Content Moderation and Local Stakeholders in Bosnia and Herzegovina

June 2022

SOCIAL
MEDIA
4PEACE

ARTICLE 19

T: +44 20 7324 2500

F: +44 20 7490 0566

E: info@article19.org

W: www.article19.org

Tw: @article19org

Fb: facebook.com/article19org

Contents

Executive summary

With a focus on Bosnia and Herzegovina (BiH) and as a contribution to the **Social Media 4 Peace** project, implemented by the United Nations Educational, Scientific and Cultural Organization (UNESCO) with partner ARTICLE 19, this report explores the local-specific contextual concerns stemming from global, non-transparent, and profit-driven content moderation processes of social media. The report analyses what happens when certain local communities and countries are 'invisible' to social media platforms and illustrates how cross-sectoral collaboration in the form of a coalition for freedom of expression and content moderation could help these communities engage with social media platforms and have a voice in content moderation cases that impact their society.

Through exploratory desk research and qualitative interviews, the report identifies trends in content moderation dynamics, as well as their harmful effects on political, societal and peace-building processes. These dynamics lead to a disconnect between the country's media landscape, civic participation and social media companies. This disconnect (referred to as 'terra nullius') prevents effective responses to the emergence of local contextual issues like information manipulation and proliferation of ethnic, gender, and other cross-cutting forms of hatred. In this environment, media outlets, civil society organisations, and the advocacy community are exposed to pressure and abuse, including character assassination, orchestrated harassment, and penalties from false copyright claims and false reporting. Some stakeholders have devised methods to successfully tackle these problems (referred to as 'guerrilla tactics'), although these tactics can be little more than a reactionary plaster to what amounts to systemic problems.

There is a concerning lack of meaningful response, as well as of public and expert debate, on solutioning the numerous problems arising from content moderation processes in the country. Content moderation is predominantly perceived as an impenetrable system. Nonetheless, content moderation is a transversal issue that falls under the scope of different state and non-state actors, private businesses, civil society, and social media companies. A list of stakeholders and an assessment of their work and needs, developed during the research process, fails to indicate a single civil society organisation that

includes Internet freedom issues in its portfolio. Additionally, there are no in-state monitoring or reporting programmes, policy-focused advocacy initiatives, or digital and open-source communities. This situation gives rise to another obstacle, namely the lack of expertise when it comes to Internet freedom and digital rights, and the absence of stakeholder proactive engagement with harms arising from current content moderation practices. Existing expertise is mostly found within academia, civil society and media organisations, and among the activist community.

Given this background, a coalition for freedom of expression and content moderation, and the general concept of a multi-stakeholder platform, emerges as a necessary mechanism for closing problematic loopholes, to support the establishment of channels of communication and cooperation with social media companies, and to address those issues that threaten freedom of expression online, media freedom, societal cohesion, and peace.

Introduction

This publication has been produced as part of the United Nations Educational, Scientific and Cultural Organization's (UNESCO) project **Social Media 4 Peace** funded by the European Union (EU).

About the project

This report is part of the **Social Media 4 Peace** project that UNESCO is implementing in Bosnia and Herzegovina (BiH), Kenya, and Indonesia with support from the European Union (EU). The overall objective of the project 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 contributing to the promotion of peace through digital technologies, notably social media. ARTICLE 19's contribution to the project focuses on concerns raised by the current practices of content moderation on dominant social media platforms in the three target countries.

In addition to the three country reports elaborated with external research consultants, ARTICLE 19 also prepared a summary report for the **Social Media 4 Peace** project: this report compares the learnings and recommendations.

ARTICLE 19 considers that social media companies are, in principle, free to restrict content on the basis of freedom of contract, but that they should nonetheless respect human rights, including the rights to freedom of expression, privacy, and due process. While social media platforms have provided opportunities for expression, a number of serious concerns have come to light. The application of community standards has led to the silencing of minority voices. The efforts of tech companies to deal with problematic content are far from being evenly distributed: for instance, it has been shown that 87% of Facebook's spending on misinformation goes to English-language content, despite the fact that only 9% of its users are English speaking. It has also been revealed that most resources and means in terms of content moderation are being allocated to a limited number of countries. Generally speaking, the transparency and dispute resolutions over content removals have so far been inadequate to enable sufficient scrutiny of social

media platforms' actions and provide meaningful redress for their users. Finally, it is doubtful that a small number of dominant platforms should be allowed to hold so much power over what people are allowed to see without more direct public accountability.

This report specifically looks at the situation of local actors who, while they are impacted by the circulation of harmful content on social media or the moderation thereof, often find themselves unable to take effective action to improve their situation in that respect. They may feel frustrated by the inconsistencies of platforms' application of their own content rules; they may feel that global companies ignore their requests or misunderstand the current circumstances of the country or region. Some may lack understanding of content rules or of content moderation, but that is not the case of all local stakeholders.

The research then seeks to test, through the views of local stakeholders, the assumption that a local Coalition on Freedom of Expression and Content Moderation could play a role to fill the gap between the realities of local actors and companies that operate on a global scale. The idea for such a coalition is based on ARTICLE 19's work on the development of Social Media Councils, a multi-stakeholder mechanism for the oversight of content moderation on social media platforms. ARTICLE 19 suggested that Social Media Councils should be created at a national level (unless there was a risk that it would be easily captured by the government or other powerful interests) because this would ensure the involvement of local decision-makers who are well-informed of the local context and understand its cultural, linguistic, historical, political, and social nuances. While the development of a self-regulatory, multi-stakeholder body such as a Social Media Council is a long-term and complex endeavour, a local Coalition on Freedom of Expression and Content Moderation would be a lighter approach that could be supported within a shorter timeframe. Basing its work on international standards on freedom of expression and other fundamental rights, such a coalition could provide valuable input to inform content moderation practices, notably through its knowledge and understanding of the local languages and circumstances. As a critical mass of local stakeholders, it could engage into a sustainable dialogue with social media platforms and contribute to addressing flaws in content moderation and improving the protection of fundamental rights online. The coalition could provide training and support on freedom of expression and content

moderation to local civil society actors that are impacted by content moderation. Finally, it could possibly pave the way to the creation of a Social Media Council in the country at a later stage. Through this research, at the initial stage of the **Social Media 4 Peace** project, the idea of a local Coalition on Freedom of Expression and Content Moderation was submitted to local stakeholders, whose views have enabled the formulation of recommendations on how to approach the facilitation of a pilot coalition in the specific context of Bosnia and Herzegovina. In order to guarantee the effective ownership of the coalition by its members, the process facilitating its creation will necessarily include a validation exercise that ensures that potential members have the opportunity to discuss the findings of the research.

To that end, and while focusing on the local voices in BiH, this report examines local-specific content moderation issues and the position, knowledge, and needs of various state and non-state actors (see Annex A for more details). It outlines the current state of social media to identify key content moderation issues and their effects on the country's political, societal, and cultural processes, including topical case studies (see The state of content moderation in BiH). Each stakeholder interviewed in this report is categorised by their affiliation and scope of work to assess their position, knowledge, and needs in addition to their capacity to engage with coalitions, advocacy, and other activities in relation to freedom of expression and content moderation (see Analysis of the stakeholders). To ensure that the **Social Media 4 Peace** project is able to yield foreseen results, this report concludes with a workplan for a local freedom of expression and content moderation coalition, and an assessment of the risks and obstacles future leadership and members may face (see Recommendations).

For the purposes of this report, we rely on the following definitions:

- **Content moderation** includes the different sets of measures and tools that social media platforms use to deal with illegal content and enforce their community standards against user-generated content on their service. This generally involves flagging by users, trusted flaggers or 'filters', removal, labelling, down-ranking or demonetisation of content, or disabling certain features.

- **Content curation** is how social media platforms use automated systems to rank, promote, or demote content in newsfeeds, usually based on their users' profiles. Content can also be promoted on platforms in exchange for payment. Platforms can also curate content by using interstitials to warn users against sensitive content or applying certain labels to highlight, for instance, whether the content comes from a trusted source.

Methodology

This report presents the findings of policy and literature review, which provide insight into the BiH freedom of expression and media freedom landscape and key trends therein. The report combines two methodological components: explorative desk research and qualitative interviews with stakeholders. The collection and analysis of (scarce) BiH-specific research and international and regional resources as the key focus of desk research helped establish a baseline understanding from which to tease out significant local narratives. The initial desk research revealed a knowledge gap and overall lack of meaningful engagement in BiH with Internet freedom related topics. To address this gap, qualitative interviews and field research were carried out in September 2021. To ensure that key actors were included, this research relied on purposive sampling methodology. In total, 23 interviews and focus groups were organised, with interviews lasting between 60 – 180 minutes, and utilising semi-structured questions. The field research and qualitative interviews proved essential elements in this research process. The researcher would like to express her gratitude to all who shared their knowledge and took part in the research process (see Annex C for a detailed list of respondents).

Bosnia and Herzegovina at a glance

To understand the dynamic of content moderation and freedom of expression in BiH more broadly, one needs to first comprehend the country's complex political and governing system. The current governing structure of BiH was codified in the Dayton Peace Agreement in 1995 during the peace negotiation processes in the aftermath of the armed conflict in the former Yugoslavia[1] during which more than 100,000 people died and two million people – 'more than half the population' – fled.

