaints received and the sanctions issued by the CRA on a yearly basis have pertained to the breaches of these standards, but this has changed in recent years. Experts state that the lack of fines for television and radio station breaches of the norms on fairness and impartiality is connected to increased political influence on the CRA (Bubonjić 2022a).

Case

In 2018, the CRA conducted a one-month monitoring of the three public service broadcasters' (BHRT, FTV and RTRS) news content and concluded that RTRS continuously and tendentiously promoted the interests of one party, constantly presenting the party's members in a positive and affirmative way and mostly reporting critically on the activities of the opposition (CRA 2019, 9).

6.4.2. The Election Law

The Election Law regulates the election process, including the activities of the media in the election campaign. Chapter 16, entitled Media in the Election Campaign, lays down provisions relevant to the work of the media during the election campaign. According to these provisions, the electronic media will cover election activities in a balanced, fair and impartial manner (16.2). It is forbidden to favor political entities, including officials at all levels of government who are also candidates in elections.

The law also contains provisions emphasizing that the principles of balance, fairness and impartiality should be respected in news programmes, interviews and political debates that could affect the mood of voters (Article 16.4). Journalists and presenters are prohibited from disclosing their party affiliation or preference (Article 16.6). The law regulates paid political advertisements and states that electronic media shall provide equal conditions for paid political advertisements of political entities (commercials, public calls, jingles, video-clips and any other type of promotion of a political entity) in the period of 30 days prior to the Election Day (16.12).

The Election Law addresses general issues of fair and impartial media reporting during elections, but its rules are not precise and the law does not contain specific provisions for online media (Džihana 2018). The Law states that electronic public media need to allow political subjects to address the audience for free (16.14) and that private electronic media also may do so but under equal conditions for all (16.5). The Rulebook on media representation of political entities further specifies that every political subject should be given at least three minutes to address the audience for free, that public electronic media should enable all political subjects in a constituency that is covered by their signal to address the audience and that the order of their representation should be decided by lot (Article 7). In 2020 during the Local Elections, the Communications Regulatory Agency fined four television stations and issued three warnings for breaching the Election Law and the Rulebook on Media Representation of Political Entities (CRA 2021a, 26-27).

Case

In the programming of Radio Rama in 2020, the mayor of the Municipality Prozor-Rama was a guest on one of its shows and spoke about the budget of the municipality, its projects and his achievements. Later in the show he mentioned his candidacy and his plans for the next period and sent messages to other candidates. The Communications Regulatory Agency ruled on breaches of principles of balance, fairness and impartiality in an information programme, according to the Election Law. The CRA concluded that the show in question belonged to a programme that could impact on voter sentiment, in particular bearing in mind that the show was used for the purpose of the political promotion of the current mayor, thus placing the mayor in a privileged position in relation to his opponents (CRA 2021a, 28).

6.4.3. International standards

The provisions of the Election Law follow the standards set out in the Council of Europe Recommendation No. R (99) 15 on measures concerning media coverage of election campaigns. For example, in the recommendation it is stated that with due respect for the editorial independence of broadcasters, regulatory frameworks should also provide for the obligation to cover electoral campaigns in a fair, balanced and impartial manner in the overall programme services of broadcasters (II, 1). The Council of Europe Recommendation CM/REC (2007) states that member states should apply the principles concerning the broadcast media and rules on “fairness,

balance and impartiality” to “non-linear audiovisual media services of public service media”. The recommendation also says that states should ensure that “there is an effective and manifest separation between the exercise of control of media and decision making as regards media content and the exercise of political authority or influence” (I, 2). In addition, “all media are encouraged to develop self-regulatory frameworks and incorporate self-regulatory professional and ethical standards regarding their coverage of election campaigns” (I, 5).

Broadcast media have historically been subject to the most regulation during elections, but the shift to social media in election campaigning has been posing challenges, as the rules need to be updated to reflect the importance of online campaigning and online media. There is a considerable amount of regulation in different European countries applicable to online media and many rules applicable to online media are contained in election and referendum legislation, but specific attention should be placed on online political campaigning and political advertisements (Council of Europe study, DGI 2017, 11). In addition to this, it is important to encourage media pluralism and media independence to counter biased media reporting. In its progress reports, the EU Commission recommended that BiH should adopt legislation on media ownership transparency and criteria on public advertising (EU Commission 2021, 27).

The European Commission has set up a number of initiatives to ensure media freedom and pluralism. It also adopted the European Democracy Action Plan and will put forward further measures to support media pluralism and to strengthen transparency of media ownership and state advertising, among other issues, through the new Media Ownership Monitor.³⁸ The Commission propose³⁹ legal action on political advertising that will address the sponsors of paid content and of production and distribution channels, including online platforms, advertisers and political consultancies, clarifying their respective responsibilities.

6.5. REGULATION OF THREATS, ATTACKS AND SMEAR CAMPAIGNS

Parts of smear campaigns such as threats and sexual harassment are prohibited by different laws. For example, online threats can be processed under criminal codes. Article 183 of the FBiH Criminal Code prohibits endangering security by threatening to attack the life or body of a particular person (1). Article 150 of the RS Criminal Code prescribes punishment for endangering the safety of a person by seriously threatening to deprive him or her or a person close to them of life, cause bodily harm, or cause harm by arson, explosion or any other dangerous act or means. Laws on public order and peace also prohibit verbal and physical attacks and threats against persons, but the existence of prohibitions in criminal codes and laws on public order enable different practices and interpretation of laws. When journalists report cases of attacks and threats, these (especially those online) are rarely processed by the

³⁸ See: https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2250

³⁹ See the proposal: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0731>

prosecutor's offices, the court proceedings are lengthy, and the charges against perpetrators are minimal (Radević 2021, 9). Since 2019, a number of cases filed by journalists have either been dismissed or were still ongoing at the time of finalizing this review, while the outcome of some cases was limited to disciplinary procedures against public officials.

In addition, sexual harassment is prohibited by the Law on Prohibition of Discrimination. It defines sexual harassment as any form of unwanted verbal, non-verbal or physical behavior of a sexual nature whose purpose or effect is to violate the dignity of the person, especially when it creates a frightening, hostile, degrading, or offensive environment (2).

The Law on the Protection of Personal Data prohibits the publication of personal data, including in the online sphere, but the jurisdiction of the Agency for the Protection of Personal Data of Bosnia and Herzegovina is limited to companies registered in Bosnia and Herzegovina.⁵¹ Thus, according to the representative of the Agency, the Agency cannot sanction unlawful publication of personal data on social networks, as they are not registered in BiH, nor do their provisions apply to the media that are regulated by the codes of the Regulatory Agency of Communication. Websites and webpages, and social networks that are registered in other countries remain out of the reach of such legislation.⁴⁰

In 2006, BiH ratified the 2001 Council of Europe Convention on Cybercrime, dealing particularly with infringements of copyright, computer-related fraud, child pornography and violations of network security. The provisions of the Convention are included in entity criminal laws (such as the provisions on damage to computer data and programs; computer sabotage). The provisions of the State Criminal Procedure Code and the entity criminal procedure laws do not include the procedural solutions contained in the Convention, which results in numerous dilemmas in the implementation of the said laws (Gačanica and Arnautović 2018, 9). In the Criminal Code of RS cybercrime includes crimes such as computer sabotage (Article 408); development and introduction of computer viruses (Article 409); computer fraud (Article 410); and exploitation of a computer network or communication by other technical means for committing the criminal offenses of sexual abuse or exploitation of a child (Article 178).

Case

In 2020 the Prosecutor's Office in Banja Luka suspended further investigations in the case of a female journalist who had received a death threat on Facebook under her photograph with migrants and refugees. In the decision of the Prosecutor's Office it is stated that the statements did not contain the elements of a criminal act of endangering security but were expressions of personal attitude and dissatisfaction, and there were no elements of spreading hatred.

⁴⁰ Silvije Fućec, Agency for the Protection of Private Data, interview.

International standards and recommendations

In its yearly progress reports the European Commission recommends that the country should ensure the protection of journalists and a systematic institutional follow-up on threats and violence against them (EU Commission 2021, 29). In its recommendations on ensuring the protection, safety and empowerment of journalists and other media professionals in the European Union, the EU Commission stated that public authorities have a duty to protect freedom of expression and journalists' safety by providing an appropriate legal environment, taking criminal threats against journalists seriously, vigorously prosecuting the perpetrators of any attack and ensuring appropriate investigation (4-5). It also recommended that the authorities consider setting up special units within police forces to investigate crimes against journalists, and the appointment and training of coordinators within prosecution offices (5).

6.6. OTHER MECHANISMS - THE INSTITUTION OF THE HUMAN RIGHTS OMBUDSMAN

Other mechanisms for countering and protecting against harmful content are the Institution of the Human Rights Ombudsman of BiH, the Ministry of Human Rights and Refugees of BiH and the Agency for Gender Equality of Bosnia and Herzegovina (Gender Centre of Republika Srpska and Gender Centre of the Federation of Bosnia and Herzegovina) (Ivandić-Ninković 2021, 33).

According to the Law on Prohibition of Discrimination (Article 7), the central institution competent for protection from discrimination is the Institution of the Human Rights Ombudsman of Bosnia and Herzegovina, which, among other things, receives individual and group complaints on discrimination, provides persons who have filed a complaint against discrimination with the necessary information on their rights, obligations and possibilities in terms of judicial and other types of protection, collects and analyses statistics on cases of discrimination, and issues recommendations to competent institutions for further action in cases of discrimination. Competent institutions in BiH are required to cooperate with the BiH Ombudsman Institution and present their responses and notifications in writing (Article 7). In 2019, for example, the Institution of the Human Rights Ombudsman received three complaints, and in 2020 four complaints on hate speech (Institution of the Human Rights Ombudsman 2021b, 36). What is problematic, however, as stated in the EU Commission's report, is that these recommendations remain rarely implemented (EU Commission 2021, 14). The Institution of the Human Rights Ombudsman in BiH also concluded that such a small number of reports of hate speech cases is due to citizens' lack of trust in institutions and that institutions for the protection of human rights should have adequate materials and capacities not only to handle complaints but also to conduct activities for awareness raising, education and prevention of hate speech (Institution of Human Rights Ombudsman 2021b, 44).

Case

In 2019, for example, the Ombudsman Institution issued a recommendation to the Prosecutor's Office of BiH, the Communications Regulatory Agency and the Federal Police Administration to act in accordance with their powers regarding the operation of the Antimigrant portal. This was done after the Institution determined that the articles on the Antimigrant portal were full of statements that spread and incite hatred both towards migrants and among the constituent peoples of BiH (Institution of Human Rights Ombudsman 2020). In addition to that, in 2018, the Ombudsman Institution sent a recommendation to a private optician from Banja Luka to immediately remove the Instagram post in which they announced that their stores would show a large inscription "stop migrants", and that in the future when posting on their Instagram page they must fully respect the Law on Prohibition of Discrimination (Institution of Human Rights Ombudsman 2020, 111).

Recommendations

In its progress reports the EU Commission stated that the Ombudsman has not made use of its prerogative to open minor offence proceedings for non-compliance with its recommendations in the past 5 years, nor has it initiated or intervened in administrative or judicial proceedings. The EU Commission recommended that the independence and effectiveness of this body should be improved (EU Commission 2021, 26).

7. SELF-REGULATION AND MONITORING OF HARMFUL CONTENT

Bosnia and Herzegovina has a self-regulatory body for the print and online media and a number of initiatives of mostly non-governmental organizations that periodically monitor some types of harmful content. The OSCE Mission to Bosnia and Herzegovina publishes a monthly visualization of hate crimes monitoring data, which include instances of hate speech as well. The work of the self-regulatory body is restricted to mediation between online and print media and their users, and the country lacks a comprehensive monitoring system of harmful content online.

7.1. THE PRESS COUNCIL OF BOSNIA AND HERZEGOVINA

The Press Council of Bosnia and Herzegovina self-regulates the work of online and print media but its activities are restricted to mediation and non-binding decisions about media content that violates the standards of the Press Code of Bosnia and Herzegovina, including, for example, hate speech. The Press Code of Bosnia and Herzegovina was amended at the end of 2021, and included a number of new articles, including the use of information technologies and disinformation. Article 3 states that publishers, editors and journalists of print and online media are obliged to use information technologies, platforms, social networks and various tools in good faith, for true, accurate, objective, verifiable and timely information to the public, for the common good, civilization values, humanism and ethical norms of this code. It also states that it is not acceptable to use information technologies, platforms, social networks and tools to spread hate speech, untruth and misinformation, stigmatization, or discredit others based on race, gender, age, language, nationality, ethnic and religious affiliation, opinions and political beliefs, social background, social status, sexual orientation or in any other way. The new provisions also expanded editorial responsibility to the overall content of the print and online media, including user generated comments (Article 2). According to the new provisions, the editor is obligated to erase user generated comments that contain hate speech, incitement to violence, harassment, threats and all other forms of inappropriate and socially unacceptable communication. The new Press Code defines disinformation as a gross violation of the basic rules of the journalistic profession and states that the editor will not be absolved from responsibility for publishing misinformation taken from another source.

Article 3 of the Code (Incitement) states that journalists will at all times be aware of the dangers that arise when the media incites discrimination and intolerance through hate speech, and do their best not to incite hatred and/or inequality based on ethnicity, nationality, race, religion, gender, sexual orientation, physical disability or mental condition, and will not in any circumstances incite crime or violence. The Code also requires accuracy and fair reporting (Article 5) and that journalists do not publish inaccurate or misleading materials in the form of photographs, texts or other materials, that photographs and documents must not be falsified and/or used in a misleading manner, and that journalists have a professional obligation to correct in a timely manner any published information that is found to be inaccurate.