Presently, BiH is a state consisting of two entities (the **Federation of Bosnia and Herzegovina** and the **Republika Srpska**) and a semi-autonomous **Brčko District**. Furthermore, the Federation of BiH consists of 10 cantons. As a parliamentary democracy, BiH's executive power is exercised through a Presidency and Council of Ministries, and its legislative power is vested in the Parliamentary Assembly. The Presidency serves as Head of State and is held by the representatives of three constituencies of BiH (Bosniaks, Croats, and Serbs). The country's decentralised political system and deconcentration of power is further replicated in other structures, thus the two-state entities and cantons enjoy a high level of autonomy and have independent governing structures. This decentralised model provides ethnic group representatives within BiH's institutions the ability to block decisions and legislative proposals made by representatives of other ethnic groups – resulting in a political dead-lock and instability. According to the report respondents,[2] this political model has given birth to three parallel political (party) structures, each of which is divided on the basis of ethnicity.[3] This tripartite and multi-polarised political and societal system and intersected interests are clearly visibly embedded within the digital sphere (see The state of content moderation in BiH for an extensive discussion of this).

The country signed the Stabilization and Association Agreement (SAA) with the EU in 2015 as the country's first step to becoming a member of the EU. In 2021, the European Commission, in their Annual Report – which includes an overview of trends related to and progress of structural reforms – stated that BiH's 'European integration has not been turned into concrete action, as political leaders continued to engage in divisive rhetoric and unconstructive political disputes […].' In practice, this means that, for part of 2019, work in the BiH Parliament was halted by members of the Parliament from Republic Srpska who failed to recognise state institutions and refused to participate in its work. In 2021, including also 2022, the country's key institutions were blocked once again following a decision issued by the Office of the High Representative (an international body that ensures the implementation of the Dayton Agreement) on amendment to the Criminal Code of Bosnia and Herzegovina.[4] After refusing to acknowledge this decision, a representative of Republika Srpska and member of the Presidency decided to invalidate

state institutions. In response, state officials from the same political party rejected the work of Parliament, thereby obstructing legislative processes. In addition to this political dead-lock, there are upcoming National Parliamentary and Presidential elections that will be held in October 2022. Leaders from the ruling political party in Republika Srpska have indicated that they might abstain from these elections if the electoral system has not been reformed by then.

This politically unstable political system, fraught with ethnic tension, has created structural loopholes, inefficient rule of law that perpetuates ethnic divisions, and erodes societal cohesion, tolerance, and, ultimately, peace. In a system such as this, the human rights legislative framework and institutional infrastructure – the Ministry for Human Rights and Migration, and the Ombudsperson at state level – have limited capacity to effect action and change. In addition, numerous human rights and peace-building organisations, international organisations – primarily the Office of the High Representative, but also the field offices of the Organization for Security and Co-operation in Europe, Council of Europe, United Nations System Agencies, Funds and Programmes, and the EU delegation are often involved in key structural state media reforms and cooperate with the a range of actors in the country, including media organisations and civil society. As noted by the 2021 European Commission report, BiH must 'ensure full respect, protection and promotion of freedom of assembly and expression, and refrain from further actions that adversely impact the exercise of these rights.' There is no media legislation related to issues including media pluralism and ownership or advertising and state financing, and existing laws are not sufficiently aligned with the European and international standards, nor are they enforced. This also holds true for hate speech legislation, which still requires harmonisation with EU standards. The insufficient effectiveness of the judiciary system further deters citizens to pursue these cases. When it comes to online platforms, the European Commission report notes that they are used to spread hate speech and disinformation while there is limited judiciary response to address religious and ethnic hatred spreading through Internet and social media. However, some positive steps have recently been taken. According to LGBTQI+ (lesbian, gay, bisexual, transgender, queer, and intersex) and feminist activist Lejla Huremović, there was less

hateful engagement in response to this year's Pride protest compared to previous years, given that Sarajevo canton's law enforcement department showed more willingness to cooperate and help with policing on and offline hate speech.[5]

There are no state-wide laws directly regulating Internet and online communication generally. There have, however, been several political and legislative efforts to curb Internet freedom. In 2015, the Republika Srpska amended its Law on Public Order and Peace to include a definition of public space that encompasses the Internet and social media networks. In September 2021, a similar legislative initiative was proposed by the BiH Parliament and, if adopted, it would task law enforcement to mandate public peace and order online. How this will be practically enforced remains to be seen. This type of intrusive regulation, which is happening within a politicised public space, and the lack of critical voices and active citizenship takes place against a background that has been well characterised as follows:

> "*There is an intention to strengthen the entity levels so that they reach a state structure. This would in turn centralise their powers and weaken the actual state-wide level, so as to keep it as a shell authority. In that case, the real power would stay within the entities, and they would like to see this mirrored on the Internet and social media, so we end up with three Internets and Facebooks. But the social media platforms, due to their infrastructure, are incompatible with this intention, and that is why they are prone to manipulation and misuse.*"[6]

The need for structural and systemic reforms necessary for economic prosperity has only increased with the severe recession caused by the Covid-19 pandemic: the job market has stagnated with unemployment rates soaring to approximately 30%[7] and up to 60% for younger populations.[8] Average salaries are low and 16% of the population live in poverty. When it comes to ecommerce, the market is underdeveloped (online fashion retailers hold the largest share), with only 15% of the population participating (only 9.7% have a credit card). Contrarily, Internet infrastructure within the country is relatively evenly distributed: of 3.29 million citizens, more than 3 million individuals use the Internet, and, of these, 1.7 million are active on social media.

It is important to read this report through the lens of deeply rooted structural and systemic ethnic division, situated within a fractured political environment, weak rule of law, judiciary system and economy. Taken together, these socio-political and technological factors shape the digital (media) landscape. This environment is exacerbated by three key political and ethnic power structures that seek to enforce a self-serving and provincial governing model, designed to gain and hoard power. The impact on and reflection of this system within public discourse is most accurately explained by the director of CIN, an investigative media outlet:

> "*These three (political) parties and their ethnic institutional structures are able to control the public narratives by 'trading in fear', as relevant public issues (e.g. poverty, poor living conditions, social justice problems) are ignored or used as political kindling – inciting historical ethnic tensions and inflaming divisive rhetoric. These tensions are reported by the media and social media, and thus continue to dissolve these public interest topics.*"[9]

Or, in the words of a director of a long-established media outlet: "This deflection tactic – that's all we live in – unfinished war, frozen space – it's the only constant in the media and social media space."[10] The key ethnic-political parties have well-utilised this deflection tactic but, more importantly, they have cemented fear, apathy, and ethnic division in the conversations, framework, and general tapestry of the country. This strategy, carried out by BiH political elites, has created a public space prone to manipulation, as analysed in the following chapter.

The state of content moderation in BiH

Social media landscape in BiH

With 67 Internet providers and 87.25% of the population using mobile Internet access, the number of Internet users in BiH (3,336,591) is nearly the same as the number of the people living in the country (3,531,159).[11] BiH's high Internet penetration is partly a result of widely accessible Internet infrastructure and competitive market, ensuring access is affordable. In April 2020, the 'roam like at home' programme was introduced, which removed high roaming prices across and within the Western Balkans, removing additional obstacles impeding Internet penetration.[12] Despite the country's geographical and ethnic diversity, there is no significant digital divide, including between urban and rural communities in the country.

According to the Agency for Statistics, almost two-thirds of households (72.8%), and more than 74% of individuals have access to and use information-communication technologies. Citizens mostly use the Internet for communication purposes (e.g. phone calls, 90.6%), chatting over private messaging apps (78.0%), accessing social media platforms (68.3%), and reading online Internet portals and news (65.6%).[13] Although there is no official data about social media usage, the widely quoted data aggregator, Data Portal, reported 1.70 million social media users in BiH in 2020. More than 50% of BiH's population have a social media account, and the vast majority access social media via mobile phone (97%).[14] Of those using social media platforms, 1.40 million have Facebook accounts, which has been cited as the most visited and utilised platform, with Instagram (860,000) the second most used, and Twitter (166,500) the third. These data are in line with 2021 Communication Regulatory Authority research on media habits in the country which found, relying on a representative sample, that Facebook is the most popular social network (72.5%), followed by Instagram (39.4%), and YouTube (38.1%), and that all the other social networks are used by less than 10% of population (Twitter 3.5% and TikTok 8.8%). Facebook and Instagram are the main social networks and are used by all ages, though Instagram is used more among the younger generation (18–24 years old) as are TikTok and Twitter. Given that Meta owns Instagram, it is clear that nearly the entire social media market is in

the hands of one company. As noted by the media editors interviewed for this report,[15] Facebook's dominance is a problem for both the media and citizens.

In fact, the top search term entered in Google browsers in 2020 was 'Facebook', followed by 'weather forecast', 'Klix.ba' (one of the most popular online news portal), 'YouTube', and 'Avaz' (a local news outlet). YouTube is used by a relatively large number of people – mainly for entertainment and fun – to watch local reality-shows and popular music competitions (Zadruga, Zvezde Granda, and similar 'low quality' content that promotes gender-based violence and ethnic hatred).[16] A 2019 survey on media habits and the perception of media, conducted in the country's three largest cities (Banja Luka, Sarajevo, and Mostar), also indicated a similar breakdown of Internet usage. The study also found that film, music, and other entertainment-oriented content on social media draws the biggest audience (46.7%), followed by sporting events and related news (38%). The same study indicated that almost 95% of those surveyed stated that they were either constantly online or that they use the Internet more than once a day. To access news content, those surveyed use online portals (42%), social media networks (30%), and television (11%). The older audiences (above 60 years old) rely more on television, radio, and print media, whereas younger audiences (18–40 years old) primarily use social media networks as their main news resource. Strikingly, the audience aged between 18 and 28 do not read print newspapers at all. A more recent study from 2020, using representative sampling, showed that television is now the most-used news source (71%) followed by social media platforms (50%) and online news portals (33%). This state-wide survey indicated a significant generational gap: 88% of those surveyed aged 60 and above access news content via television compared to 45% of those aged 18–29. These results were reproduced by a 2020 UNESCO and Communication Regulatory Authority survey that looked at informational awareness during the Covid-19 pandemic, and confirmed television as the primary source of information, followed by online news portals and social media platforms.[17] As is the case in most parts of the world, in times of crisis, like during the latest Covid-19 pandemic, television broadcast content is the primary news resource in BiH. However, in BiH, trust in media is a complicated issue, with only half of the population

espousing trust in news media (50%), broadcast media (43%), and even less in social media (33%).