The work of the Press Council is limited to mediation between the users and the media outlets and decisions that certain codes have been breached and that a media outlet should retract it, or publish a correction or a rebuttal, without any

power to sanction those who refuse to do so. Even if this mechanism is important for raising professional norms among professional content producers without imposing sanctions, its reach is still limited, particularly regarding non-professional content producers and anonymous portals. In 2021, the Press Council of BiH, the self-regulatory body for the print and online media, received 1,073 complaints, out of which 559 referred to texts published in print and online media, while 505 complaints referred to user generated comments, mostly hate speech, on online media (Išerić 2022). A large number of complaints – 766 of them – were resolved by self-regulation and most cases of hate speech were removed by the media. Most of the complaints pertain to hate speech in user generated content and to defamation, for which complaints mostly come from persons who are directly targeted, while complaints about biased media reporting or disinformation are rare. In addition, despite the fact that some media outlets have breached the Code many times, their manner of reporting has still not improved.

Case

The Complaints Commission of the Press Council did not find a violation of the Code when an online media outlet published a short video showing the dance of a member of the City Council of the City of Zenica, with the title: “She showed new talent: watch the ‘dirty dance’ of council member Tufekčić (Video)”. The Appeals Commission emphasized that membership in the council makes the person a public figure, and therefore that person could suffer greater criticism in the public. This decision, however, was criticized by non-governmental organizations.

7.2. CIVIL SOCIETY AND INTERNATIONAL ORGANIZATIONS

7.2.1. Fact-checking and media watchdogs

There are a number of organizations that monitor the content on the internet, social media, and online media, acting as fact-checking organizations and media watchdog organizations. Media watchdog organizations in Bosnia and Herzegovina include Mediacentar, which publishes thematic articles on issues related to the media and social networks, including disinformation and hate speech, and Analiziraj.ba which monitors the content of television broadcasters and also publishes articles on various topics related to the media sector.

Istinomjer, of the civil society organization Why Not, deals with the verification of the credibility of politicians’ statements and the fulfilment of election promises, and *Raskrinkavanje*, of the same organization, verifies the accuracy of media content. *Raskrinkavanje* checks the content of online media and social networks, and operates according to a methodology that recognizes 15 forms of problematic

media content, including disinformation and fake news, but not hate speech.⁴¹ Since its establishment in 2017 until today, the fact-checking platform has debunked thousands of examples of problematic media content on portals and social networks, especially during the Covid-19 pandemic. In 2020 the platform started working in partnership with Facebook, wherein, after fact-checkers mark content as fake, Facebook puts a mark on the disputed content and reduces the reach of these posts, also informing those who spread the questionable content.⁴² In this way, the public is warned about problematic content, and the reach of those who spread disinformation, and thus their financial resources, is reduced. Media representatives have complained that labelling content as fake is done without prior warning and that such measures drastically reduce their reach (Stokić 2021a).

7.2.2. Monitoring and reporting mechanisms

Other international and non-governmental organizations that deal with human rights and have monitoring and reporting mechanisms for harmful content include:

- Sarajevo Open Center, which monitors hate speech against the LGBTIQ community;
- Association for Democratic Initiatives,⁴³ that has an online form for reporting hate crimes and hate speech and gives free legal advice;
- Network for Building Peace, a coalition of more than 220 associations from BiH, which on certain occasions reacts to hate speech and reports it to the Institution of the Ombudsman and the Press Council;
- Association of BH Journalists, which receives complaints from journalists on attacks, threats and other violations of their rights and gives free legal advice including about online harassment. It also publishes reactions against attacks and harassment against journalists;
- Safejournalists.net, which publishes cases of threats and attacks against journalists in the Western Balkans, including Bosnia and Herzegovina, and is a network made of journalists’ associations across the Western Balkans.

⁴¹ Satire, clickbait, product placement, error, biased reporting, conspiracy theories, pseudo-science, manipulation of facts, disinformation, spin, fake news, transmitting fake news, corrected, censorship, unverified. In the methodology, disinformation is described as a media report that contains a “mix” of facts and inaccurate or semi-true content, and the media are not necessarily aware of the inaccuracy of the information; and fake news is described as an original media report that contains factually false claims or information created with the intention of misinforming the public. About the methodology see: <https://raskrinkavanje.ba/metodologija>

⁴² For details on the programme, see: <https://www.facebook.com/journalismproject/programs/third-party-fact-checking/how-it-works>

⁴³ The Association does not receive many reports through their online form, but more by telephone and e-mail. After receiving the report, they contact a lawyer who gives advice. Previously they reported cases directly to the police but say that it took a long time for them to respond (Dajana Čelebić, Association for Democratic Initiatives, interview).

- The Interreligious Council of BiH, a non-governmental organization that connects four religious communities in BiH (Muslim, Orthodox, Catholic and Jewish), has an online form for reporting attacks on religious objects, including hate speech and hate crimes.⁴⁴
- The OSCE mission to Bosnia and Herzegovina⁴⁵ has a mechanism for monitoring hate crimes (which can include also hate speech as an indicator) based on police records and court proceedings.⁵⁶ The Hate Monitor is a monthly visualization of the hate crimes monitoring data of the OSCE mission to BiH, which presents the latest data on all known bias-motivated incidents and responses to these incidents by the justice sector, local authorities, and civil society. The statistics include victim characteristics, the type of incident and the corresponding community response as well as data on ongoing and completed hate crimes cases. It also includes hate incidents that are committed with a bias motivation, but may not reach the threshold of a hate crime.
- The Balkan Investigative Reporting Network initiated a Mapping of Hate at the end of 2021, with data on hate incidents, hate speech, discriminatory speech and denials of crimes collected through media reporting, and mostly pertaining to political actors.
- The Share Foundation from Serbia monitors violations of digital rights in Croatia, Hungary, North Macedonia, Romania and Serbia, as well as in Bosnia and Herzegovina. The database includes categories such as destruction and theft of data and programs, computer fraud, publishing information about private life, publishing falsehoods and unverified information with the intention to damage reputation, insults and unfounded accusations etc.⁴⁶
- The Coalition Pod lupom, which monitors elections for the prevention of election fraud and election irregularities.

Even though a number of actors monitor harmful content, their work is fragmented and focused on certain types of content. BiH lacks a mechanism for monitoring that would monitor more types of harmful content online, particularly content that is specific to the BiH context, such as hate speech and hate narratives against ethnic groups and ethno-biased media reporting, and the country also lacks an intersectoral coalition that would include more actors and initiatives.

Case

As of March 2022, the Mapping Hate database of BIRN BiH contained 53 cases of harmful content including: 33 incidents, 7 cases of discriminatory speech and 13

⁴⁴ Stefan Terzić, Interreligious Council, interview.

⁴⁵ According to the monitoring mechanism, every year there are 130 cases of incident reports to the police, out of which some are processed under the clause of “inciting hatred”. Halisa Skopljak, OSCE mission to BiH, interview.

⁴⁶ See the data base for BiH: <https://bih.bird.tools/>

cases of denial of war crimes. The map shows that despite the amendments to the Criminal Law of BiH that prohibit public condoning, diminishing and justification of war crimes and glorification of war criminals, such incidents still occur, and come even from high-ranking politicians in BiH.

7.3. ONLINE MEDIA

Online media have different practices in moderating their comments sections, including media that do not have comment sections, that do not allow certain words in comments sections, that prior to publishing review all comments, or that delete comments after users report them. However, most online media lack capacities to fact-check their online content and to monitor and moderate their comments sections online, including on their accounts on social networks. While some online media are dedicated to professional standards, others are governed by political agendas or financial motivation that is conducive of harmful content.

Case

By the end of 2021, the most popular online media in BiH, klix.ba, had set up an internal monitoring system of hate speech in user generated content according to which it publishes new data every hour. By March 2022, the website had more than 1.3 million reported comments, more than 2.5 million had been removed and 30 thousand users banned. The disclaimer above their comments section states that the comments reflect the views of their authors alone, who may be prosecuted for hate speech.

7.4. RECOMMENDATIONS

Self-regulatory mechanisms are important tools for countering harmful content online, especially content that is not considered illegal but that can harm individuals, groups and democracies. Bosnia and Herzegovina has a functioning self-regulatory system and a number of media watchdog and other organizations that collect data on different types of harmful content. However, their work is fragmented for the most part, and the effect of their work on raising professional standards in the online media or reducing harmful content online has not been researched.

A coalition to combat harmful content online comprised of civil society representatives, media representatives, representatives of the Communications Regulatory Agency of BiH, the Press Council of Bosnia and Herzegovina, the Institution of the Human Rights Ombudsman of BiH and academics could be established and a monitoring mechanism set up that would include the monitoring of different types of harmful content. Civil society organizations, together with government institutions, should work together to counter harmful content online and protect freedom of expression. Civil society organizations should be

additionally capacitated in developing expertise in digital rights and be supported by international organizations, foreign donors and authorities in BiH.

The Press Council of BiH should be capacitated to monitor online media occasionally, and the members of the Council Commission should be trained in the new provisions of the Press Code and capacitated to deal with different cases. The Press Council of BiH should promote the use of the new provisions of the Press Code and promote its complaints mechanism among the general public.

Additional efforts should be placed at raising media and information literacy skills of citizens, and providing tools and knowledge to journalists regarding how to fact check and moderate user generated content. In addition, specific guidelines regarding harmful content should be developed for the online media, giving instructions to the media as to how to deal with statements that contain hate speech or denials of war crimes, how to avoid biased media reporting and provide constructive content, and how to counter disinformation in their content.

8. COMMUNITY GUIDELINES OF SOCIAL NETWORKS AGAINST HARMFUL CONTENT

Social networks, particularly Facebook, are widely used in Bosnia and Herzegovina. A recent research study showed that 52% of respondents access information through social networks and that 73% of respondents use at least one social network. The most popular social network is Facebook, on which almost three quarters of adults have an account (73%), followed by Instagram (39%) and YouTube (38%). All other social networks are used by fewer than 10% of adults – TikTok (8%) and Snapchat (8%), Twitter (4%), Pinterest (2%) and LinkedIn (2%) (Hasanagić et al 2021, 35).

According to the study, BiH adults (more than a quarter of them) “only” read content on social networks for the most part, and rarely perform any activity on them (such as “liking” or posting content). Less than a fifth of the respondents in the study (19%) said that they sometimes create/post content, while slightly fewer (16%) sometimes share/comment on content, and 16% said that they sometimes “like” content (Hasanagić et al 2021, 37). Young people often consume information on the internet through the processes of peer sharing and scrolling (Hodžić and Sokol 2019, 50-51). They comment more often on content posted by their friends rather than the media and rarely report problematic content on social networks (Sokol and Alibegović 2021, 30-31). The participants of the consultative meetings pointed out that people in Bosnia and Herzegovina mostly do not understand online algorithms and do not know the terms and conditions of social networks.

Social networks are also important tools for reaching audiences. Media outlets use social media platforms to reach their audiences, and journalists increasingly use social media for the dissemination of the content that they produce in addition to searching for information and contacting potential interviewees (Đelilović 2021, 8-9). Civil society organizations also rely highly on social networks for the promotion of their own work and for reaching out to their audiences. Many, however, do not have social media managers or guidelines for using social networks, or sometimes even the skills and capacities to use them more efficiently.⁴⁷

Representatives of civil society organizations and the media in Bosnia and Herzegovina face a number of issues in using social networks that impact their work, including harmful content and attacks, threats and smear campaigns.⁴⁸ Studies also point to digital rights violations on social networks. For example, the monitoring of BIRN and the Share Foundation, conducted between 1 August 2019 and 30 November 2020, found 94 cases of digital violations in Bosnia and Herzegovina, among which 39 were cases of pressure through expression and activities on the internet (Perkov et al 2021, 21). In addition, studies show that harmful content often remains on Facebook and Twitter in the Western Balkan countries despite the fact that the harmful content is being reported. A poll conducted by BIRN showed that, according to the survey, 43% of content reported as hate speech remained online, while 57% was removed. When it comes to reports of threatening violence, content was removed in 60% of cases and around half of reports of targeted harassment resulted in removal (Jeremić and Stojanović 2021). Experts state that platforms perform less effectively in detecting harmful content that is not in English and

⁴⁷ Conclusions of the three meetings with civil society representatives in Sarajevo, Banja Luka and Mostar.

⁴⁸ Ibid.

that that there is inconsistent application of community guidelines (Jeremić and Stojanović 2021).

There have been cases when the content of civil society organizations and the media from Bosnia and Herzegovina were removed or suspended because they were detected to violate community standards. Representatives of civil society organizations state that inaccurate automatic detection of problematic content can curb freedom of expression in the country.⁴⁹ For example, in 2021 the Facebook account of the local investigative online portal Direkt was banned, but the representatives of the media outlet stated that it was not clear what rule they violated and they have not been able to appeal the decision in a satisfactory manner.⁵⁰ In May 2018, Facebook blocked the profile of Bosnian journalist Dragan Bursac for 24 hours after he posted a photo of a detention camp during the collapse of federal Yugoslavia in the 1990s (Jeremić and Stojanović 2021).

Interviewed experts criticize the broadly defined rules of social networks and automated communication with platforms, the lack of social network representatives in Bosnia and Herzegovina, inconsistent application of community guidelines and inadequate redress mechanisms.⁵¹ International organizations have also raised concerns over the impact of community guidelines of social networks on freedom of expression. The UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression stated that common concerns are that definitions of harmful content on social networks are very broad and are not specific about what kind of harm and what likelihood of harm will lead to content removal, labelling or other sanction. Other issues include over-reliance on automated filters that are unable to capture nuance or understand context, and lack of transparency and access to data, which makes it impossible to assess the effectiveness of the measures adopted by the companies and their impact on human rights (A/HRC/47/25). Another issue pointed out in the UN special report on disinformation is that companies do not invest sufficient resources in understanding the local factors and that greater transparency is needed on agreements to government removal of content (A/HRC/47/25).

Social networks are important platforms for freedom of expression and are used by individuals, groups and organizations for different purposes, including promotion of reconciliation and peace building in Bosnia and Herzegovina. However, they are also spaces where harmful content targets different groups, which calls for more transparent and accountable content moderation, especially in countries with weak democracies. The following is an overview of community guidelines of the most commonly used social networks in Bosnia and Herzegovina and those where harmful content is present, and the rules they set for dealing with different types of harmful content posted to their platforms, with a focus on hate speech and disinformation.

⁴⁹ Conclusions of the three meetings with civil society representatives in Sarajevo, Banja Luka and Mostar.

⁵⁰ Nikolija Bjelica, interview, Direkt portal.

⁵¹ Bojan Perkov, Share Foundation, interview.

8.1.1. Overview of community guidelines of social networks

When creating an account with any of the social media services today, users must agree to the Terms of Service, a contract that specifies the terms under which the user and the platform interact, the conditions of this relationship and liability, intellectual property, and arbitration of disputes to regulate content moderation activities and the behaviour of its users (De Streef et al 2020). This document is usually written in legal vocabulary and can thus be difficult to read or understand. It is accompanied by a page of “Community guidelines” (or similar), a more informal sounding text that uses plain language to explain to users what sort of behaviour and content is allowed and what is appropriate and what is not, as well as to list prohibitions, violations of which can result in content or page removal, with accounts being disabled altogether, and, in some cases, being reported to authorities. These rules usually fall into the following categories of forbidden or restricted content: sexual content (nudity, sex, pornography), graphic content (violence and obscenity), harassment (abuse, trolling, direct threats), hate speech, self-harm, illegal activity, and disinformation.

8.1.2. Facebook

Facebook’s Community Standards “apply to everyone, all around the world, and to all types of content”. The listed prohibitions include violence and incitement, coordinating harm and publicizing crime, fraud and deception, suicide and self-injury, child sexual exploitation, abuse and nudity, sexual exploitation of adults, bullying and harassment, human exploitation and privacy violations. Objectionable content includes hate speech, violent and graphic content, adult nudity and sexual activity and sexual solicitation.⁵² Additionally, Facebook requires that its users use their real name and will sometimes ask for an official ID to be submitted in order for the identity to be verified. This policy has often been criticized as being inconsistent with international standards on free expression and privacy and disproportionately affecting members of marginalized communities (who might want to use a name that is different from their legal name, or need to be anonymous online), and has been amended to a degree after interventions by human and digital rights organizations in 2015.

Users are given a set of reporting tools to flag and report inappropriate, offensive or abusive content or behaviour. According to their transparency reports,⁵³ Facebook’s technology proactively detects and removes more than 90% of violating content before anyone reports it and since November 2020, the prevalence of hate speech on the platform, as well as data about the removal of such content, is included in

⁵² See: <https://transparency.fb.com/en-gb/policies/community-standards/>

⁵³ See: <https://transparency.fb.com/en-gb/enforcement/detecting-violations/technology-detects-violations/>

the company's Community Guidelines Enforcement Report.⁵⁴ In cases where it is not entirely clear whether content violates community standards, it is forwarded to human moderators (Facebook has around 15,000 human moderators worldwide, according to their community standards) or escalated to content policy teams. Review systems use technology to prioritize high-severity content with the potential for offline harm and viral content which is spreading quickly. These automated systems and review teams also check content posted on Facebook's image and video sharing platform, Instagram.⁵⁵

Facebook's community standards contain a hate speech policy,⁵⁶ detailing what kind of expressions and statements users are discouraged from posting and which are subject to reporting and removal. Defining hate speech "as a direct attack against people – rather than concepts or institutions – on the basis of what we call protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease", Facebook's community standards tell users not to post:

- Content targeting a person or group of people (including all subsets except those described as having carried out violent crimes or sexual offences) on the basis of their aforementioned protected characteristic(s) or immigration status with violent speech or support in written or visual form; or dehumanizing speech or imagery in the form of comparisons, generalizations or unqualified behavioural statements (in written or visual form); mocking the concept, events or victims of hate crimes, even if no real person is depicted in an image; designated dehumanizing comparisons, generalizations or behavioural statements (in written or visual form);
- Content targeting a person or group of people on the basis of their protected characteristic(s) with: generalizations that state inferiority (in written or visual form), expressions of contempt, dismissal, disgust or cursing;
- Content targeting a person or group of people on the basis of their protected characteristic(s) with any of the following: segregation in the form of calls for action, statements of intent, aspirational or conditional statements, or statements advocating or supporting segregation; exclusion in the form of calls for action, statements of intent, aspirational or conditional statements, or statements advocating or supporting economic, social, political or explicit exclusion.

As of October 2020, Holocaust denial was added to the hate speech policy and it now prohibits "any content that denies or distorts the Holocaust".⁵⁷

⁵⁴ See: <https://transparency.fb.com/data/community-standards-enforcement/hate-speech/facebook/>

⁵⁵ See: <https://transparency.fb.com/hr-hr/policies/improving/prioritizing-content-review>

⁵⁶ See: <https://transparency.fb.com/en-gb/policies/community-standards/hate-speech/>

⁵⁷ See: <https://about.fb.com/news/2020/10/removing-holocaust-denial-content/>

When it comes to disinformation, Facebook has a policy related to false news⁵⁸ and manipulated media,⁵⁹ asking users not to post videos that have been edited or synthesized, beyond adjustments for clarity or quality, in ways that are not apparent to an average person, and would likely mislead an average person to believe: videos where the subject of the video said words that they did not say, and the video is the product of artificial intelligence or machine learning, including deep learning techniques (e.g. a technical deepfake), that merges, combines, replaces and/or superimposes content onto a video, creating a video that appears authentic (this policy does not extend to content that is parody or satire).

False news, when detected, is not necessarily removed from Facebook, but distribution is significantly reduced by showing it lower in News Feed.⁶⁰ In 2016, Facebook launched a partnership with International Fact-Checking Network (IFCN),⁶¹ a group of third-party fact-checking organizations, whose members are able to review content across both Facebook and Instagram, including organic and boosted posts. They can also review videos, images, links or text-only posts. The fact-checking initiative Raskrinkavanje, part of the Association Why Not, joined the programme in 2020 and is the only fact-checking organization from Bosnia and Herzegovina included in it.

Facebook's community standards also list these strategies aimed at addressing viral misinformation:

- Disrupting economic incentives for people, pages and domains that propagate misinformation
- Using various signals, including feedback from our community, to inform a machine learning model that predicts which stories may be false
- Reducing the distribution of content rated as false by independent third-party fact-checkers
- Empowering people to decide for themselves what to read, trust and share by informing them with more context and promoting news literacy
- Collaborating with academics and other organizations to help solve this challenging issue

Facebook Oversight Board, a body formed in 2020 to make precedent-setting content moderation decisions and consisting of twenty international experts,⁶² has so far made 23 decisions⁶³ on a case by case basis, nine of which concerned hate speech and two concerning misinformation.

⁵⁸ <https://transparency.fb.com/en-gb/policies/community-standards/false-news/>

⁵⁹ <https://transparency.fb.com/en-gb/policies/community-standards/manipulated-media/>

⁶⁰ See <https://transparency.fb.com/hr-hr/policies/community-standards/false-news/>

⁶¹ See: <https://www.facebook.com/journalismproject/programs/third-party-fact-checking/how-it-works>

⁶² About the Oversight Board see: <https://oversightboard.com/>

⁶³ See: <https://oversightboard.com/decision/>

In June 2020, Facebook opened journalist registration as part of its Facebook Journalism programme,⁶⁴ allowing registered journalists access to additional tools and increased account protection against online abuse. The programme is so far only available in the United States, Mexico, Brazil, and the Philippines. Through Facebook Protect,⁶⁵ high-profile accounts are encouraged to adopt stronger account security protections by simplifying security features. In October 2021, the platform announced that they would grant journalists and activists the status of “involuntary public figures”, with the aim of increasing the safety measures against harassment and bullying of these user groups (Reuters 2021).

As the most widely used social media platform in Bosnia and Herzegovina (according to recent data, there are 1.8 million Facebook users in BiH, accounting for 99% of social media visits in the country),⁶⁶ Facebook’s policies significantly impact civil society and media, who rely on this service to promote their activities and causes. During consultation meetings conducted with representatives of local NGOs, the most common complaints about the functioning of the platform involved the difficulty of contacting platform representatives, insufficient clarity of procedures for removal of content or suspension of pages and subsequent appeals, as well as declining organic reach of their content.⁶⁷

According to their latest transparency report for the first half of 2021, Facebook received a total of 55 requests for user data from the government of Bosnia and Herzegovina, 50 of which were legal process requests and five were emergency disclosure requests.⁶⁸ In 72-80% of cases, some data was produced as a result of the request. During the same period, three restriction requests were filed for pages/accounts on Facebook and Instagram, based on local law.⁶⁹ According to the report, Facebook “restricted access to one item in Bosnia and Herzegovina for publishing misinformation” in 2020. However, upon a routine review, it was determined that access to this item was restricted in error and the mistake has been corrected since.

8.1.3. Instagram

The short version of Instagram’s Community Guidelines simply says:⁷⁰ “Post only your own photos and videos, and always follow the law. Respect everyone on Instagram; don’t spam people or post nudity.” Photo sharing social media platform Instagram (owned by Facebook since 2012) promises to remove content that contains credible threats or hate speech, content that targets private individuals

⁶⁴ See: <https://www.facebook.com/journalismproject/journalist-registration-facebook>

⁶⁵ See: <https://www.facebook.com/gpa/blog/facebook-protect>

⁶⁶ Taken from statista.com

⁶⁷ See also Article 19 (2018).

⁶⁸ See: <https://transparency.fb.com/data/government-data-requests/country/BA/>

⁶⁹ See: <https://transparency.fb.com/data/content-restrictions/country/BA/>

⁷⁰ See: <https://about.instagram.com/blog/announcements/instagram-community-guidelines-faqs>

to degrade or shame them, personal information meant to blackmail or harass someone, and repeated unwanted messages. Using a built-in reporting feature on user profiles, content items and direct messages, users can report spam, nudity or sexual activity, hate speech or symbols, racist language or activity, violence or dangerous organizations, bullying or harassment, selling illegal or regulated goods, intellectual property violations and suicide and self-injury. Reports are anonymous (except in the case of intellectual property infringement) and having an Instagram account is not mandatory to submit a report.

Posts and accounts that may contain false information are identified through both technology and user feedback, and Instagram is also a part of Facebook’s network of third-party fact checkers that work on flagging false information on the platform. When third-party fact checkers identify false information, altered content or content with missing context on Instagram, the platform may make it harder to find this content by filtering it from Explore and hashtag pages and reducing its visibility in Feed and Stories.⁷¹ When a post is rated false, partly false, missing context, or satire by third-party fact-checkers on Instagram or Facebook, the platform uses image matching technology to find further instances of this content and automatically labels identical content if it is posted elsewhere on Instagram or Facebook as false.⁷²

8.1.4. YouTube

YouTube’s Community Guidelines apply to all types of content on the platform, including videos, comments, links, and thumbnails.⁷³ The guidelines cover spam and deceptive practices, sensitive content, violent or dangerous content, regulated goods, and misinformation. YouTube’s approach to moderation of harmful content is summarized as “four R’s”: Removal of content that violates the policies as quickly as possible, Reduction of the spread of harmful misinformation and content that brushes up against the policy lines, Raising up authoritative sources when people are looking for news and information, and Rewarding trusted, eligible creators and artists.

To achieve this, YouTube uses a combination of human moderators and machine learning to detect potentially problematic content on a massive scale, as well as relying on the YouTube community as well as experts in their Trusted Flagger program to help them spot potentially problematic content by reporting it directly to the platform.⁷⁴ Once such content is identified, human content reviewers evaluate whether it violates YouTube’s policies (reviewers evaluate flagged videos against all of their Community Guidelines and policies, regardless of why the video was originally flagged). If it does, the content is removed and used to train machines for better coverage in the future. Offending videos can be removed or age-restricted

⁷¹ See: <https://help.instagram.com/2109682462659451>

⁷² See: <https://about.instagram.com/blog/announcements/combating-misinformation-on-instagram>

⁷³ <https://www.youtube.com/howyoutubeworks/policies/community-guidelines/>

⁷⁴ <https://support.google.com/youtube/answer/7554338?hl=en>

when deemed not appropriate for all audiences. YouTube's reports indicate that the most common reasons for removing content are child safety, nudity, violence and spam.⁷⁵ YouTube's content reviewers are instructed to protect content that has a clear educational, documentary, scientific, or artistic [EDSA] purpose. "A YouTube channel is terminated if it accrues three Community Guidelines strikes in 90 days, has a single case of severe abuse (such as predatory behavior), or is determined to be wholly dedicated to violating our guidelines (as is often the case with spam accounts). When a channel is terminated, all of its videos are removed."⁷⁶ From July until September 2021, YouTube removed more than 4 million channels.

YouTube defines hate speech as "content that promotes violence against or has the primary purpose of inciting hatred against individuals or groups based on certain attributes, such as race or ethnic origin, religion, disability, gender, age, veteran status, sexual orientation/gender identity." The policy informs users that "there is a fine line between what is and what is not considered to be hate speech. For instance, it is generally okay to criticize a nation-state, but not okay to post malicious hateful comments about a group of people solely based on their race" and that "not everything that's mean or insulting is hate speech."