Despite the low level of trust in media, surveyed citizens of Sarajevo, Banja Luka, and Mostar mostly agree (77%) that the media is capable of inflaming national tensions and inciting conflict in the country, and that there should be more positive content that promotes public values and society cohesion through individual and public authority stories and narratives (95.3%). The fact that the media is seen as a catalyst of ethnic tensions is significant, and must be viewed through the lens of political segregation, and there is strong political influence over media outlets and the media economy, generally.[18] Social media networks, and Facebook in particular, are not exempt from this influence. This is also true for YouTube, a platform that has become a 'super-spreader' of disinformation during the pandemic (see Case study 2). As far as Twitter and TikTok are concerned, but also YouTube outside the context of disinformation, there is almost no research on usage. For example, the survey respondents mentioned that Twitter is mainly used as a 'closed group'[19] – people from similar communities or with shared opinions and perspectives.

On a macro-level, the analysed data and surveys reveal that people in BiH use information-communication technologies to communicate, visit social media profiles, or scroll through their social media accounts to look for entertaining programmes and other personal content shared and produced via social media. As a result, online public deliberation is relatively low, even though social media platforms – predominantly Facebook given its dominance in the country – are the most used fora for public discussion. This has led one respondent to conclude that: "For many people, the Internet is Facebook, but people usually do not see Facebook as a powerful company, for them it is just a platform."[20] As much as social media is an important source of news, the Covid-19 pandemic has altered the media habits in this country, propelling television into first place as the primary and most trusted source of information. However, media habits are constantly changing, and it is hard to estimate for how long television will remain the primary source of information. The fact that young people, but also younger to middle-aged adults, are increasingly accessing news content and other information via social media algorithmically selected

content calls for a greater attention and better understanding of the impact that these platforms and their content moderation policies and practices have on the forming of public opinion.

Overview: Impact of content moderation on peace and stability

In BiH, perhaps more than in other countries, ethnic division, fragmented audiences, and frictions create fertile ground for different forms of manipulation and lead to an uneven distribution of communication power, thus strengthening the positions of those already dominating public discourse. However, there is a dearth of substantive and holistic studies on the problems related to content moderation,[21] which creates a serious obstacle for a meaningful intervention. This significant obstacle comes as a result of a lack of sustainable funding and provision of resources for meaningful engagement with those problems inherent in the operations of social media platforms. Indeed, project-based engagement has created a plethora of knowledge and expertise loopholes (a lack of consistent social media monitoring practices, data, tools, and archive practices, including in-depth analysis or content moderators in the newsrooms, etc (see Analysis of the stakeholders for an extensive discussion). In an effort to fill these loopholes, this report will attempt to synthesise existing resources and will rely on interviews for a more complete understanding of the content moderation dynamic in the country. As a starting point, and recalling the fragmented socio-political system, a particular paradox that characterises the social media landscape requires unpacking.

The decentralised political system has the potential to create space for a vibrant and participatory public debate – a cornerstone of democratic societies. Paradoxically, despite the high level of decentralisation of political power in BiH, ethno-political forces have still managed to fracture and atomise public debate both off- and online. The 'atomized identities' and 'dispersive public discussion online',[22] coupled with low levels of media and information literacy, have further suppressed public debate, and silenced those voices critical of political elites.[23] Within this environment, a number of local-sensitive contextual issues have emerged as a spectrum of inter-connected and dependent concerns: inter-ethnic manipulation and hatred; targeted assaults on under-represented and oppressed

groups; media-related obstacles related to news dissemination on social media; as well as vital social media initiatives that foster diversity and peace.

In relation to the removal of different forms of identified harmful content, state institutions and judiciary provide little to no help. After filing a complaint, the investigation and prosecution can take a long time to make any meaningful progress, and there is little experience when it comes to litigating these cases.[24] The Journalists' Association has had some success – out of nine litigated cases on the violence against journalists, six concerned online harassment. Another worrying trend is emerging that could result in greater capacity of law enforcement and the judiciary to police speech online. Amendments to the Law on Public Order and Peace that seek to expand the scope of public space to include Internet and social media are currently being debated. This legislative proposal merely mimics the provision of the existing law in Republika Srpska, adopted in 2015, which has been described by critics as an 'attack on free speech' and more strictly enforced during the Covid-19 pandemic. Under the new draft and existing law, anyone who publishes and disseminates content, including user-generated content on the Internet and social media deemed a disturbance to public order, indecent, offensive or disturbing, or with intent to insult, or engage in rude or insolent behaviour, can be fined.[25] In an analysis published on the citizen's portal Javna rasprava ('Public discussion'), the following concern was raised:

> *Imprecise and open-ended provisions of the law can, in practice, be misused and deter citizens from receiving and imparting information on the Internet for fear of possible consequences, even when the content they spread is not hate speech. Therefore, it is imperative that the provisions be precise and clear, in order to prevent arbitrary interpretation and abuse in practice.[26]*

To fill the problematic gaps, and respond to these concerning and complex issues, media outlets, civil society organisations, and activists have made use of a variety of different methods in order to reach social media platforms. Even for those who have been able to establish human contact with them, this contact does not necessarily result in the resolution of content disputes or a remediation of process failures (see more on this in the

following sub-section).[27] The seeming ambivalence of platforms, and clear lack of willingness to address content moderation issues in BiH, renders them 'enablers' – unaccountable – of hatred and discriminatory speech and obstacles that prevent the full public from accessing content of public interest and exercising freedom of expression online. From the perspective of the social media companies, given the lack of safeguards preventing the monopolising of power, BiH can even be considered as a perfect 'test zone'[28] where they can – in real time – introduce new and change old content moderation policies and practices (or not, as the case may be) under a veil of impunity. But this dangerous disconnect means that fraught country dynamics and noted local-sensitive contextual issues place severe limitations on freedom of expression and media freedom and can result in severe consequences for the society of BiH as a whole.

Applicability and relevance of content moderation rules/regimes

Underpinning local-sensitive contextual issues and related content moderation concerns are a plethora of [intersectional pluralistic interests](#) that shape individual experience and the ability to express, impart, and receive information online. Interview respondents talked extensively about the noticeable effects of a lack of content moderation of political and malicious inflammatory speech, and its negative effect on public debate (see 'Content moderation: local implications of global rules'). However, this is not to say that content moderation, as such, can be seen as a 'value-neutral' exercise. On the contrary, it is governed by powerful social media companies with profit-driven policies and decisions that impact not only the digital ecosystem, but also our 'analogue' society and values. Thus, the logical question arises: how do social media companies accommodate and respond to existing content moderation challenges in BiH? This question is discussed in more detail in the following paragraphs, but the answer seems to be rooted in the disconnect between the social media companies and the communities, media organisations, and individuals that utilise them ('on the ground') and, simply, that many concerns remain unaddressed in this 'terra nullius'. This concerning disconnect is explained by the interview respondents: "Social media companies do not even know that we exist, that their policies are affecting us."[29] Another respondent noted, "We are on the

margins of the global social media processes, and we are not recognized as important actors."[30]

In practice, this disconnect is also reinforced by the fact that social media companies do not have representatives in the country. Except for the debunking organisation Raskrinkavanje, an official third-party checker for Facebook, there is little to no information about their relationship to country stakeholders and whether media organisations are involved in the journalism initiatives of social media (e.g. Facebook Trust Program or Google News Initiative). Existing internal documents (e.g. community standards, terms of service) – key contractual documents between the individual and platform – are not available in their entirety in local languages, and the parts that are available are poorly translated and they are not necessarily framed to respond to the identified local contextual issues. However, given the broad and open-ended nature of these documents, a range of provisions could be potentially applicable in addressing some of the challenges like gendered or ethnic hate speech.[31] The same is true for community standards addressing serious breaches such as, for example, the promotion of violence or criminal behaviour, including orchestrated and organised assaults, threats, and slurs. From a normative perspective, these community standards and internal documents could potentially be applicable to respond to some local-specific content moderation issues. However, the lack of data and insight into the enforcement of these global standards, especially their ability to address local-context, makes it extremely difficult to assess their relevance and their effectiveness.

A survey of Western Balkans social media users, published by Balkan Investigative Reporting Network's (BIRN) Balkan Insight news portal, provides an insight into the social media platform response in the region. The study found that reported threats of violence and harassment in local languages were removed in 60% and 50%, respectively, of the cases in four Western Balkans countries, including BiH. However, up to half of reported hate speech remained online.[32] The report included input from Meta and Twitter representatives who stated that the platforms primarily use algorithms and human moderators to detect and remove hate speech, in addition to proactive removal measures. In a similar vein, when asked about the experience of reporting content to social media

platforms during the interview, one respondent said: "When I report content in English, the reaction of platforms is much faster and more accurate than when I report a problem in our language."[33] While this particular experience refers to Facebook, when asked about YouTube, the respondent said the platform is "even worse, since their reporting labels do not make sense, and there are no adequate or functional reporting mechanisms."[34] However, they also noted some progress regarding Facebook's hate speech recognition algorithm, and more consistent demotion or take downs of obvious hate speech and threats, but there is no official data to confirm this assumption.[35] This progress is often limited, given the linguistic and cultural proximity to Serbia and Croatia, because disinformation and narratives of hate populate social media spaces, adding to the existing internal layers of complexity.