When it comes to complying with local laws, YouTube offers an option to report content that violates a local law or someone's privacy and safety. Being part of Google, the platform is included in their transparency report about government requests to remove content, by country.⁷⁷ According to Google's transparency report, the government of Bosnia and Herzegovina has filed a total of six removal requests since 2011, naming 28 items, most for the reason of defamation and bullying/harassment. There is no data on how many of these were removed in response to requests.⁷⁸

In addition to Community Guidelines, creators who want to monetize content on YouTube must comply with its Monetization Policies, an additional set of terms of service and policies regulating copyright and Google's ad policies.⁷⁹

However, research conducted by Mozilla showed that YouTube's algorithm is recommending videos considered disturbing and hateful that often violate the platform's very own content policies and that people in non-English speaking countries are far more likely to encounter videos they consider disturbing (Mozilla 2021).

⁷⁵ <https://www.youtube.com/howyoutubeworks/progress-impact/responsibility/>

⁷⁶ YouTube Community Guidelines enforcement, <https://transparencyreport.google.com/youtube-policy/removals?hl=en>

⁷⁷ Government requests to remove content – Google Transparency Report

⁷⁸ Ibid.

⁷⁹ See: <https://www.youtube.com/howyoutubeworks/policies/monetization-policies/>

8.1.5. Twitter

Twitter's community guidelines come under the title of "Twitter Rules" and contain a list of rules aimed at preserving the users' safety, privacy, and authenticity. The list of forbidden types of content includes threats and glorification of violence, terrorism/violent extremism, child sexual exploitation, abuse/harassment (including wishing or hoping that someone experiences physical harm), hateful conduct (promoting violence against, threatening, or harassing other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease), promotion or encouragement of suicide or self-harm, sensitive media (graphic violence and adult content) and selling, buying, or facilitating transactions in illegal goods or services.

Users are also not allowed to disseminate spam, to impersonate other individuals, groups, or organizations in a manner that is intended to or does mislead, confuse, or deceive others, to post non-consensual nudity, or use the service for the purpose of manipulating or interfering in elections or other civic processes.

In enforcing these rules, Twitter's guidelines emphasize that context matters, and the following factors are included when considering a possible violation and enforcement of consequences: is the behaviour directed at an individual, group, or protected category of people; has the report been filed by the target of the abuse or a bystander; does the user have a history of violating Twitter policies; how severe is the violation; is the content a topic of legitimate public interest.

Users are given a wide range of reporting tools, giving them the option to report spam and fake accounts, sensitive and offensive content, and abusive behaviour. Users are also given clear instructions on how to appeal account suspension.⁸⁰

Upon violation, Twitter first notifies the offender of the rule that was broken and requires them to delete the offending tweet before being able to tweet again. Repeating an offense provokes a stronger enforcement action, such as removal of content or suspension of account. Certain severe violations, such as posting violent threats, non-consensual intimate media, or content that sexually exploits children, lead to immediate and permanent suspension of an account.

In November 2021, Twitter announced an addition to their private information policy, stating that photos or videos of private individuals that are posted without their permission will be taken down at their request.⁸¹ Twitter's rules also prohibit the posting of private information like addresses, phone numbers, IDs, and financial and medical records, in order to inhibit an online harassment practice called "doxxing" (publishing of private information without consent).

Like Facebook, Twitter's policy enforcement notes "newsworthiness" of content, meaning that public impact of the content, its source and availability of coverage are considered (for instance, if the tweet has the potential to impact the lives of

⁸⁰ <https://help.twitter.com/forms/general?subtopic=suspended>

⁸¹ https://blog.twitter.com/en_us/topics/company/2021/private-information-policy-update

large numbers of people, the running of a country, and/or it speaks to an important societal issue or is part of an ongoing situation and provides valuable information).⁸²

In the drafting of their community policies, Twitter is advised by a body called the Trust and Safety Council, which is made up of several advisory groups, each dedicated to issues critical to the health of public conversation, such as online safety and harassment, human and digital rights, suicide prevention and mental health, child sexual exploitation, and dehumanization.⁸³ Information about requests for removal of content or profiles based on legal demands are published in Twitter's Transparency Report.⁸⁴ Twitter received three legal demands for removal of content from Bosnia and Herzegovina in the period of their latest transparency report (January-June 2021), 50% of which were complied with.⁸⁵ Even though Twitter is not popular in Bosnia and Herzegovina, it has been used by opinion makers and there have been cases of blocking Twitter accounts because they violated community guidelines.

8.1.6. TikTok

TikTok's Community Guidelines "apply to everyone and to everything on TikTok". The guidelines state that the platform proactively enforces the guidelines using a mix of technology and human moderation even before content gets reported.⁸⁶ The guidelines prohibit violent extremism, hateful behaviour, illegal activities, violent and graphic content, promotion of suicide, self-harm and dangerous acts, harassment and bullying, and adult nudity and sexual activities.

Considering its audience, TikTok's community guidelines have a special section dedicated to the safety of minors. Users under the age of 16 are not allowed to use direct messaging and cannot host live streams or be featured in recommended content. The list of banned behaviours involving minors includes sexual exploitation, grooming, nudity and sexual activity, harmful activities (such as illegal substances or dangerous physical challenges), physical and psychological harm of minors, and crimes against children.

Like other platforms, TikTok also makes an exception for "educational, documentary, scientific, or artistic content, satirical content, content in fictional settings, counterspeech, and content in the public interest that is newsworthy or otherwise enables individual expression on topics of social importance".

TikTok's policy against the spread of misinformation emphasizes exchange of ideas but does not permit "misinformation that causes harm to individuals, our

community, or the larger public regardless of intent". Specifically, it discourages users from posting these types of content:

- Misinformation that incites hate or prejudice
- Misinformation related to emergencies that induces panic
- Medical misinformation that can cause harm to an individual's physical health
- Content that misleads community members about elections or other civic processes
- Conspiratorial content that attacks a specific protected group or includes a violent call to action, or denies a violent or tragic event occurred
- Digital Forgeries (Synthetic Media or Manipulated Media) that mislead users by distorting the truth of events and cause harm to the subject of the video, other persons, or society

According to their latest transparency report (January-June 2021), TikTok did not receive any government requests from Bosnia and Herzegovina in that period.⁸⁷ Even though it is far less popular compared to other social networks, it is still becoming used by more young people and there are high concerns about disinformation spread on its platform.

⁸² <https://help.twitter.com/en/rules-and-policies/enforcement-philosophy>

⁸³ <https://about.twitter.com/en/our-priorities/healthy-conversations/trust-and-safety-council>

⁸⁴ <https://transparency.twitter.com/>

⁸⁵ <https://transparency.twitter.com/en/reports/countries/ba.html>

⁸⁶ <https://www.tiktok.com/community-guidelines?lang=en#31>

⁸⁷ <https://www.tiktok.com/transparency/en-us/government-removal-requests-2021-1/>

9. EXAMPLES OF REGULATIONS OF HARMFUL CONTENT FROM EU COUNTRIES

Harmful content online is a global issue and states around the world have been trying to find ways to counter it. In recent years, several states have adopted laws that grant the authorities powers to compel social media platforms to remove content that they deem illegal, but the risk with such laws is that companies could make errors and “over-remove” content for fear of being sanctioned (A/HRC/47/25). In the EU, The Digital Services Act will require platforms and other intermediaries to adopt transparency and due process measures that could help address the issues of harmful content online. However, civil society organizations argue that states must refrain from unnecessary content regulation, “and, when such regulation is necessary, they must do so in the least-restrictive manner possible to protect media pluralism, diversity, and human rights –and that transparency, accountability, and human rights should be central to any regulatory framework” (Article 19 2021). The following chapters give an overview of some of the legislative frameworks and self-regulatory tools and initiatives that have been proposed or adopted in different states and by different organizations, which could serve if similar frameworks or initiatives are set up in Bosnia and Herzegovina.

9.1. NETWORK ENFORCEMENT ACT

The Network Enforcement Act (*German: Netzwerkdurchsetzungsgesetz or NetzDG*) is a German law introduced in 2017 with the aim to combat online hate speech and disinformation on social networks.⁸⁸ The Act obligates social media platforms that have two million or more registered users in Germany to remove content that is “clearly illegal” within 24 hours after receiving a user complaint. The “unlawful content” is defined in 22 provisions of the Criminal Code and ranges widely, from insult of public officials to actual threats of violence. If the content is not illegal in an obvious manner, the platform has seven days to investigate and delete it and noncompliance can result in a fine of up to 50 million euros. NetzDG has been criticized, however, as unconstitutional, in particular with regard to free speech, and as not user-friendly, in particular with regard to the complaint mechanism (Human Rights Watch 2018). Since the law entered into force, several political parties have submitted proposals to amend or repeal it.

In June 2021, the law was amended to introduce an appeals procedure and arbitration, increase user friendliness of complaint procedures about unlawful content, and expand the powers of the Federal Office of Justice. The amendment introduces an appeals procedure for measures taken by the social network provider. The powers of the Federal Office of Justice are expanded to include supervisory powers. Due to new requirements under the EU Audiovisual Media Services Directive, video-sharing platform services are included in the scope of the Network Enforcement Act (Gesley 2021). In 2022, Germany introduced harsher measures that oblige online platforms to report certain types of “criminal content” to the Federal

⁸⁸ See: Bundesministerium der Justiz und für Verbraucherschutz, *FAQ: Act to Improve Enforcement of the Law in Social Networks*, 2017, <https://www.bmjv.de/SharedDocs/FAQ/EN/NetzDG/NetzDG.html>.

Criminal Police Office, but a German court ruled against it and online companies have taken legal action against this version of the law.

9.2. COMMUNICATIONS PLATFORM ACT

The new Communication Platforms Act that came into force in 2021 in Austria applies to domestic and foreign providers of for-profit communication platforms that have more than 100,000 users in Austria or revenues in Austria of more than EUR 500,000.⁸⁹ The new law came into force on 1 January 2021. The rules require platforms to: appoint and notify the authority of a) an authorized representative who is responsible for compliance with the new laws, and b) a representative who is able to accept official communications on behalf of the platform (it may be the same person, but they must speak German); set up an “effective and transparent procedure” for reporting and deleting illegal content whereby: a) deletion must take place within 24 hours if the illegality is “obvious to a legal layman (...)”, or within 7 days if a detailed examination is necessary; b) there must be a complaints procedure for users affected by deletion or blocking to avoid “overblocking”; c) illegal content comprises, for example, defamation, harassment, pornography involving minors, racist, discriminatory, or national-socialist content, unauthorized photographs, and stalking by means of telecommunication. The Law has been criticized in that it is likely to impede freedom of expression, delegating censorship power to private companies (Staber and Stutz 2021; Article 19 2021).

9.3 THE DIGITAL SERVICES ACT

The Digital Services Act is a legislative proposal by the European Commission submitted to the European Parliament and the European Council in December 2020, which the European Parliament approved in January 2022. The DSA sets out rules for online intermediary services, including social media platforms, and notes that “very large online platforms pose particular risks in the dissemination of illegal content and societal harms – specific rules are foreseen for platforms reaching more than 10% of 450 million consumers in Europe” (Facebook, Google, TikTok and Twitter all meet this requirement). The DSA maintains the current rule according to which companies that host others’ data are not liable for the content unless they actually know it is illegal, but when illegal content is found and flagged, companies are required to remove it. The DSA also aims at improving the transparency of social networks – it will oblige platforms to disclose to regulators how their algorithms work, how decisions to remove content are taken and the way advertisers target users, and companies that do not comply with the new obligations will risk fines

⁸⁹ European Commission, *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*, December 15, 2020, <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN>

of up to 6% of their annual turnover. Very large online platforms will be subject to specific obligations because they impose risks of harmful content which might not be illegal, and the spread of disinformation will be tackled by provisions on mandatory risk assessments, risk mitigation measures, independent audits and the transparency of so-called “recommender systems” (algorithms that determine what users see).⁹⁰ Digital rights activists however have criticized the Act because the whole process relies heavily on automated filtering in real time, which has been proven not to be very reliable and will likely result in over-blocking of content (Penfrat 2020).

9.4. THE EU CODE OF CONDUCT ON COUNTERING ILLEGAL HATE SPEECH ONLINE

The European Commission and Facebook, Microsoft, Twitter and YouTube agreed on a “Code of conduct on countering illegal hate speech online” in May 2016. Since then, Instagram, Snapchat, TikTok and LinkedIn have also joined the Code of Conduct. The implementation of the Code is conducted through regular monitoring, with input from a network of organizations from different EU countries and the tech companies themselves. According to the latest round of reports from October 2021, the companies involved are now assessing 81% of flagged content within 24 hours and 62.5% of the content deemed illegal hate speech is removed.

⁹⁰ See: <https://www.europarl.europa.eu/news/en/press-room/20220114IPR21017/digital-services-act-regulating-platforms-for-a-safer-online-space-for-users>

10. ADDITIONAL INITIATIVES TO INFLUENCE SOCIAL MEDIA CONTENT MODERATION

10.1. THE SANTA CLARA PRINCIPLES

The Santa Clara Principles are principles of content moderation proposed in 2018 by a group of organizations, advocates, and academic experts who support the right to free expression online.⁹¹ The three principles aim to address the issue of how best to obtain meaningful transparency and accountability around internet platforms' moderation of user generated content, and outline the initial steps that companies engaged in content moderation should take. The proposed principles are: 1) Numbers - Companies should publish the numbers of posts removed and accounts permanently or temporarily suspended due to violations of their content guidelines; 2) Notice - Companies should provide notice to each user whose content is taken down or account is suspended about the reason for the removal or suspension; 3) Appeal - Companies should provide a meaningful opportunity for timely appeal of any content removal or account suspension. Twelve major tech companies—including Apple, Facebook (Meta), Google, Reddit, Twitter, and Github—have endorsed the Santa Clara Principles, and in 2021 an updated and more detailed iteration of the principles was published. The new iteration of the principles “expands the scope of where transparency is required with respect to what is considered ‘content’ and ‘action’ taken by a company. The term ‘content’ refers to all user-generated content, paid or unpaid, on a service, including advertising. The terms ‘action’ and ‘actioned’ refer to any form of enforcement action taken by a company with respect to a user’s content or account due to non-compliance with their rules and policies, including (but not limited to) the removal of content, algorithmic downranking of content, and the suspension (whether temporary or permanent) of accounts.”