The only publicly available BiH relevant information on social platforms are in relation to state requests to access certain (personal) information. For example, Meta's Transparency Center shows that from July to December 2020 they received 77 state requests (71 legal process requests and 6 emergency disclosure requests) and 112 user account requests from state authorities. The data from previous years show a steady increase in state requests. During the same period [Facebook restricted access to several pages, groups, and profiles](12 in total).[36] Similarly, [Twitter published a transparency report] that shows that only three 'legal demands' were issued by state authorities in 2020. This data lacks context – as much as in the context of content moderation – and as such cannot contribute to any substantive findings. The following case study offers some practical insights into the loopholes created in this disconnect between social media platforms and the country through the lens of tension between Raskrinkavanje and the Press Council.

Case study 1: Fact-checkers and the Press Council – same goal, different means

There is an increasing tension between media outlets and the official third-party fact-checking platform, Raskrinkavanje – part of the well-established civil society organisation, Zašto ne?. The main reason for this tension is the potential penalties outlets face when their content is flagged by Raskrinkavanje following a substantive analysis. These penalties can be a serious blow for an already economically vulnerable media outlet, and so the media organisations have reacted negatively although the importance of preventing disinformation from circulating in the online public discourse is clear. The Press Council, of which almost all media outlets are registered members, has emerged as a facilitator, seeking to find a feasible solution for media outlets' growing dissatisfaction.

The Raskrinkvanje platform (in local language, the name means 'Debunking') originated from the Istinomer project which assesses the statements of politicians compared to stated political agenda announced during election campaigns. Istinomer operated until 2017 when the extent of media-manipulated content necessitated the establishment of Raskrinkvanje to enforce media accountability for content creation. Soon after, Raskrinkvanje became an official member of the International Federation of Fact-Checkers Network, through the Poynter Institute in Florida. The latter established cooperation with Facebook which led to the official process of accreditation for fact-checkers. In practice, this means that only a small number of organisations that meet proposed standards (e.g. solid methodological framework that guarantees transparency, impartial processes, and objectivity) can become official fact-checking partners. In turn, Facebook is required to make certain parts of their platforms available to fact-checkers. As a result, Raskrinvakvanje is allowed to rate and flag content they assess as false or misleading media content on the platform. Once content has been flagged (followed by a further assessment), the system of sanctions is set in motion and Raskrinkavanje has no further control. Sanctions are algorithmically imposed and, in practice, only Facebook can intervene in this process. Sanctions vary from a decreased visibility of the particular content, to the decreased visibility of an entire page, or inability to sponsor content, if the number of negative ratings continues to increase over time. The subject of debunking is not only related to media content, but also to publicly available user-generated content reported by either individual users (as is usually the case), or Raskrinkvanje journalists. To

date, Raskrinkvanje has rated around 10,000 news articles and pictures, encompassing more than 2,000 disinformation sources.

In the words of the editor-in-chief, they believe that, through fact-checking, media outlets should be encouraged to respect professional standards and that the goal of this process is not to chill the media freedom, but to ensure that social media infrastructure is not an instrument for monetisation in the hands of media that produce disinformation and other manipulative content. However, the media negatively reacted to their factual assessment, and in recent months, a trend of violence against fact-checkers has also emerged, as two of the female journalists have faced serious threats coming from a journalist and a public figure. The editor is also painfully aware that the group of six journalists cannot make a meaningful change and that in many ways their efforts are only 'a drop in the ocean'.[37] But in the situation where there are no effective mechanisms to deal with such content – debunking remains the only viable option.

For years, the Press Council has worked to strengthen journalism and media independence. The executive director of the Press Council, which operates on a shoe-string budget, spoke about the vulnerabilities of their current position. The media are not satisfied with the rating and penalty system (a number of outlets are on the verge of economic ruin) and have no complaint mechanism method for challenging these decisions, and according to the Press Council director, some media have called the process non-transparent, arbitrary, and error-prone. They also point to a glaring double-standard: when they report hate speech found in the comment section to Facebook, the platform provides no support and the hate speech remains online, but for media outlets' 'missteps', there is a real price. It should be noted, however, that the Press Council is not involved in the fact-checking process, nor do they have any power to influence Raskrinkavanje as they are not members of the Council.

The fact-checker's work and resulting sanctions, put pressure on the media to produce news in line with professional standards of journalism. The Press Council, as a self-regulatory body, houses the Complaint Commission which also assesses and adjudicates the individual complaints against media content, including disinformation, by offering both parties a place at the table. Media content is assessed in accordance with the Journalism Ethical Code of Conduct – a set of standards that media outlets are required to adhere to

upon becoming members of the Press Council. Therefore, both organisations do assess and rate content using different mechanisms. However, for media outlets (members of the Press Council), this is the only policy against which their work and content should be assessed. In addition, media outlets do not oppose fact-checking, but object to their exclusion from, and lack of transparency of, processes that both fall outside the scope of the Ethical Code and are established by an external stakeholder. Essentially, this tension comes down to the question of the legitimacy of the Raskrinkavanje platform.

According to the Press Council, the work of Raskrinkavanje has limited the impact of the Complaints Commission. According to a project coordinator at the Press Council,[38] fact-checking work is unidirectional – from the fact-checker to the media outlet, mimicking the relationship between the media and social media platforms in the country. For some, however, the Complaints Commission is toothless and unable to contribute to real change. It is also true that fact-checking has uncovered numerous sources of disinformation and their harmful effects on media freedom and even public health (see Case study 2).

The goals of the Press Council and Raskrinkavanje are the same: sustainability and strength of independent journalism. There is some sort of jurisdictional conflict between the organisations that both employ mechanisms of assessment and have mandates to assess and ensure adherence to professional standards of journalism. Any resolution of this tension genuinely involves the role of Facebook.

The absence of social media companies – when it comes to representation and interest in the region – is clearly seen via their discriminatory and non-uniform[39] monitoring response to local content moderation problems, and the differential treatment of certain forms of speech and countries.[40] For an editor from Raskrinkavanje, this lack of platform engagement in BiH can be explained by simply looking at numbers: "There is power in numbers – if there is a large group of speakers of your language, your position vis-à-vis platform changes, these companies look at you differently, but this also changes your possibilities and opportunity."[41] There also seems to be a lack of political will to address some of these issues, as there are no mechanisms for public oversight and political

scrutiny in place that would mandate that these profit-driven practices align with the human rights legislative framework.

Consequently, this content moderation dynamic is clearly unbalanced in its implementation: with the unidirectional implementation from the (social media) companies to stakeholders and individual users, with no oversight or moderation in the other direction. On a macro-level, this disconnect has a 'visible' impact on democratic processes, as social media platforms have an increasing amount of 'opinion power', that is, power to dictate what information is visible and hidden, and for whom. And there is a lot at stake here – the proliferation of hate speech and disinformation, coupled with a lack of (visibility of) content that promotes peace and tolerance, profoundly colours public debate and public participation.

Content moderation: Local implications of global rules

A variety of local-sensitive contextual issues related to content moderation practices have emerged and, for the sake of this report, they can be generally categorised as follows:

- **Ethnic and political information manipulation**: inter-ethnic hatred, inflammatory narratives and related disinformation campaigns (see Case study 2);

- **Gender harassment**: assaults against women and under-represented groups (e.g. LGBTQI+ community, refugees, etc);

- **Complex and unresolved challenges** for media outlets vis-à-vis social media platforms, in particular Facebook;

- **Community-led initiatives** (It's All Witches and Love tales) that 'game' social media for 'good' (see Case study 3).

Ethnic and political information manipulation: inter-ethnic hatred, inflammatory narratives, and related disinformation campaigns

On a collective level, ethnic and political information manipulation are the result of the collision of political, economic, and societal interests and their influences over social media. This collision has established a constellation where "virtual wars on the Internet have transitioned from the physical world, and there the war is still unfinished."[42] This is a crucial segment of the social media landscape, which has resulted in ethnic and political information manipulation: 'through their bots, political actors have opened new battlefields in comment sections, which serve to raise tensions, fuel dominant narratives, defame political opponents and, essentially, narrow, rather than open space for pluralism of opinion.'[43]

Political elites in this environment are 'often sources of divisive messages, polarising and problematic statements' where 'both ethnic and political party media bias' are widespread. A 2018 study conducted by fact-checking platform Raskrinkavanje provided concrete data on this type of manipulation through analysis of readers' comments on the most visited news platform. It revealed the existence and modus operandi of political party 'bots'' and, over the course of a week, identified 259 user profiles that supported a particular ethnic-political party.[44] Most recently, a prominent politician from Republika Srpska spoke openly about his party's plans to conquer social networks and the methods by which they will do so.

In 2019, research carried out by the fact-checking platform on the scope of disinformation and related patterns analysed more than 450 online media articles, including those available on social media platforms, and found nearly 60% of the content to be inaccurate or manipulated, including political news and information. This study revealed two types of disinformation:

1. 'Opportunistic disinformation' which relies on anonymous online media and social media for monetisation, and
2. Disinformation spread by 'political and state actors', using public and private media 'to favour their political agendas'.