10.2. THE CHRISTCHURCH CALL

The Christchurch Call to Action to Eliminate Terrorist and Violent Extremist Content Online is a commitment by Governments and tech companies launched by the New Zealand government, in partnership with France, following a terrorist attack that was streamed online in May 2019. As part of the initiative, online service providers pledge to “take transparent, specific measures seeking to prevent the upload of terrorist and violent extremist content and to prevent its dissemination on social media and similar content-sharing services, including its immediate and permanent removal, without prejudice to law enforcement and user appeals requirements” and “provide greater transparency in the setting of community standards or terms of service”, “enforce those community standards or terms of service in a manner consistent with human rights and fundamental freedoms”, “implement immediate, effective measures to mitigate the specific risk that terrorist and violent extremist content is disseminated through livestreaming”, “implement regular and transparent

⁹¹ See: <https://santaclaraprinciples.org/>

public reporting” and “review the operation of algorithms and other processes that may drive users towards and/or amplify terrorist and violent extremist content”.⁹²

10.3. THE MANILA PRINCIPLES ON INTERMEDIARY LIABILITY

The Manila Principles are a set of standards for censorship and takedown laws, developed in 2015 by Electronic Frontier Foundation (EFF) and other NGOs to guide governments towards implementing laws that protect free expression. The Principles state that intermediaries should be shielded from liability for third-party content, content must not be required to be restricted without an order by a judicial authority, requests for restrictions of content must be clear, be unambiguous, and follow due process, laws and content restriction orders and practices must comply with the tests of necessity and proportionality, laws and content restriction policies and practices must respect due process, and transparency and accountability must be built into laws and content restriction policies and practices.⁹³

10.4. SOCIAL MEDIA COUNCILS

Proposed in 2018, the idea of social media councils was first brought forward by the international human rights organization Article 19, which proposed that “This model could include a dedicated ‘social media council’⁹⁴ – a multi-stakeholder, voluntary-compliance mechanism inspired by the effective self-regulation models created to promote journalistic ethics and high standards in print media. The key objectives of the SMC, as set out by Article 19, are to review individual content moderation decisions made by social media platforms on the basis of international standards on freedom of expression and other fundamental rights; provide general guidance on content moderation practices to ensure they follow international standards on freedom of expression and other fundamental rights; act as a forum where stakeholders can discuss recommendations; use a voluntary-compliance approach to the oversight of content moderation that does not create legal obligations. By acting as a multi-stakeholder forum the model proposes to solve several identified challenges of current content moderation practices: antagonism between stakeholders, no external oversight of content moderation decisions, no remedy for individual users, opacity and unilateral decision making.⁹⁵ A pilot of the model is being developed in Ireland, by a council made up of representatives of five categories of stakeholders: social media companies, media and advertising industries, journalists, academics, and civil society organizations.

⁹² See: <https://www.christchurchcall.com/call.html>

⁹³ See: <https://www.eff.org/sites/default/files/manila-principles-background-paper-0.99.pdf>

⁹⁴ See: [Social Media Councils: One piece in the puzzle of content moderation – ARTICLE 19](#)

⁹⁵ See: <https://www.article19.org/wp-content/uploads/2021/10/A19-SMC.pdf>

Article 19 launched a campaign called Missing Voices in 2019, calling on Facebook, Twitter and Google to provide more transparency about the number of content removals, types of flaggers, reasons for removal, the number of appeals they receive and the outcome of the appeals, to provide more transparency about the algorithms used to remove content, which can perpetuate bias or fail to detect context, and to give all users a right to appeal when their content is removed or their account closed down.⁹⁶

⁹⁶ See: <https://www.article19.org/campaigns/missingvoices/>

11. CONCLUSION AND RECOMMENDATIONS

The scope of harmful content online in Bosnia and Herzegovina is worrying but it often remains out of the reach of the legislative, regulatory and self-regulatory frameworks. Legislative and regulatory frameworks are often not applicable to the online sphere, while the community guidelines of social networks are not effective enough in removing harmful content online, as studies and interviewed experts point out. In many cases the existing legislative frameworks are not aligned with international standards, and what is worrying is that there are initiatives to introduce more punitive measures that could curb freedom of expression.

The legislative framework for the regulation of hate speech in BiH is fragmented and unaligned with international standards, and hate speech cases are mostly unprocessed. Crime codes of BiH, FBiH and the Brčko District are limited to prohibiting inciting and inflaming hatred on national, racial and religious grounds, but do not include other protected categories such as skin colour, gender, sexual orientation, and disability. Appropriate legislation has not been set up that would criminalize distributing and making available racist and xenophobic material, threats motivated by racism and xenophobia and insults motivated by racism and xenophobia through a computer system. Criminal codes do not sanction the creation or leadership of a group which promotes racism or the support for such a group and the participation in its activities. In addition, there are no initiatives that would address hate narratives that are not illegal but can impact democratic processes as most of the efforts so far have been focused on legislative and punitive measures.

Bosnia and Herzegovina lacks a strategy and an action plan that would combat disinformation and involve diverse actors such as governments, non-governmental organizations, regulators, self-regulators and academics. Efforts aimed at combating disinformation are mainly restricted to self-regulatory frameworks, individual media outlets, media watchdog platforms, the fact-checking platform and social networks. Media and information literacy is not part of the school curriculum across the whole country and the public generally lacks media and information literacy skills that would enable them to identify and report disinformation and problematic media content on the internet and social networks.

Ethno-nationally and/or politically biased media reporting is largely present both in the traditional and online media and the efforts of the regulators and self-regulators to combat such reporting have not provided the needed results. The media are divided along ethno-national and political party lines, influenced by political party groups through media financing and ownership lines, and the media selectively report on current affairs and past war events. Such reporting further polarizes the public, giving space to divisive inflammatory and nationalistic rhetoric of politicians that raises hatred towards communities in Bosnia and Herzegovina and incites and legitimizes hate speech among the citizens.

In 2021 the Criminal Code in Bosnia and Herzegovina was amended to prohibit public condoning, denial, gross diminution or attempts to justify the crimes of genocide,

crimes against humanity and war crimes established by final judgments, directed against a group of persons or a member of a group determined by race, colour, religion, background or national or ethnic origin, in a manner that could incite violence or hatred and for glorifying persons sentenced for genocide, crimes against humanity or a war crime. Even though these amendments follow international standards, they provoked a political crisis, as the government of Republika Srpska introduced a law on non-application of the decision and boycotted the work of central government institutions. Initial media monitoring showed that initially instances of genocide denial decreased but at the beginning of 2022 there were incidents of glorifying war crimes, and the failure to prevent and sanction these crimes poses harms to democracy.

The amount of harmful content online and the lack of appropriate tools and sanctions call for a comprehensive response to address issues that affect democratic processes and raise hatred towards others. Such a response needs to involve diverse actors from governments, the regulator, self-regulators, civil society organizations, international organizations, the media, social networks and educational institutions. Bosnia and Herzegovina should adopt measures that are in compliance with international human rights law and refrain from introducing tools that might affect freedom of expression.

Recommendations:

State level authorities

1. State level authorities, authorities of the entities and the Brčko District should amend the criminal codes in order to align them with international standards, the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD), and the recommendations of the European Union Commission and the European Commission against Racism and Intolerance. The International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) requires of states to incriminate any spread of ideas based on racial superiority or hatred, any incitement to racial discrimination, any act of violence or incitement to such violence, directed against any race or group of persons of another colour or other ethnic origin, as well as providing assistance to racist activities, including their financing, and requires them to declare as illegal and to ban organizations as well as organized and all other propaganda activities that encourage racial discrimination and participation in those activities.
2. The state and entity level authorities, including authorities of the Brčko District, should implement the provisions from the Council of Europe Convention on Cybercrime and its additional Protocol, that require states at the national level legislature to adopt as criminal offences the following actions done through a computer system: distributing and making available racist and xenophobic material, threats motivated by racism and xenophobia, and insults motivated by racism and xenophobia.

3. The Council of Ministers of BiH (particularly the Ministry of Communication and Transport of Bosnia and Herzegovina) should prepare appropriate legislation on media transparency ownership, media financing and media ownership concentration. It should include all types of media, including online, and all media-related companies (such as advertising companies).
4. The governments of the entities should amend defamation laws in accordance with the case law of the European Court of Human Rights to incorporate different standards of acceptability, tolerance and proof when it comes to alleged defamation against public figures in relation to private persons.
5. The government of the Brčko District and cantonal governments should repeal provisions in the laws on public order that contain misdemeanour sanctions for spreading false news or allegations that can cause harassment of citizens and endanger public order and peace, as these have been used for violating freedom of expression of citizens.
6. As pointed out in the European Union Commission's progress report in 2021, the authorities need to take concrete measures to counter hate speech particularly in official discourse. These could include voluntary codes of conduct for political parties and government representatives at different levels.
7. The authorities should refrain from unnecessary content regulation, and should protect media pluralism, diversity, and human rights – including the right to freedom of expression. Any laws aimed at regulating harmful content online should comply with international human rights law. The authorities should also report their involvement in content moderation decisions and requests to remove content on social networks. Any further legislation on harmful content online should comply with international human rights law.

The Central Election Commission

8. With a view to preventing undue limitation of the freedom of expression and ensuring consistent decisions, the definition of prohibited speech in the Election Law should be made more precise. The Central Election Commission could adopt further guidance for stakeholders and its own decision making. The Central Electoral Commission's mandate to monitor the use of hate speech should be expanded to the entire duration of election campaigns, not just for thirty days before the date of the elections. It should be extended to allow its oversight of early campaigning, including through online content, which includes strengthened procedures for determining the responsibility of political subjects in cases of hate speech, disinformation and other inappropriate material.

The Communications Regulatory Agency

9. The Communications Regulatory Agency should improve its Code to include a more precise definition of harmful content and their types, especially taking into consideration the local context and the delicate political situation in the country. It should include provisions on condoning, denial and justification of crimes of genocide, crimes against humanity and war crimes in line with the amendments to the Criminal Code. Since ethno-nationally and/or politically biased media reporting has been the main problematic media content in traditional and online media, the regulator should increase its capacities to conduct periodical monitoring of the traditional media with the aim to sanction the media that do not follow the principles of fairness and impartiality. It could also consider including a definition of disinformation in its Code, especially in cases of adverse consequences for human safety and health. It needs to implement the EU Directive 2018/1808, which would include its co-regulation over video-sharing platforms and should particularly take into consideration co-regulating content that is particularly harmful in the context of Bosnia and Herzegovina. It should also consider co-regulating the online media and sanctioning the most extreme examples.

The judiciary and law enforcement agencies

10. The judiciary in Bosnia and Herzegovina should follow the practice of the European Court of Human Rights according to which only serious and extreme examples should be criminalized. The Rabat Plan of Action suggests a high threshold for defining restrictions on freedom of expression and incitement to hatred, and for the application of Article 20 of the International Covenant on Civil and Political Rights. The courts need to sanction examples of hate speech and denials of war crimes and glorification of war criminals particularly of those who have a large influence, such as politicians and public officials. Police enforcement agencies and the members of the judiciary should receive education on illegal harmful content, international standards in countering illegal harmful content and on Article 10 of the European Convention on Human Rights.

The Press Council of Bosnia and Herzegovina

11. The Press Council of Bosnia and Herzegovina should promote the use of the new Press Code in order to increase professional standards in the online media, particularly in relation to disinformation, hate speech, the use of information technologies and content in user-generated comments. Together with online media and media organizations it should consider preparing guidelines for online media on moderating user-generated comments and following the new provisions in the Code. It should strengthen its capacities to follow cases of disinformation and politically and ethno-nationally biased media reporting and could periodically conduct monitoring of online media.

The members of the Complaints Commission should receive education on freedom of expression and improve the transparency of their decision-making.

Media and civil society organizations

12. Media watchdog organizations and civil society organizations should set up a monitoring system of harmful content online and react to harmful content online. Together with the Press Council of Bosnia and Herzegovina, the Regulatory Agency for Communication and the media, they could develop editorial guidelines for the media on reporting on inflammatory and divisive ethno-national language, on reporting on denials of war crimes and issues that are sensitive in the context in Bosnia and Herzegovina. They could provide training to journalists on constructive media reporting and develop tools and manuals for the media on how to fact-check information and provide information to the public in a constructive manner without further inciting hate. Media organizations along with other stakeholders should consider preparing a nation-wide campaign on harmful content online and the tools that could be used to address it. Media and civil society organizations should be capacitated to gain expertise in digital rights and digital technologies.

The media

13. The media should be capacitated and trained to fact-check information and report on issues that are sensitive to the society in Bosnia and Herzegovina. They should adopt editorial guidelines on reporting on inflammatory statements of politicians, on denials of war crimes and glorification of war criminals, and on moderating comments sections online.

Political parties

14. Politicians should refrain from using harmful content. They should receive appropriate training on harmful content online and adopt codes of conduct (self-regulatory tools) with the aim to tackle it. They should also receive training on the use of social networks and on the rules of political campaigns and advertising especially during the election periods.