The authors also found that 'the coincidence of media misinformation and specific political interests indicates the possibility of the existence of targeted disinformation campaigns in the online space, including those related to foreign actors and sources.'[45]Another study, published by Mediacentar Sarajevo on problematic content online, found that mainstream media often ignore or work against the public interest.

The effects of artificial and self-interest-driven public manipulation on social media are diverse. In some cases, anonymous online media[46] ('a primary contributor'[47] to political disinformation and manipulation online and on social media) boost particular ethnic narratives.[48] Specifically, they create clusters of disinformation about a particular ethnic group, supported by country and neighbouring country media. In effect, these clusters result in broader networks of manipulation and disinformation that tweak the public narrative, given the sheer number of comments, thus creating the impression of broad public support. There are also other tactics that 'do not take a form of disinformation, but are manipulative in other ways',[49] such as the production of fabricated interviews to promote certain political parties, biased reporting of online media, and a proliferation of anonymous media that bolster certain political views. These manipulative practices were devised and applied in pre-election periods in 2018 and 2020. The same study found political statements, made throughout both elections, to be full of false claims (62 in 2018 and 80 in 2020). This study concludes that 'citizens of BiH have been 'bombed' with the manipulation that came from mainstream media (primarily from broadcast media), informational portal and anonymous Internet pages […] and social media, including political subjects participating in the election process.'[50]

These manipulative practices also target individuals that 'confront dominant ethno-national narratives'[51] – also referred to as 'character assassination'[52] – as was recently the case for two prominent women experts.[53] In these cases, the party responsible for online harassment was found to be a 'party bot', well-informed not merely about the lives and whereabouts of his targets, but also about the nuances of hate expression that do not carry the potential for sanctions. In other words, this person and many others like him have found ways to continue to harass and incite hate online while staying under the radar of content moderation and reporting mechanisms.[54] These types of information

manipulation undermine critically important journalistic content, such as investigative stories, using ethnic and other types of hate in comment sections to decrease their visibility and relevance. Another tactic, though not widespread, is the misuse of redress mechanisms on social media platforms for false reporting. Based on, for example, Terms of Services of social media companies, malicious users are able to block certain content, as in the cases of media outlets detektor.ba and CIN, or suspend accounts belonging to journalists and activists.[55] One such activist had to leave the country after being inundated with hate and death threats.

Consequently, independent media outlets, academic and civil society representatives, journalists, activists, human rights defenders, and members of minority groups – 'those that present an alternative to the existing mainstream discourse and political attitudes'[56] – face an increasingly shrinking civic space constrained by harassment and interest-driven manipulation, and must navigate an environment prone to manipulation and censorship.[57] The Internet and social media platforms have increasingly become spaces not for democratic debate and participatory citizenship, but for ethnic-driven manipulation, politically motivated hatred, and targeted group harassment.[58]

According to the media and civil society representatives interviewed for this report, these concerns are exacerbated by a lack of adequate content moderation, a situation that has further deteriorated given the Covid-19 pandemic. It is now clear that not all online speech is spontaneous and that harmful narratives surrounding religion and ethnicity, including homophobia and misogyny, among others, are now mainstays within the flood of disinformation[59] and anti-EU narratives – the West is largely presented as unprepared and lacking solidarity, especially when 'juxtaposed with Russia's strength, China's clever planning and the generosity of both.'[60] In the context of BiH, disinformation has evolved into a hybrid form of manipulative practices that blend conservative ideology, ethical-political agendas, and anti-western and EU views, together with anti-vaccination and conspiracy theory narratives.

Case study 2: The world of disinformation in the time of pandemic

Disinformation and 'right-oriented radicalising share pools'[61] in the country have mostly flourished on social media, Facebook,[62] and YouTube. The majority of the studies analysing the impact and scale of the problem, produced by Raskrinkavanje, indicate an increase in the spread of disinformation.[63] Some of this content has attracted enormous attention and, according to Raskrinkavanje's editor, some of the content is so inflammatory as to be dangerous: "disinformation can in fact kill people",[64] especially when it is related to public health, and in particular when it discourages people from adhering to and downplaying the importance of public health (e.g. anti-vaccination disinformation).

These studies identified a new generation of 'YouTubers', referred to as 'one hit wonders',[65] as a relatively new phenomenon for fact-checkers to contend with. It is not unusual for these YouTubers to hire online marketing and web experts, or engage with public figures, including politicians, to increase their reach and profitability.[66] Much of this content is produced in neighbouring countries, particularly in Serbia,[67] by YouTubers previously involved in the production of controversial content. Many of these content creators have used the platform to spread Covid-19 conspiracy theories, and other controversies to increase engagement and financial gain.[68] In so doing they have created 'a closed world that is not accessible or visible to a more critically-oriented audience'[69] in which pseudoscience, together with anti-Western sentiment, and fear-driven political targeting (QAnon content, for example) are the dominant narratives.

These studies highlight another significant concern – the more extreme and sensational the content, the more viral (when compared to less inflammatory content). The most prevalent misinformation revolves around the pandemic – assigning blame to China, fake remedies for the virus, conspiracy theories about Bill Gates creating the virus, and falsehoods about the effectiveness of masks. These narratives were assessed by fact-checkers as disinformation and fact manipulation, as baseless claims, but also as more sophisticated methods of manipulation, especially prevalent when it comes to Covid-19.

To respond to these occurrences, public officials have warned of legal sanctions for spreading disinformation.[70] In Republike Srpska, an individual was investigated and another fined for publishing content on social media inciting panic and spreading

disinformation. Another person was fined for criticising the government's handling of the pandemic. In addition, the Republika Srpska adopted a new decree that further strengthens the provision of the Law on Public Order and Peace and increases the sanctions for inciting panic and spreading false information online. Meanwhile, social media platforms, media outlets, and state actors continue to play a passive role, increasing 'the potential for disinformation to exacerbate internal conflicts, including, and especially, regional and international conflicts to which they are inextricably linked.'[71]

Gender harassment: assaults against women and under-represented groups

Despite the fact that online hate speech that targets individuals based on ethnic group affiliation remains the most pervasive form, gender harassment and hate speech against minority and isolated groups are also widespread. An analysis of digital rights violations, published by BIRN Investigative Resource Desk (BIRD) and the Share Foundation, covering the pandemic period (August 2019 – December 2020), identified and verified approximately 800 violations of digital rights and recorded 94 cases in BiH. Most cases (34) were violations in the form of 'Pressure exerted because of expression and activities on the Internet, including the publication of falsehoods and unverified information with intent to damage reputation, cause insult, threaten and endanger security', with the last category targeting mostly women, immigrants, and LGBTQI+-identifying groups and individuals.

Online misogyny takes many different forms, including direct threats, incitement to violence, and sexually-explicit and discriminatory speech. In the words of the director of feminist art collective Crvena:

> "*The position of women online is the same as their position in society, and the Internet is predominantly a man's space. Educational and other critically needed content cannot reach women and other groups through Facebook. The platform is not meant to be educational, nor does it allow people to rethink their positions and attitudes. It is a place for entertainment, and feminist thinking cannot be reduced to a funny video.*"[72]

For these reasons, women utilise private chat groups and offline spaces as safe and collaborative spaces. According to a Raskrinkavanje study, groups including immigrants, Roma, and LGBTQI+ communities are also regularly targeted on social media platforms by the same actors involved in the manipulation of political content (anonymous online media and via online media's comment sections), and often for political purposes. Through assessment of 628 news articles and 1,200 analyses over a three-year period (2017–2020), this study was able to identify several trends in hate and other harmful speech. Migrants and refugees are often labelled as threats to security, economy, and public health, in both articles/news content and comment sections, including and particularly via an anonymous media portal 'Anti-immigrant'. Women are primarily targeted with sexually-explicit and misogynistic content, as well as trivialisation and romanticisation of gender violence, while LGBTQI+ groups and individuals face hate stemming from bigoted cultural and morality norms.[73] The study found that there is 'a great chance that these assaults are in fact 'disinformation campaigns' capable of changing public narratives and attitudes towards these groups, but also, as 'clickbait' content, are economically incentivised to manipulate narratives to target marginalised communities.'[74]

Despite some positive changes, noted by prominent LGBTQI+ and feminist activists, advocates and members of marginalised communities still face regular threat, especially related to digital security. Members of these communities have, for the purpose of this research, shared their experience and lessons learned through monitoring efforts, primarily on Facebook and Instagram. They note that content directly inciting violence or hatred will most likely be taken down after flagging, but more subtle forms of hate – misogyny, bigotry, racism – often remains online because 'Facebook algorithms do not have the capacity for interpretive and qualitative analysis'.[75] Moderators noted that if they flagged content as related to 'Sex and Gender Identity', this led to a quicker reaction from the social media platforms, than when they flagged content related to 'Sexual Orientation'. They also claim that mass reporting (simultaneously reporting dozens of posts and comments from the same individual) is more efficient than reporting the individual account. All agreed that content removal and redress mechanisms were hard to grasp,

prone to constant change, and in need of urgent reform.[76] They report all hateful content and incitements to violence to the police on a daily basis and have found a method to engage with Meta: namely, utilising personal contacts within the company to flag content for quicker removal.[77]

Community-led initiatives

Case study 3: The social media for peace, equality, and diversity

This case study presents a different face of social media companies. Imagined as participatory, inclusive, and alternative, these online spaces, including the feminist and activist-initiated 'Sve su to vještice' ('It's All Witches') and the Balkan Discourse Platform from the well-known peace-building organisation Post-Conflict Research Center (PCRC), emerged as research-backed examples of healthy initiatives that 'game for good' comment moderation and social media algorithms in a way that virality of content entails more positive comments and healthy interactions.