Social networks

15. Social networks should:
 - appoint specific representatives, or make their representatives more available, in Bosnia and Herzegovina and the Western Balkans.
 - follow the recommendations for content moderation set out in the Santa Clara Principles that call companies to respect human rights in all stages of the content moderation process.

- take into account the specific local context of Bosnia and Herzegovina and ensure, as the principles set out, that users “have access to rules and policies and notice, appeal, and reporting mechanisms that are in the language or dialect with which they engage”.
- publish information about accounts actioned, broken down by the categories of rules violated, as well as about the automated processes used.
- “provide a meaningful opportunity for timely appeal of decisions to remove content” as set out in the Santa Clara Principles.
- set out clear and precise rules and policies, in line with these recommendations, relating to when action will be taken with respect to users’ content or accounts.
- provide clearly accessible means of appealing content moderation decisions.

Specific recommendations:

Media and information literacy

16. The state level government should adopt a Strategy on Media and Information Literacy that would set out guidelines and action plans for its introduction in the educational system of Bosnia and Herzegovina, including provisions for teaching materials and education of teachers. The government of Republika Srpska should consider including the subject Digital World, which was introduced in the school year 2021/22, in other grades as well and introduce Media Literacy in all high schools across the country. The government of the Federation of Bosnia and Herzegovina and cantonal governments should introduce subjects in primary and secondary education on media and information literacy. The non-formal coalition of media and information experts that has been set up by the Communications Regulatory Agency should prepare or give expert opinion on educational materials that should be provided for education on media and information literacy. Non-governmental organizations should provide non-formal media and information literacy workshops especially on harmful content online to all age-groups including youth and the elderly. The workshops should include information on the algorithms on social networks, on harmful content online, digital security, and teaching and developing skills to recognize and report harmful content online. Education should also be provided on freedom of expression and international standards in regulation of harmful content.

Coalition against harmful content online

17. A coalition to counter harmful content online comprised of civil society representatives, media representatives, representatives of the Communications Regulatory Agency, the Press Council of Bosnia and Herzegovina, the Institution of the Human Rights Ombudsman and academics could be established. It could propose activities, make recommendations for the improvement of the tools in place in Bosnia and Herzegovina to combat harmful content online, work with the media in establishing editorial guidelines and in preparing manuals and educational activities, mediate with Meta (Facebook) on removals of content and banning of the accounts of the representatives of civil society and journalists, and on cases of smear campaigns targeting journalists and civil society representatives that aim to intimidate and discredit them. The capacities of the Coalition against Hate Speech and Hate Crimes in BiH could be expanded to form such a coalition. The Coalition could be based on the model of social media councils prepared by Article 19 - a multi-stakeholder, voluntary-compliance mechanism inspired by the effective self-regulation models created to promote journalistic ethics and high standards in print media. The coalition could also conduct periodic monitoring of harmful content online and present the reports to the public, social networks and other stakeholders.

Monitoring mechanism

18. International and/or civil society organizations should set up a monitoring mechanism for harmful content online and focus on content that is harmful in the context of BiH. The mechanism could involve the five types of content identified in the research. They should regularly or periodically conduct monitoring of diverse platforms online, including online media, comments sections and social networks. They should collaborate and take data from actors such as the Institution of the Human Rights Ombudsman and the police. The results of the monitoring initiatives should be used to follow trends and also to set up advocacy initiatives aimed at protecting freedom of expression and preventing the spread of information that could be harmful to others.

REFERENCE

- Adilagić, Rea. (2019). *Od viktimizacije do demonizacije: Gdje je istina? Istraživanje o načinima izvještavanja medija o migrantima i izbjeglicama*. Udruženje Bh novinari. https://safejournalists.net/wpcontent/uploads/2019/02/od_viktimizacije_do_demonizacije_gdje_je_istina_BHN_feb_2019.pdf
- Aktis, Aktek and Mediacentar. (2018). *Radicalisation in Flux: Unpacking the Role of the Internet in Violent Extremism (Bosnia and Herzegovina)*. Unpublished. Available for Reference. Article 19. (2020). *Austria: the draft Communication Platforms' Act fails to protect freedom of expression*. <https://www.article19.org/resources/austria-draft-communication-platforms-act-fails-freedom-of-expression/>
- Article 19 (2018). *Facebook Community Standards: Analysis against international standards on freedom of expression*. <https://www.article19.org/resources/facebook-community-standards-analysis-against-international-standards-on-freedom-of-expression/>
- Article 19. (2020). *Austria: the draft Communication Platforms' Act fails to protect freedom of expression*. <https://www.article19.org/resources/austria-draftcommunication-platforms-act-fails-freedom-of-expression/>
- Article 19. (2021). *Watching the Watchmen: Content Moderation, Governance, and Freedom of Expression*. <https://www.article19.org/resources/international-regulating-content-moderation-who-watches-the-watchmen/>
- Aščerić, Selma. (2021). *Šta bi za građane i građanke FBiH značilo proglašavanje interneta javnim mjestom? Zašto ne?* <https://zastone.ba/sta-bi-za-gradane-igradanke-fbih-znacilo-proglasavanje-interneta-javnim-mjestom/>
- Association BH Journalists. (2021). *A member of the AF BiH sentenced to three months probation for hate speech against Martina Mlinarević*. <https://bhnovinari.ba/en/2021/06/11/a-member-of-the-af-bih-sentenced-to-three-months-probation-for-hate-speech-against-martina-mlinarevic/>
- Augustinović, Marija and Milorad Milojević. (2021). *Vakcine propadaju, samo 20 posto vakcinisanih*. Radio Free Europe. <https://www.slobodnaevropa.org/a/bih-vakcine-unistavanje-vakcinacija/31555241.html>
- Council of Europe study, DGI(2017)11. *Study on the Use of Internet in Electoral Campaigns*. <https://rm.coe.int/use-of-internet-in-electoral-campaigns-/16807c0e24>
- Blagovčanin, P. (2019). *Stranački botovi i kontaminacija online prostora dezinformacijama*. MC Online. Media.ba <https://media.ba/bs/magazin-novinarstvo/stranacki-botovi-i-kontaminacija-online-prostora-dezinformacijama>
- Blažević, Jozo. (2019). *Izveštće o pojavama govora mržnje i kaznenih djela učinjenih iz mržnje u Bosni i Hercegovini u razdoblju od lipnja 2018. godine do lipnja 2019. godine*. Sarajevo Open Centre. <https://soc.ba/site/wpcontent/uploads/2019/10/Izve%C5%A1%C4%87e.pdf>
- Boračić-Mršo, Selma. (2022). *Napadi na novinarke u BiH bez sudskih epiloga i reakcije javnosti*. Media.ba. <https://www.media.ba/bs/magazin-novinarstvo/napadi-nanovinarke-u-bih-bez-sudskih-epiloga-i-reakcije-javnosti>
- Bubonjić, Mladen. (2022a). *Otkako je bivši direktor RTRS na poziciji direktora RAK, nema objavljenih kazni za RTRS*. Media.ba. <https://media.ba/bs/magazin-novinarstvo/otkakoje-bivsi-direktor-rtrs-na-mjestu-direktora-rak-nema-objavljenih-kazni-za>
- Bubonjić, Mladen. (2022b). *U RS-u "specijalna vojna operacija", u Federaciji "agresija na Ukrajinu*. Media.ba. <https://media.ba/bs/magazin-novinarstvo/u-rs-u-specijalna-vojnaoperacija-u-federaciji-agresija-na-ukrajinu>
- Coalition for the Fight Against Hate Speech and Hate Crimes. (2021). *Reakcija civilnog društva na presudu Fatmiru Alispahiću: Govor mržnje nije sloboda govora*. Network for Building Peace. <https://www.mreza-mira.net/vijesti/aktivnosti-mreze/reakcija-civilnog-drustva-na-presudu-fatmiru-alispahicu-govor-mrznje-nije-sloboda-govora/>
- Cvjetičanin et al. (2021). *Izveštaj o negiranju genocida u Srebrenici 2021*. Srebrenica Memorial Centre. <https://srebrenicamemorial.org/assets/files/1625819630-izvjestaj-o-negiranju-genocida-za-2021-godinu-bosanski-jezik.pdf>
- Cvjetičanin et al (2019). *Disinformation in the Online Sphere: The Case of BiH*. Zašto ne? https://zastone.ba/app/uploads/2019/05/Disinformation_in_the_online_sphere_The_case_of_BiH_ENG.pdf
- De Stree et al (2020). *Online Platforms Moderation of Illegal Content Online*. European Union. 2020. [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652718/IPOL_STU\(2020\)652718_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652718/IPOL_STU(2020)652718_EN.pdf)
- Dragičević, Franjo. (2019). *Govor mržnje, izborni proces i pravo na slobodu izražavanja*. Foundation Public Law Centre. http://fcjp.ba/analize/Franjo_Dragicevic_Govor_mrznje_izborni_proces_i_pravo_na_slobodu_izrazavanja.pdf
- Đelilović, Zinaida. (2021). *Društvene mreže i novinarstvo u Bosni i Hercegovini*. Press Council BiH. https://www.vzs.ba/images/2021/drustvene_mreze_i_novinarstvo_u_BiH.pdf
- Džekman, Vildana. (2021). *Podnošenje pa povlačenje tužbi kao način pritiska na medije*. Media.ba. <https://www.media.ba/bs/magazin-novinarstvo/podnosenje-pa-povlacenje-tuzbi-kao-nacin-pritiska-na-medije>
- Ferhatović, Amila, and Davor Trlin. (2019). *Krivičnopravni aspekt govora mržnje kao zloupotreba prava na slobodu izražavanja*. Pregled, social issues magazine, year LX, No. 2, May August 2019, p. 133-157.
- Filipović, Ljiljana. 2019. *Krivičnopravni aspekti govora mržnje na elektronskim medijima i na internetu*. Foundation Public Law Centre. http://fcjp.ba/analize/Ljiljana_Filipovic_Krivicnopravni_aspekti_govora_mrznje_na_elektronski_m_medijima_i_na_internetu.pdf
- First Draft. (2020). *Understanding Information disorder*. <https://firstdraftnews.org/long-form-article/understanding-information-disorder/>
- Gaćanica, Lejla. (2021). *Šta zabrana poricanja genocida, ratnih zločina i zločina protiv čovječnosti znači za medije?* Media.ba https://www.media.ba/bs/magazin-novinarstvo/sta-zabrana-poricanja-genocida-ratnih-zlocina-i-zlocina-protiv-covjecnosti-znaci#_edn1
- Gaćanica, Lejla and Caroline Finkeldey. (2019). *Calling War Atrocities by their Right Name: Regulating a Ban on Denial, Trivialisation, Justification or Condonation of Genocide, The Holocaust, Crimes Against Humanity or War Crimes*. ForumZDF – Trial International. <https://trialinternational.org/wp-content/uploads/2020/11/Policy-Brief-%E2%80%93-Regulating-a-Ban-on-Denial-of-War-Crimes.pdf>
- Gaćanica, Lejla and Marija Arnautović. (2018). *Mehanizmi zaštite od online nasilja: Priručnik za novinare i novinarke*. Mediacentar. https://www.media.ba/sites/default/files/mc_mehanizmi-zastite_web.pdf

Gesley, Jenny. (2021). "Germany: Network Enforcement Act Amended to Better Fight Online Hate Speech," *Library of Congress*. <https://www.loc.gov/item/global-legal-monitor/2021-07-06/germany-network-enforcement-act-amended-to-better-fight-online-hate-speech/>.

Halilović, Mehmed and Amer Džihana. (2017). *Medijsko pravo*. (Media Law). Internews. <https://www.mediaobservatory.net/sites/default/files/Medijsko%20pravo%20u%20BiH%20bos.pdf>

Hasanagić, Snježana, Maja Popović and Erna Lević. (2021). *Research on Media Habits of Adults in BiH*. Council of Europe. <file:///C:/Users/New%20guest/Desktop/Posljednji-final-NEW-CC-for-CoE-BHS-Final-Report-211021-002.pdf>

Hasečić, Nejra. (2021). *Novi ZOSPI bi mogao dodatno ograničiti slobodan pristup informacijama u BiH*. Media.ba. <https://www.media.ba/bs/magazin-novinarstvo/novizospi-bi-mogao-dodatno-ograniciti-slobodan-pristup-informacijama-u-bih>

Hodžić, Sanela and Anida Sokol. (2017). *Obrasci finansiranja medija iz javnih budžeta: Politički pritisci i finansijska nestabilnost*. Mediacentar Sarajevo. https://www.media.ba/sites/default/files/mc_obrasci_finansiranja_medija_-_final_0.pdf

Hodžić, Sanela and Anida Sokol. (2019). *Surfanje po tankom ledu: Mladi, mediji i problematični medijski sadržaji*. [Surfing on This Ice: Youth, Media and Problematic Media Content]. Mediacentar. https://media.ba/sites/default/files/mladi_mediji_problematichni_medijski_sadrzaji_web.pdf

Hodžić, Sanela, Brankica Petković and Sandra Bašić-Hrvatini. (2019). *Medijska pismenost u Bosni i Hercegovini: Brojne inicijative civilnog sektora i nedostatak javnih politika*. Mediacentar. https://www.media.ba/sites/default/files/medijska_i_informacijska_pismenost_u_bosni_i_hercegovini_final.pdf

Human Rights Watch. (2018). *Germany: Flawed Social Media Law: NetzDG is Wrong Response to Online Abuse*. <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law>

Ivanović, Aleksandar. (2021). *Studija mapiranja institucionalnog kršenja ljudskih prava u Bosni i Hercegovini*. Pravni fakultet u Lukavici. https://pravnikafakultet.edu.ba/wp-content/uploads/2021/08/Studija-mapiranja-institucionalnog-kršenja-ljudskih-prava-u-Bosni-i-Hercegovini_23082021.pdf