'Sve su to vještice' ('It's All Witches') was founded in 2015 by a young academic and feminist activist from Sarajevo. At that time, when the diversity of female experience was not represented on the Internet, videos, gifs, and memes (altered and funny videos) had become a popular method for expressing opinion. The platform became the first in the region to offer humorous perspectives on the female experience as a method of media activism (communication of feminist and empowerment messages via humour and satire). It's All Witches is her response to, and disagreement with, the pervasive violence and scrutiny that women experience.[78]

Her posts are explosive, combining academic feminist thinking and use of vernacular language to expose oppression, patriarchal culture, and pervasive misogyny. Focusing on a range of topics (including the uneven distribution of housework between men and women, emotional labour, etc), she has built a safe and accessible space for women to share their experience and connect through humour.

In her words:

> "*Some women tag their husbands below the posts and put a heart next to the tag. Under some posts, girlfriends make fun of each other, and it makes me happy to see these interactions. That means that Witches is doing something good for the women in BiH. Humor is the key – every time misogyny pops up and it threatens someone else – I use my humorous approach, my modus operandi to discourage haters.*"[79]

This approach seems to resonate with her audience. Surprisingly, the page does not attract negative reactions: "that's because 'Witches' are multi-layered beings – they are well educated, but are also ordinary women, and not everyone will always understand them."[80]

It's All Witches main presence is on Facebook, with more than 60,000 followers, of which 18% are men. Since last year, they are also active on Instagram (around 10,000 followers) and YouTube, where they publish a podcast about women's diverse experiences for roughly 300–500 listeners per episode. Their reach and interaction with the audience is completely organic, and they do not allow sponsorships for their content. Unfortunately, Facebook has changed a lot since 2015, in ways that could mean a much smaller audience and decreasing reach. "I notice that my reach has decreased. I used to get 20,000 views per post, and now it's about 7,000–8,000, because of their (Facebook's) algorithms that encourage me to pay to promote content. And that's also why I work offline."[81] In a country where journalists, activists, and critical thinkers face threats and hate speech online, Facebook's profit-driven strategies can have a serious chilling effect on community activism. This particular risk was highlighted by a professor of communication at the Faculty of Political Science of the University of Sarajevo who observed that community activism and activists are returning to offline spaces to express themselves and create change.[82]

According to the activist, Witches "serves as a mechanism for communicating the patriarchal 'cover-up'. It's not a storytelling platform – it's not reactive. I give them didactic tools to improve their daily lives. And if it also inspires laughter at the memes, that's enough for me."[83]

The PCRP, known for its peace-building research, documentaries, and engagement with local youth communities, developed a project in 2011 to collect local and community stories of courage, love, and heroism during the time of the armed conflict. The project,

'Ordinary Heroes', included young activists, human rights defenders, and high-school students, and supported the production of multimedia material and exhibitions showcasing these diverse and truly human experiences. Because they felt that so many other stories deserved to be told, the Balkan Discourse Platform was created to provide a virtual space for conversation and content production. As the director of PCRP explained:

> "Platform reports about the stories of communities are not available in the mainstream media because they are not explosive. But they are very important for small communities and for these young people. This was our first leap into the online world, and this is where we realized that audiences respond well to these stories."[84]

Utilising social media, primarily Facebook, to share the young activists' stories, PCRP have cultivated a network of 125 young people who are provided basic journalism and photography training. PCRP's activists produce local stories that address issues relevant for their communities. PCRP's editorial agenda is different from that of traditional media, since young activists chose the stories they cover, many of which cover experiences during conflict, rather than explicit peace-building processes.[85]

A similar project, 'Love Tales', which forms part of the Balkan Discourse Platform, was also launched as a storytelling platform for sharing experiences of inter-ethnic and queer relationships. Initially, platform staff were concerned that the project would attract excessive hate, given ethnic tensions in the country. "We were really wonderfully surprised that we used sponsored and organic reach and there was no problem. We did not have to remove one single comment."[86] For decades, PCRP has worked to establish a healthy relationship with their audience. They have been able to rely on this relationship and the potential of social media to communicate important community content. These efforts have opened up new ways of communicating with audiences, and cooperating with other youth activists, as well as similar regional and international projects. "During the pandemic our reach and work tripled, but it gave us a new perspective on how to talk about peace and what peace means to people. 'Love tales' have shown us how important this content is for people from all over the region, and, without social media networks, it would not be possible to reach these people."[87]

In essence, the Balkan Discourse Platform is not only a peace platform, but also a space for the local community to explore, communicate, and produce stories that transcend

country divides. Local audiences return to it because it focuses on real problems: "it is what they see in their community – discrimination, gender issues, ethnic conflicts, and this is what social networks should offer – space for young people to create and consume community content."[88]

Complex and unresolved challenges for media outlets vis-à-vis social media platforms, in particular Facebook

In BiH, the fragmented nature of political discourse, the strategy of 'trading in fear', and the fractured public debate, have together resulted in a media landscape in which political, business, and other profit and power-driving interests overshadow meaningful media freedom. 'Hence it is difficult to talk about the media landscape as separate from politics.'[89] Public broadcasters are in a [weak position economically](), and some are seen as a mouthpiece of the government,[90] whereas print media is in serious decline and, as the data from the previous section suggests, social media platforms and online media have become key information resources for a large part of the population. As a result, and in effect, there is no genuine media pluralism in the country: "We can talk about media pluralism, because there are Serbs, Croats and Bosniaks, but not because we have an independent media space or real pluralism."[91]

When it comes to vulnerability due to financial instability, as much as in any other country context, media in BiH have fallen prey to social media business models; namely, that inflammatory content and comments attract engagement and are economically profitable, and many media outlets depend on social media companies to reach their audience, which in turn generates profit. The biggest generator of comments, including hate speech, is the country's most visited online platform, klix.ba.[92] In addition, media outlets report that they lack the capacity to hire content moderators to facilitate their online news platforms and social media accounts and must allocate the task to a journalist, leaving an empty space in the newsroom. Some media outlets claim that content reach on social media decreases

if they carry out active content moderation by banning and hiding users. This points to the connection between algorithmic-driven moderation and a saturated but heavily under-financed media market effectively supporting a lack of moderation and bolstering harmful and hateful narratives.

In addition, media outlets are desperately trying to tame the social media algorithms for content moderation and curation to ensure that their news content reaches their audience. In the case of CIN, an investigative media outlet, the problem began after YouTube took down a documentary produced and promoted by the outlet on social media, citing violation of company copyright regulations. It took them 24 hours, and a complex process, including communication with an artificial intelligence chat bot and filing of official purchasing copyright documents, before the documentary was restored. But the damage had already been done (as the European Court of Human Rights found: 'the news is a perishable commodity'[93]) and CIN missed the crucial window for reaching a wide audience to inform the public about the film. This case is relevant for another reason: the company that reported the copyright infringement claim to YouTube was the same company known for making other false claims for the purpose of content removal in Serbia. Specifically, in 2014, the company successfully requested the removal of a political satire video and other content from YouTube. This company (and others with similar goals) constitutes a threat to independent media outlets in the country as long as YouTube regulations allow anyone the ability to interfere with news dissemination processes. As the director of CIN noted, "these internal control checks are meant for good, but you never know when they will be abused by different power centres."[94]

In another case, media outlet BiH BIRN (detektor.ba) published an extensive report on the activities of an Islamist group in the country, revealing their financial malversations. In response, the group posted a how-to video to Facebook for their followers on reporting the BIRN document for take-down. While the video remained available on Facebook, it did not lead to the removal of the BIRN report. Still, the media outlet had no tools or effective redress procedures at their disposal to alert the platform to the false reports. They did attempt to contact Meta Europe offices, but without much success.[95]

However, for media that depend on Facebook to reach their audience, as is the case for media watchdog Analiziraj.ba, they have repeatedly experienced Facebook as a censor. In 2021, Analiziraj.ba regularly tried different methods to work around Facebook's algorithm in order to be able to sponsor content, including in August 2021 when they were denied the option to pay for the promotion of content for allegedly attempting to circumvent Facebook policies. As a result, the Analiziraj.ba account was suspended, though they were later able to request a review of the decision, exchange emails with a human employee, and reactivate the account. Upon sending a simple thank you email to Facebook for the reactivation, the platform issued another suspension of the account, and the process started again. On other occasions, they were not allowed to sponsor content with extended audience reach to Serbia or post an article with a metaphor title interpreted as hate speech. According to Analiziraj.ba, this situation is untenable, with media paying to reach audiences, but with no control of promotion and reach, and no understanding of these processes:

> "*We are not aware of the rules behind this since their rules aren't transparent. They do not even have an office in the country, and yet they get to earn money and make decisions about the visibility of our content. This is an invisible power that is exercised every day and that surely impacts what citizens see online.*[96] *Facebook can even decide to close our page for reasons we are not aware of, and we would not be able to do anything about it.*"[97]

In a similar vein, BiH BIRN (detektor.ba) attempted to promote content on Facebook related to a war crimes judgment issued by the International Tribunal for the former Yugoslavia (ICTY), which was very relevant for the process of reconciliation and truth finding in post-conflict countries. Facebook's evaluation of the article took several days, and then the platform declined promotion.[98] In another case, a media outlet, Raporter, from Sarajevo was prevented from publishing content related to the ICTY war crimes judgment in the case of Ratko Mladić, a notorious war criminal charged with genocide in Srebrenica, because of a 'misinterpretation'. Facebook's algorithms concluded that the article sought to promote 'criminal organisations', which resulted in this label being applied to the ICTY, thus the article was not published. The irony is that the Facebook

algorithm wrongly classifies the ICTY as an criminal outlet, while allowing outlets like Despotovina.info (known for its [glorification of Ratko Mladić](#)) or SAFF (a Muslim religious magazine known for their anti-LGBTQI+ articles and affiliation and Bosniak ethno-national political agenda) to operate active social media pages.[99]

Independent media outlets in the country seem to be increasingly aware that their content is treated differently than clickbait content, though there is little they can do to change this situation. Outlets can pay for a larger reach, though the structure of their audience is likely to remain the same given that algorithmic processes tend to create certain feedback loopholes.[100] The guerrilla tactics of media outlets, but also of activists devised to address the issues in the noted disconnect between social media platforms and the country, are a painful reminder of the darkness in which social media users and media are being left in by these companies. It is not known, however, how these processes limit individuals from receiving crucial (especially in a post-conflict society) public interest content. Media actors have clear reason to perceive social media companies as 'invisible and powerful censorship machinery',[101] given the dozens of experiences that indicate profoundly negative impacts on media freedom in the country.

Interim conclusion

Harms stemming from (a lack of) content moderation are real and they have great potential for amplification, thus reinforcing societal tensions and divisions.[102] While the understanding of societal context is crucial to the resolution of any content case, the algorithmic-driven content moderation is often seen as context-blind, resulting in certain content remaining online even if it constitutes a violation of international human rights principles and/or social media platform internal rules, while other pieces of content may be blocked from appearing online despite their public relevance, as noted in the previous examples. But social context is also crucial for another reason: in post-conflict countries, the harms linked to algorithmic content moderation scale differently – often to the detriment of freedom of expression and media freedom, not to mention an increase in the risk of further societal tensions.

To address these issues, ARTICLE 19 and other civil society organisations have developed recommendations:[103]

- Companies should ensure that their content rules are sufficiently clear, accessible, and in line with international standards on freedom of expression and privacy, and it is of key importance that social media companies' content rules be made accessible and available in local languages.

- Companies should also provide more detailed examples or case studies of the way in which their community standards are applied in practice and conduct reviews of their standards to ensure human rights compliance:

- Companies should be more transparent about their decision-making processes, including the tools they use to moderate content, such as algorithms and Trusted Flagger schemes.

- Companies should ensure that sanctions for non-compliance with their Terms of Service are proportionate.

- Companies should put in place internal complaints mechanisms, including for the wrongful removal of content or other restrictions on their users' freedom of expression. In particular, individuals should be given detailed notice of a complaint and the opportunity to respond prior to content removal. Internal appeal mechanisms should be clear and easy to find on company websites.

- Companies should publish comprehensive transparency reports, including detailed information about content removal requests received and actioned on the basis of their Terms of Service. Additional information should also be provided in relation to appeals processes, including the number of appeals received and their outcome.

- Companies should collaborate with other stakeholders to develop new independent self-regulatory mechanisms, such as a social media council, modelled after effective self-regulation archetypes in the journalism field.

ARTICLE19

Of relevance for this research is the Principle of Cultural Competence in the [Santa Clara Principles on Transparency and Accountability in Content Moderation](#), which posits that 'those making moderation and appeal decisions understand the language, culture, and political and social context of the posts they are moderating.' This Principle requires that:

- Users should have access to rules and policies and notice, appeal, and reporting mechanisms that are in the language or dialect with which they engage. Users should also have confidence that:

  - Moderation decisions are made by those familiar with the relevant language or dialect;

  - Moderation decisions are made with sufficient awareness of any relevant regional or cultural context;

  - Companies will report data that demonstrates their language, regional, and cultural competence for the users they serve, such as numbers that demonstrate the language and geographical distribution of their content moderators.

The question of local representation may raise specific concerns. ARTICLE 19 has observed that local laws sometimes require social media companies to establish local presence in the country. This is a matter of concern, particularly in countries with governments with a poor record on the protection of freedom of expression, since such national establishments may facilitate pressure towards the removal of content that would be considered legitimate under international human rights law. However, social media companies could nonetheless make themselves easily accessible to local stakeholders through online means that would enable local actors to effectively engage with the companies.

Analysis of stakeholders

During the course of the research project, a number of interviews with experts and stakeholders were carried out to further understand the content moderation concerns and dynamics that have a profound impact on digital public space, freedom of expression, and media freedom in the country. The interviews also touched upon a number of topics related to stakeholders' knowledge, positioning in the broader societal landscape, needs, and objectives. Based on their responses, personal insights, and the overall political and societal situation in the country, this chapter provides an in-depth analysis of stakeholder interests by categorising them into groups that emerged during the interview process as key drivers of change and expertise:

1. Media organisations and associations
2. Civil society organisations
3. Academics and scholars
4. Activists and human rights defenders

While this initial research has focused on these stakeholders, the role of a Coalition on Freedom of Expression and Content Moderation would include meaningful engagement with social media platforms[104] and state authorities.

Against this backdrop, this section delves more deeply into the noted groups of stakeholders' categories by analysing their scope of work, programme orientation, and capacities, including obstacles and gaps (see the following for details of stakeholders and a short description of their portfolios). The goal of this analysis is to create a list of initial stakeholders whose involvement and active participation should enable the coalition to address the previously noted disconnects and gaps in content moderation emerging within the 'terra nullius'. The identified stakeholders are not the only actors in the country working in this field but have been identified given their particular expertise and practical engagement with the issues. In addition, in line with an inclusive approach to building the coalition, these actors are well-positioned to identify and attract others (see Annex B).

Media stakeholders

Key actors

- Investigative media outlet: **CIN**

- Media watchdog: **Analiziraj.ba**

- BiH BIRN media outlet: **Detektor.ba**

- Fact-checking platform: **Raskrinkavanje**

- Most popular online media portal: **Klix.ba**

- **Mediacentar Sarajevo**

- **Press Council**

- Journalists' Association: **BH Novinari**

Programmatic focus

- **CIN**, **Analiziraj.ba**, and **BIRN** focus on the production of news content, especially of public interest, such as corruption and organised crime (**CIN**) and extremism; rule of law, war crimes adjudications and corruption (**detektor.ba**); and monitoring media content (**Analiziraj.ba**). They constitute the few independent outlets and are funded primarily on a project basis.

- **Klix.ba** is especially relevant as it is one of the largest producers of media content but is also a generator of harmful comments.

- **Mediacentar Sarajevo**, a civil society organisation and media outlet, focuses on journalists' capacity-building, advocacy, research, and media and information literacy. They run the media-focused website media.ba, and they also host diskriminacija.ba, a website focusing on hate speech and minority rights in collaboration with local civil society organisations, and digital media archive repository, infobiro.ba.

- The **Press Council** is a self-regulatory body and their core objective is to strengthen media freedom and foster the professional independence of and respect for journalistic standards.

- The **Journalists' Association** primarily focuses on safety and protection of labour rights.

Flagship projects

- **Mediacentar Sarajevo** is heavily involved in a number of regional projects, like the [South East Network for Professionalization of Media (SEENPM) Resilience programme](#), which is funded by the EU and seeks to strengthen media freedom and counter disinformation and hateful propaganda. They are also in collaboration with local civil society organisations, coordinating the website that focuses on marginalised groups and hate speech (diskriminacija.ba); however, this local collaboration has faced funding and support problems recently.

- **Raskrinkavanje**, in addition to fact-checking, also uses analysis to bring public attention to lesser-discussed topics, including the lack of adequate content moderation.

- The **Press Council** has, through the Council of Europe project '[Reinforcing Judicial Expertise on Freedom of Expression and the Media in South-East Europe (JUFREX)](#)', offered support and trainings to legal experts, members of the judiciary, and media actors that address a range of freedom of expression issues, including [hate speech](#) – which is also the topical focus of their monitoring project, '[Stop! Hate-speech!](#)'.

- The **Journalists' Association** has also been involved in the JUFREX project, but with a primary focus on the protection of journalists, leading to the creation of a [network of media lawyers that help with cases of online harassment](#). They also offer mental health services, legal support, and are involved in research processes that touch upon a range of content moderation issues.

ARTICLE 19

Capacity and needs

In the identified organisations and media outlets, human capacities to support the work of the coalition may be limited due to their current programmatic activities, which absorb all of their resources, limiting their abilities to engage in additional and non-program focused activities. The funding model of some of the identified civil society's organisations is reliant on external sources of funding, whereas Klix.ba, as a private media outlet, depends on the market. These models may cause risks in terms of the sustainability of the coalition and internal capacity within these organisations to further its work.

Media outlets and media organisations are aware of and are knowledgeable about when it comes to the problems caused by the current state of content moderation. There is a growing interest in this topic but there are no in-house programmes (e.g. research, advocacy initiatives).

Mediacentar Sarajevo organised discussions on the topic of social media and freedom of expression.

The **Journalists' Association** conducts research and provides protection against online harassment.

However, two critical issues surfaced in the interviews:

1. System response to the local contextual issues (hateful comments, inability to promote content on Facebook) is reactive and fails to effectively address the problem; and
2. There is an obvious gap in terms of meaningful cooperation across and within all stakeholder sectors, hindering effective solutions for content moderation problems.

Thus, there is a need for genuine support for programmes and initiatives that actively and comprehensively address these problems and involve social media companies in the resolution of problems.