Ivandić-Ninković, Snježana. (2021). *Priručnik za državne službenike u Bosni i Hercegovini za prepoznavanje i postupanje u slučajevima govora mržnje*. Council of Europe. <https://rm.coe.int/prirucnik-za-web-final/1680a2cb52>

Išerić, Harun. (2022). *Šta donose izmjene Kodeksa za štampu i online medije?* Media.ba <https://media.ba/bs/magazin-novinarstvo/sta-donose-izmjene-kodeksa-za-stampu-i-online-medije-bih>

Jeremić, Ivana and Milica Stojanović. (2021). *Facebook, Twitter Struggling in Fight against Balkan Content Violations*, Detektor.ba. <https://detektor.ba/2021/02/16/facebook-twitter-struggling-in-fight-against-balkan-content-violations/?lang=en>

Karadžuz, Amar et al. (2022). *Bora protiv dezinformacijskih narativa i mapiranje teorija zavjere. Slučaj BiH*. Association Why Not. <https://zastone.ba/app/uploads/2022/04/Borba-protiv-dezinformacijskih-narativa-i-mapiranje-teorija-zavjere-Slucaj-BiH.pdf>

Kurspahić, Kemal. (2003). *Zločin u 19:30. Balkanski mediji u ratu i miru*. Mediacentar Sarajevo. https://www.media.ba/sites/default/files/2003_-_zlocin_u_19.30_balkanski_mediji_u_ratu_i_miru.pdf

Krupalija et al. (2020). *Disinformation during COVID-19 Pandemic*. Friedrich-Neuman Foundation. https://www.freiheit.org/sites/default/files/2021-05/disinformation_covid_19_march_2021.pdf

Lučić-Ćatić, Marija and Bajrakterević Pajević, Dina. (2017). *Nedostaci kaznene politike Tužiteljstva BiH I Suda BiH u slučajevima zločina iz mržnje i govora mržnje. Kriminalističke teme*, Journal of Forensics, Criminology and Security Studies, No. 5. <http://krimteme.fkn.unsa.ba/index.php/kt/article/view/163>

Mediega, Tambiama. (2020). *Reform of the EU liability regime for online intermediaries: Background on the forthcoming digital services act*, European Parliamentary Research Service. [https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/649404/EPRS_IDA\(2020\)649404_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/649404/EPRS_IDA(2020)649404_EN.pdf)

MCOOnline (2021a). *Mreža mira poziva medija na odgovorno izvještavanje i smirivanje*. Media.ba <https://media.ba/bs/vijesti/mreza-mira-poziva-medije-na-odgovorno-izvjestavanje-i-smirivanje>

MCOOnline. (2021b). *Urednik portala Antimigrant.ba negirao krivicu i odbio nagodbu sa Tužilaštvom BiH*. Media.ba <https://media.ba/bs/vijesti-i-dogadaji/vijesti-urednik-portala-antimigrantba-negirao-krivicu-i-odbio-nagodbu-sa>

Memorijalni centar Srebrenica. 2021. *Memorijalni centar Srebrenica: Smanjen broj slučajeva negiranja genocida u entitetu RS*. <https://srebrenicamemorial.org/bs/aktuelnosti/memorijalni-centar-srebrenica-smanjen-broj-slucajeva-negiranja-genocida-u-entitetu-rs/61>

Mozilla (2021). *YouTube Regrets A crowdsourced investigation into YouTube's recommendation algorithm*. https://assets.mofoprod.net/network/documents/Mozilla_YouTube_Regrets_Report.pdf

Omerović, Enis, and Hrustić, Amna. (2020). *Sloboda izražavanja i govor mržnje: Odgovor države Bosne i Hercegovine*. Annals of the Legal Faculty of the Zenica University, Vol. 13 Issue 25.

Osmančević, Enes i drugi. (2021). *Mapiranje medijskih web portala u Bosni i Hercegovini*. CPCD. <https://medijskapismenost.ba/objavljeni-rezultati-prvog-sveobuhvatnog-istrazivanja-o-broju-i-karakteristikama-web-portala-u-bih/>

Penfrat, Jan. (2020). "The EU's attempt to regulate Big Tech: What it brings and what is missing," *EDRI*. <https://edri.org/our-work/eu-attempt-to-regulate-big-tech/>

Perkov, Bojan i drugi (2021). *Panemic for Digital Rights*. Share Foundation and BIRN. <https://bird.tools/wp-content/uploads/2020/11/Pandemic-for-Digital-Rights-2020.pdf>

Petković, Brankica. (2021). *Regulatory and Self-Regulatory Frameworks against Hate Speech and Disinformation in the European Union*. Examples and Challenges. SEENPM. <https://seenpm.org/wp-content/uploads/2021/12/Resilience-Factsheet-EU.pdf>

Petković, Brankica, Sandra Bašić-Hrvatini and Sanela Hodžić. (2021a). *Hate narratives in the Western Balkans and Turkey*. SEENPM and Peace Institute. <https://seenpm.org/wp-content/uploads/2021/04/Resilience-E-book-research-2-all-reports-with-CIP-April-2021.pdf>

Petković, Brankica, Sandra Bašić-Hrvatini and Sanela Hodžić. (2021b). *Media Trust in the Western Balkans: Together Apart*. SEENPM and Peace Institute. <https://seenpm.org/research/regional-review-of-media-trust-in-the-western-balkans/>

- Petković, Brankica and Sanela Hodžić (2019). *Sustainability of Professional Journalism in the Media Business Environment of the Western Balkans*. TACSO. <http://tacso.eu/wp-content/uploads/2020/09/Sustainability-of-Professional-Journalism-in-WB-Study-National-Data-Overview-BiH-final.pdf>
- Radević, Maja. 2021. *Bosna i Hercegovina: Indikatori nivoa medijskih sloboda i sigurnosti novinara 2020*. <https://bhnovinari.ba/wp-content/uploads/2021/07/BiH-BiH-2020.pdf><https://www.parlament.ba/Publication/Read/15943?pagelid=238&lang=hr>
- Ti BiH. 2021. *Izveštaj o primjeni Zakona o slobodi pristupa informacijama u 2021*. <https://ti-bih.org/izvjestaj-o-primjeni-zakona-o-slobodi-pristupa-informacijama-u-2021-godini/>
- Sokol, Anida and Ehlimana Memišević (2021). *Bosnia and Herzegovina in Antisemitic Discourse in the Western Balkans*. IRI. https://www.iri.org/sites/default/files/antisemitic_discourse_in_the_western_balkans_04-15-21.pdf
- Sokol, Anida and Elvira Jukić-Mujkić. (2021). *Bosnia and Herzegovina. Vibrant Information Baramoter 2021*. IREX. <https://www.irex.org/sites/default/files/pdf/vibe-bosnia-herzegovina-2021.pdf>
- Sokol, Anida. (2019). *Media Consumption Habits in the Birač region*. Mediacentar Sarajevo. https://www.media.ba/sites/default/files/media_usage_habits_birac_region_0.pdf
- Sokol, Anida. (2020). *Propaganda, Disinformation and Hate Speech Models of Media and Communication in Bosnia and Herzegovina*. SEENPM, Peace Institute and Mediacentar. <https://seenpm.org/wp-content/uploads/2020/09/Research-publication-1-BiH-ENG.pdf>
- Sokol, Anida. (2021a). *Hate Narratives in the Media and User-Generated Content*. SEENPM, Peace Institute and Mediacentar. <https://seenpm.org/wp-content/uploads/2020/12/Resilience-research-publication-2-BiH-English.pdf>
- Sokol, Anida. (2021b). *Polarized Public's Trust in the Media and Social Networks in Bosnia and Herzegovina*. <https://seenpm.org/research-polarized-public-trust-in-the-media-and-social-networks-in-bosnia-and-herzegovina/>
- Sokol, Anida and Alibegović Senad. (2021). *Mladi za bolje medije: Stavovi, navike i potrebe mladih za medijskim sadržajima*. Mediacentar. https://www.media.ba/sites/default/files/mladi_za_bolje_medije_dtp_05.pdf
- Staber, Gabriela and Angelika Stutz. (2021). *Communication platforms face new obligations and high fines in Austria*. CMS Law Now. https://www.cms-lawnow.com/ealerts/2021/03/au-test?cc_lang=en
- Srebrenica Memorial Centre. 2021. *Memorijalni centar Srebrenica: Smanjen broj slučajeva negiranja genocida u entitetu RS*. <https://srebrenicamemorial.org/bs/aktuelnosti/memorijalni-centar-srebrenica-smanjen-broj-slucajeva-negiranja-genocida-u-entitetu-rs/61>
- Stokić, Vanja. (2021a). *Zašto pojedinim medijima ne odgovara rad fact-checkera?* (Why does the work of fact-checkers not suit certain media?). Media.ba <https://media.ba/bs/magazin-novinarstvo/zasto-pojedinim-medijima-ne-odgovara-rad-fact-checkera>
- Stokić, Vanja. (2021b). *Digitalni svijeti u osnovnim školama: Đaci uče o uzeđajima i sigurnosti ali i o bontonu*. Media.ba <https://media.ba/bs/magazin-novinarstvo/digitalni-svijet-u-osnovnim-skolama-daci-uce-o-uredajima-i-sigurnosti-ali-i>
- Sali-Terzić, Sevima. (2019). *Govor mržnje na internetu: međunarodni standardi i zaštita*. [Hate Speech on Internet: international standards and protection]. Report. Sarajevo: Foundation Public Law Centre. http://www.fcjp.ba/analize/Sevima_Sali_Terzic7_Govor_mrznje_na_internetu_medjunarodni_standardi_i_zastita.pdf

- Tadić-Stojisavljević, Sanja, and Trlin, Davor. (2021). *Govor mržnje online i drugom javnom prostoru*. Centre for Education of Judges and Prosecutors of the Federation of BiH. http://www.fbih.cest.gov.ba/images/doclink/DTP_GOVOR-MRNJE-5-FINAL-3-4-2021.pdf
- Tomić, Sladan. (2021a). *Regulacija sadržaja internet stranica televizijskih i radijskih emitera*. Media.ba. <https://www.media.ba/bs/magazin-novinarstvo/regulacija-sadrzaja-internet-stranica-televizijskih-i-radijskih-emitera>
- Tomić, Sladan. (2021b). *Slučaj Mulahusić. Hapšenje zbog govora mržnje na društvenim mrežama*. Media.ba <https://www.media.ba/bs/magazin-novinarstvo/slucaj-mulahusic-hapsenje-zbog-govora-mrznje-na-drustvenim-mrezama>
- Turčilo, Lejla. (2020). *U susret lokalnim izborima: Portali za diskreditaciju neistomišljenika i radikalizaciju birača*. Analiziraj.ba. <https://analiziraj.ba/2020/06/19/u-susret-lokalnim-izborima-portali-za-diskreditaciju-neistomisljenika-i-radikalizaciju-biraca>
- Učanbarlić, Učanbarlić. (2021). *Zakon o slobodi pristupa informacijama BiH. Prednosti i mane postojećeg Zakona, (ne)transparentnost institucija i analiza predloženih izmjena zakona*. Udruženje Bh. novinari. https://bhnovinari.ba/wp-content/uploads/2021/09/Analiza_ZoSPI-FINAL.pdf
- Vučić, Nikola. (2021). *Online nasilje i govor mržnje protiv novinara i novinarki u BiH*. Udruženje Bh novinari. <https://bhnovinari.ba/wp-content/uploads/2021/09/Analiza-govor-mrznje-FINAL.pdf>
- Zulejhić, Emir et al. (2020). *Preko margine: Dezinformacije o marginaliziranim grupama u bh. medijima*. [Over the Margins: Disinformation about marginalized groups in the BH media]. Zašto ne. https://zastone.ba/app/uploads/2020/10/Preko-margine_1stra%C5%BEivanje.pdf
- ZFD Forum. (2019). *Balkan Perspectives. Dealing with the Past: The Role of the Media*. 13th edition. <https://www.forumzfd.de/en/dealing-past-role-media>
- Wardle, Claire and Hossein Derakhshan. (2017). *Information Disorder. Toward an interdisciplinary framework for research and policymaking*. Council of Europe. <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-forresearch/168076277c>

REPORTS

- A/66/290
<https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/A.66.290.pdf>
- A/HRC/47/25
<https://daccess-ods.un.org/tmp/7547888.75579834.html>
- A/74/486
<https://daccess-ods.un.org/tmp/7903608.08372498.html>
- A/HRC/38/35
<https://daccess-ods.un.org/tmp/5832985.0435257.html>
- CCPR/C/GC/34
<https://daccess-ods.un.org/tmp/4002248.94285202.html>
- Commission Opinion on Bosnia and Herzegovina's application for membership of the European Union, <https://ec.europa.eu/neighbourhood-enlargement/system/files/2019-05/20190529-bosnia-and-herzegovina-opinion.pdf>

European Commission against Racism and Intolerance. 2017. ERCI report on Bosnia and Herzegovina. Fifth monitoring cycle. <https://rm.coe.int/third-report-on-bosnia-and-herzegovina/16808b5602>

European Commission, Bosnia and Herzegovina Analytical Report, <https://ec.europa.eu/neighbourhood-enlargement/system/files/2019-05/20190529-bosnia-and-herzegovina-analytical-report.pdf>

European Commission. (2021). Bosnia and Herzegovina Progress Report 2021. https://ec.europa.eu/neighbourhood-enlargement/bosnia-and-herzegovinareport-2021_en

Institution of Human Rights Ombudsman BiH. (2021a). *Godišnji izvještaj o rezultatima aktivnosti Institucije ombudsmena za ljudska prava Bosne i Hercegovina za 2020. godinu.* https://www.ombudsmen.gov.ba/documents/obmudsmen_doc2021030808580995bos.pdf

Institution of Human Rights Ombudsman BiH. (2021b). *Specijalni izvještaj o govoru mržnje u Bosni i Hercegovini.* https://www.ombudsmen.gov.ba/documents/obmudsmen_doc2021111511252845bos.pdf

Institution of Human Rights Ombudsman BiH. (2021c). *Specijalni izvještaj o govoru mržnje u Bosni i Hercegovini.* https://www.ombudsmen.gov.ba/documents/obmudsmen_doc2021111511252845bos.pdf

Institution of Human Rights Ombudsman BiH. (2020). *Godišnji izvještaj o rezultatima aktivnosti Institucije ombudsmena za ljudska prava Bosne i Hercegovina za 2019. godinu.* https://www.ombudsmen.gov.ba/documents/obmudsmen_doc2020051813152592bos.pdf

Communications Regulatory Agency. (2019a). *Izvještaj Regulatorne agencije za komunikacije o izrečenim izvršnim mjerama iz oblasti emitovanja u 2018. godini.* <https://docs.rak.ba/documents/c9e72fd0-7902-45c6-b458-85dd9d5a03f6.pdf>

CRA. (2020a). *Izvještaj Regulatorne agencije za komunikacije o izrečenim izvršnim mjerama iz oblasti emitovanja u 2019. godini.* <https://docs.rak.ba/documents/c9e72fd0-7902-45c6-b458-85dd9d5a03f6.pdf>

CRA. (2020b). *Pregled slučajeva iz oblasti emitovanja.* Januar-April. 2020. <https://docs.rak.ba/documents/168aa74f-bbb4-4a88-bd0a-5ffc344f12cf.pdf>

CRA. (2021a). *Izvještaj o izrečenim izvršnim mjerama Regulatorne agencije za komunikacije za 2020.* <https://docs.rak.ba/documents/29674686-be2c-4f40-aeaa51536a5684b0.pdf>

CRA. (2021b). *Godišnje izvješće Regulatorne agencije za komunikacije.* <https://docs.rak.ba/documents/0ecfac52-b3e7-4d1e-a080-df83226ee7d9.pdf>

CRA. (2021c). *Pregled slučajeva iz oblasti emitovanja. Periodična publikacija.* <https://docs.rak.ba/documents/5b09c9f8-f218-4c7a-8806-4a58716f461a.pdf>

Press Council BiH. (2020). *Pregled žalbi građana na uredničke sadržaje i komentare posjetilaca Internet portala 2020.* https://www.vzs.ba/images/2020/BHS_2020.pdf

Coalition Pod lupom. (2021). *Lokalni izbori u Bosni i Hercegovini. Finalno izvješće o građanskom, nestranačkom promatranju lokalnih izbora.* <https://podlupom.org/v2/bs/dokument/lokalni-izbori-2020-finalno-izvjesce-o-gradanskom-nestranackom-promatranju-lokalnih-izbora/39>

LAWS AND CODES

Action Plan for a Safer Internet 1999-2004, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM%3A124190>

Explanatory Memorandum, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020PC0825&from=en>

European Democracy Action Plan. https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2250

EU directive 2018/1808 <https://eur-lex.europa.eu/eli/dir/2018/1808/oj>

Election Law of Bosnia and Herzegovina, https://izbori.ba/Documents/documents/ZAKONI/Izborni_zakon_PRECISCENI_TEKST-bos.pdf

Code of Audio-visual Media Services and Radio Media Services, <https://docs.rak.ba/articles/333eb24f-ca18-4ef2-a9ab-8f402e8a4f40.pdf>

Code on Commercial Communications, <https://www.rak.ba/hr/articles/107>

Code of Press and Online Media of BiH, Press Council in Bosnia and Herzegovina, https://www.vzs.ba/index.php?option=com_content&view=article&id=218&Itemid=9

Criminal Code of Bosnia and Herzegovina, http://www.sudbih.gov.ba/files/docs/zakoni/ba/krivicni_zakon_3_03_-_bos.pdf

Criminal Code of the Federation of Bosnia and Herzegovina, http://www.oss.ba/dokumenti/Krivicni_zakon_FBIH.pdf

Criminal Code of Republika Srpska, https://www.paragraf.ba/propisi/republika-srpska/krivicni_zakon-republike-srpske.html

Criminal Code of Brčko District of Bosnia and Herzegovina, <https://skupstinabd.ba/3-zakon/ba/Krivic--ni%20zakon%20Brc--ko%20Distrikta%20BiH/05B19-20%20Krivic--ni%20zakon%20-precisceni%20tekst.pdf>

Law on Prohibition of Discrimination, <http://www.mhrr.gov.ba/PDF/LjudskaPrava/Zakon%20o%20zabrani%20diskriminacije.pdf>

Law on Protection against Defamation of the Federation of Bosnia and Herzegovina, <http://www.fbihvlada.gov.ba/bosanski/zakoni/2003/zakoni/13bos.htm>

Law on Protection against Defamation of Republika Srpska, https://www.parlament.ba/data/dokumenti/press-kutak/zakon_o_zastiti_od_klevete_republike_srpske_hr.pdf

Law on Protection against Defamation of Brčko District of Bosnia and Herzegovina, <https://skupstinabd.ba/3-zakon/ba/Zakon%20o%20zas--titi%20od%20klevete/000%2014-03%20Zakon%20o%20zas--titi%20od%20klevete.pdf>

Law on Public Order and Peace of RS, [https://mup.vladars.net/zakoni/rs_cir/ZAKON%20O%20JAVNOM%20REDU%20I%20MIRU\(Službeni%20glasnik%20RS%20broj%2011.15\).pdf](https://mup.vladars.net/zakoni/rs_cir/ZAKON%20O%20JAVNOM%20REDU%20I%20MIRU(Službeni%20glasnik%20RS%20broj%2011.15).pdf)

Law on Public Order and Peace of the Tuzla Canton, http://vladatk.kim.ba/Vlada/Dokumenti/Zakoni/2008/zakon_o_javnom_redu_i_miru.pdf

Law on Public Order and Peace of Brčko District, <https://skupstinabd.ba/3-zakon/ba/Zakon%20o%20javnom%20redu%20i%20miru/000%2032-09%20Zakon%20o%20javnom%20redu%20i%20miru.pdf>

Law on Public Order and Peace of the Herzegovina-Neretva Canton, <https://advokatbeganovic.ba/zakon-o-javnom-redu-i-miru-hercegovacko-neretvanskog-kantona/>

Law on Public Information of the Tuzla Canton, http://www.tk.kim.ba/Vlada/Dokumenti/Zakoni/2008/zakon_o_javnom_informisanju.pdf

Law on Public Order and Peace of the Una-Sana Canton, https://www.sluzbeniglasnikusk.com/files/sl_glasnik_2002/sl_gl_8.pdf

Law on Gender Equality in Bosnia and Herzegovina, https://arsbih.gov.ba/wp-content/uploads/2014/01/ZoRS_32_10_B.pdf

Law on Freedom of Religion and Legal Status of Churches and Religious Communities in Bosnia and Herzegovina, <http://www.mpr.gov.ba/biblioteka/zakoni/bs/ZAKON%20o%20slobodi%20vjere.pdf>

Decision of the High Representative Enacting the Law on Amendments to the Criminal Code of Bosnia and Herzegovina, <http://www.ohr.int/odluka-visokog-predstavnik-kojom-se-donosi-zakon-o-dopuni-kaznenog-zakona-bosne-i-hercegovine/>

Law on Non-Application of the Decision of the High Representative Enacting the Law on Amendments to the Criminal Code of Bosnia and Herzegovina, <https://www.narodnaskupstinars.net/?q=la/akti/usvojeni-zakoni/zakon-o-neprijemljenju-odluke-visokog-predstavnik-kojom-se-donosi-zakon-o-dopuni-krivi%C4%8Dnog-zakona-bosne-i-hercegovine>

Law on Amendments to the RS Criminal Code, <https://www.narodnaskupstinars.net/?q=la/akti/usvojeni-zakoni/zakon-o-dopuni-krivi%C4%8Dnog-zakona-republike-srpske>

Government of RS, 2020a, Prohibition of causing panic and disorder, <https://mup.vladars.net/lat/index.php?vijest=23036&vrsta=novosti>

Government of RS, 2020b, "Usvojen Prijedlog Uredbe sa zakonskom snagom o prestanku važenja Uredbe sa zakonskom snagom o zabrani izazivanja panike i nereda za vrijeme vanrednog stanja", <https://mup.vladars.net/lat/index.php?vijest=23281&vrsta=novosti>

Law on the Protection of Private Data, <https://www.paragraf.ba/propisi/bih/zakon-o-zastiti-licnih-podataka.html>

CONVENTIONS AND RECOMMENDATIONS

Recommendation No. R (97) 20, https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=0900001680505d5b

Council of Europe Recommendation No. R (99) 15 on measures concerning media coverage of election campaigns <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680645b44>

Commission of the European Communities 1997, <https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:1996:0487:FIN:en:PDF>

Council of Europe. (1997). Recommendation No. R(97)20. <https://rm.coe.int/1680505d5b>

Council of Europe Convention on Cybercrime (EST No. 185). [CEETS 185 - Convention on Cybercrime \(coe.int\)](https://www.coe.int/t/t09/cybercrime/Convention_on_Cybercrime_(coe.int))

Additional Protocol to the Convention on cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems (ETS No. 189)

The International Convention on the Elimination of All Forms of Racial Discrimination, <https://www.ohchr.org/en/professionalinterest/pages/cerd.aspx>

The Rabat Plan of Action, <https://www.ohchr.org/en/issues/freedomofexpression/articles19-20/pages/index.aspx>

International Covenant on Civil Rights and Political Rights <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>

DECISIONS AND RECOMMENDATIONS

European Court of Human Rights. (2016). Application No. 48657/16, Smajić v. BiH <http://www.mhrr.gov.ba/PDF/UredPDF/default.aspx?id=8222&langTag=bs-BA>

Human Rights Ombudsman. (2019b). Recommendation No. P-281/19. https://www.mreza.mira.net/wp-content/uploads/2019/12/Preporuka_Ombudsmen.pdf

Court of BiH (2017). S1 3 K 020812 17 Kž - Mandić Milan et al. <http://www.sudbih.gov.ba/predmet/3589/show?fbclid=IwAR2jmxw0nsa3QYjjQBo6qhW3ynhXL5aXbinfXyZQFgBiHzUN2o0jLUmM4>

Court of BiH (2018). S1 3 Iž 030580 18 Iž. file:///C:/Users/New%20guest/Downloads/1540535630RJESENJE_30580_18_IZ.pdf

Constitutional Court. (2016). AP 3430/16, Midhat Velagić's appeal. <http://www.sluzbenilist.ba/page/akt/byjFCFaeogE=>

LIST OF INTERVIEWS:

Azra Maslo, Regulatory Agency for Communication

Bekir Halilović, Adopt Srebrenica

Bojan Perkov, Share Foundation

Bojana Kostić, Article 19

Dajana Čelebić, Association for Democratic Initiatives

Edin Ibrahimfendić, Ombudsmen BIH

Halisa Skopljak, OSCE mission

Maida Zagorac, Network for Building Peace.

Mirna Miljanović, Department for Public Relations, Ministry of Interior BiH

Mirjana Ćuskić, Helsinki Committee for Human Rights

Nikolija Bjelica, Direkt portal

Sunita Dautbegović – Bošnjaković, ZFD forum

Silvije Fućec, Agency for the Protection of Private Data

Stefan Terzić, Interreligious Council

Tijana Cvjetičanin, Raskrinkavanje

SOCIAL NETWORKS COMMUNITY GUIDELINES

[Facebook standardi zajednice | Centar za transparentnost \(fb.com\)](#)

[How technology detects violations | Transparency Center \(fb.com\)](#)

[Community Standards Enforcement | Transparency Center \(fb.com\)](#)

[Govor mržnje | Centar za transparentnost \(fb.com\)](#)

[Removing Holocaust Denial Content | Meta \(fb.com\)](#)

[Lažne vijesti | Centar za transparentnost \(fb.com\)](#)

[Manipulirani medijski sadržaji | Centar za transparentnost \(fb.com\)](#)

[How Facebook's third-party fact-checking program works](#)

[Oversight Board | Independent Judgment, Transparency, Legitimacy,](#)

[Decision | Oversight Board](#)

[Introducing Journalist Registration on Facebook](#)

[Facebook Protect](#)

[YouTube Community Guidelines & Policies - How YouTube Works](#)

[YouTube Trusted Flagger program - YouTube Help \(google.com\)](#)

[YouTube Progress on Responsibility - How YouTube Works](#)

[YouTube Community Guidelines enforcement \[https://transparencyreport.google.com/youtube_policy/removals?hl=en\]\(https://transparencyreport.google.com/youtube_policy/removals?hl=en\)](#)

[Government requests to remove content - Google Transparency Report](#)

[YouTube Monetization Policies - How YouTube Works](#)

[Appeal an account suspension or locked account | Help Center \(twitter.com\) Expanding our private information policy to include media \(twitter.com\) Twitter's enforcement philosophy & approach to policy development](#)

[About Twitter | Trust and Safety Council Removal Requests - Twitter Transparency Center](#)
[TikTok](#)

TRANSPARENCY REPORTS:

<https://transparency.twitter.com/en/reports/removal-requests.html#2021-jan-jun>

https://transparencyreport.google.com/government-removals/overview?hl=en_GB

<https://transparency.fb.com/data/government-data-requests/>

<https://www.tiktok.com/transparency/en-ca/government-removal-requests-2020-1/>