Civil society organisations

ARTICLE19

Key actors

- Human rights protection and democratic participation organisation: **Zašto ne?**

- Peace-building organisation: **Post-Conflict Research Centre** (PCRP)

- Regional human rights investigative network: **BIRN Hub**

- Social justice and technology organisation: **Internet Society of BiH**

- Human rights regional organisations: **Civil Rights Defenders**

- Feminist collective:[105] **Crvena**

Programmatic focus

The above-mentioned civil society organisations are known for their dedicated activism, potential for societal change, and strong focus on public participation, inclusion, and freedom of expression. These civil society organisations have built micro-spaces both off- and online, and target different groups and audiences through a range of activities and programmes. They are also part of regional organisations and networks, such as BIRN Hub, and Civil Rights Defenders, or collaborate with regional experts and audiences like Crvena and PCRP.

Flagship projects

- **Zašto ne?** has been recognised as a key organisation and knowledge-generator in the domain of Internet freedom, not only because of programmatic work such as Istinomer (Trust Meter) and Raskrinkavanje, but also because of their genuine involvement with a range of regional and international actors via an annual conference 'Point'. They are also involved in the global and regional topical discussions and are members of the recently established SEE Digital Rights Network.

- **PCRP** is already well-established as an organisation that fosters social cohesion and strengthens local initiatives through the Balkan Discourse Project and, as such, is a crucial actor in the peace-building process, including on social media.

- **BIRN Hub** is one of the first regional human rights and investigative platforms and has garnered broad public attention for their regional work. Their current monitoring and reporting program, BIRD, provides monitoring data on digital rights and, through this program, produced a series of articles available across the region. Together with the Share Foundation, they coordinate the work of the SEE Digital Rights Network.

- The **Internet Society of BiH** is a recently established organisation and is a member of the Internet Governance forum, whose members include tech-engineers, professors, and enthusiasts. Their agenda focuses mostly on social justice and inclusion, though they are still heavily involved in the process of fundraising.

- **Civil Rights Defenders** operate across the region and their programme in BiH mainly focuses on offering support to media organisations and civic media for content production. During their annual event, '<u>Regional Journalistic Days</u>', they address a number of issues that journalists face online – particularly on social media.

- **Crvena** predominantly focuses on exploring and discovering intersectional oppression and power dominance through art and education. This work <u>puts them in touch with regional feminist collectives</u> and those working within the art scene.

Capacity and needs

The identified organisations have been operating in a financially unstable civic field for a number of years and have proven their ability to overcome a number of challenges and obstacles (political pressure, financial insecurities, etc). This by no means implies that their financial situation is stable, but that they have built resilience and showed innovation and resourcefulness, positioning themselves as important actors within their respective domains. In addition, there is growing, but still limited multidisciplinary expertise (in terms of numbers of individuals working in civil society organisations) in the field of Internet freedom, freedom of expression online, and content moderation. This expertise will be crucial in the coming period, but there remains serious concern that it is far from sufficient to address serious and systematic loopholes left by the lack of social media involvement in resolving content moderation issues.

With their specific domain and through targeted activities, these organisations have been able to find a niche within the wider civil society organisation landscape but have limited capacity (except perhaps for Zašto ne?) when it comes to involvement in broader advocacy, societal, and political initiatives. There is a need for better communication, representation, and inclusion of civil society organisation expertise in political and societal processes at the county level. There is also a need to increase the pool of experts, and to develop programmes specifically targeting content moderation and Internet freedom.

Academia

Key actors

Universities across the country, in particular **University of Mostar**, **University of Sarajevo** (Faculty of Political Science and Faculty of Philosophy), **University of East Sarajevo**, and **University of Banja Luka** have emerged as important actors that produce relevant knowledge and academic resources and are meaningfully involved in the work of civil society organisations. The coalition should include representatives from academic institutions across the country and the researcher has reached out to potential participants in that respect.

Programmatic focus

The curricula developed by these academic institutes focus, among other topics, on social media and Internet and media freedom, but also on the broader information society, addressed separately by the department of librarianship and information society at the Faculty of Philosophy. During the interview process, these faculties emerged as academic institutions with a strong interest in content moderation-related topics.

Flagship projects

The Institute for Social Science Research (at the Faculty of Political Science of the **University of Sarajevo**), with participation from academics from across the country, has been heavily involved in a number of strategic projects, one of which is a UNESCO-run project on media literacy. They have produced academic and policy resources addressing

a range of media and information literacy structural loopholes to ensure that media and information literacy is addressed as a cross-sectoral curriculum throughout the educational system.

Capacity and needs

Academic institutions and scholars often have to manage multiple projects and programme activities simultaneously. In addition, there are a number of administrative hurdles that prevent them from engaging with experts from outside academia.

In addition to the noted financial insecurity, a lack of genuine expertise and over-burdened scholars signal the need for additional expertise and capacity-building processes to increase the resource pool. While content moderation is recognised as an important topic, these constraints may result in it being swept to the margins of academic agendas.

Activists and human rights defenders

Key actors

Despite the fact that most of the identified actors/activists and human rights defenders are affiliated with the above-mentioned civil society organisations or academic institutions, their community engagement in the context of Internet freedom, including content moderation falls outside their professional tasks. The identified activists interviewed for this research have expertise in various fields, including Internet freedom and open tech/data, online feministic activism, LGBTQI+ rights, and women equal participation.

Programmatic focus

These activists use social media and other digital spaces to engage in political discussions, produce knowledge on different related topics, and build capacities of different organisations, including media outlets. They are also engaged in off- and online activism for the general improvement of the lives of marginalised community members.

Flagship projects

The researcher and activist surveyed is affiliated with the University of Sarajevo's Faculty of Philosophy. As an activist he is deeply involved in political and societal events that can affect Internet freedom, and he is regionally recognised for his work.

The efforts of the interviewed LGBTQI+ activist and her team have led to the BiH Pride March for the past three years and an improvement of their treatment online and offline.

The feminist activist is running her online profile and podcast 'It's All Witches' story and provides tools for women emancipation online, thus gaming this space for diversity and dignity.

Capacity and needs

Activism is an ongoing struggle to improve public participation and ensure equality for all. Their work and online presence make them especially valuable resources, especially in BiH, where activism for marginalised communities and interests is often limited.

Like other stakeholders, activists and individual experts are not able to garner the required public attention. There is a need to propose new ways to create those conditions to support the efforts of experts and online activists in general, but also to create space for new activists, especially those from the local/marginalised communities.

Obstacles to addressing content moderation challenges

As this list of stakeholders indicates, in BiH there are no actors, programmes, or projects that address content moderation on social media, as defined in this report. This structural gap does not come as a surprise, given that content moderation and social media networks give rise to obstacles that require a multi-pronged and multidisciplinary approach. Existing initiatives, such as research projects by media stakeholders, civil society organisations, and academia that monitor efforts and produce news content, address content moderation through a singular lens and creates a situation of 'knowledge silos' and hidden expertise. Stakeholders and their initiatives are only able to map the

terrain and react when a problem emerges, but they have limited resources to propose proactive initiatives to address problems that emerge, given the disconnect between social media companies and the country. There is also a lack of genuine cooperation among actors and a lack of suitable platforms for exchange of expertise for civil society organisations and media organisations, including academic institutions. The existing networks that partially tackle some of the issues noted in the report are the recently established SEE Digital Rights Network coordinated by Share Foundation and the BIRN Hub whose member is Zašto ne?. In addition, the Coalition to Combat Hate Speech and Hate Crimes, founded in 2013, advocates for the improvement of the legislative framework on hate speech and public awareness campaigns. Organisations mentioned in this report like Civil Rights Defenders, Mediacentar Sarajevo, Journalists' Association, Press Council, and Fondacija CURE are members of this coalition. These networks that engaged with the topics of freedom of expression and media freedom could be an important source of information in helping to shape the governing structures and cooperation models for the future coalition. Finally, all mentioned stakeholders struggle to find sustainable funding and are subject to donor-funding strategies and agendas, and they must often tailor work to address these issues rather than addressing the data-driven needs of the digital ecosystem.

Among interviewed stakeholders there is a shared perception of content moderation as a predefined and impenetrable system, meaning that they can only address (negative) consequences, reducing its impact on freedom of expression at the individual level. For this reason, there are no initiatives that propose policy ideas, practices or advocate for context-specific solutions to content moderation concerns. This expertise and programmatic gaps are exacerbated by the fact that state actors are perceived as 'observers', often lacking expertise, a legal mandate, or the capacity to engage in those discussions taking place on the margins. In addition, respondents agree that involving all the government structures in their efforts is much too complicated: "you need to go through a lot of political loops to make your problem meaningful, and still, you will not succeed."[106] Furthermore, given that public officials sometimes initiate hate campaigns, involving them in solution-oriented and advocacy work seems, to many respondents,

counterproductive and meaningless. Therefore, despite growing expertise, the lack of effective and sustainable cooperation and support – especially amongst international, European and regional communities – is a significant obstacle to addressing content moderation challenges.

ARTICLE[19]

Conclusion

This report elaborates on the ways in which human and nonhuman agents as well as external (global) and internal (country) content moderation processes shape the digital and media landscape, and the impact on individuals and communities in BiH. Content moderation in BiH has surfaced as a set of processes and practices that are primarily algorithmically driven, political in practice, and shaped by the net outcome of intersecting and fractured interests in this ethnically divided country. This report is an important contribution to a discussion that currently only exists in the margins, surfacing local and country-specific implications of and responses to globally transformative processes.

Throughout the country's history, especially its recent history, mass media, primarily television and public broadcasts, have played an important role in promoting intolerance, ethnic divisions, and hatred. For this reason, BiH communities are extremely vulnerable to hateful narratives, especially those coming from media, political officials, and influential (online) figures, and resulting implications for societal cohesion and peace-building processes. The emergence of non-accountable actors, like social media platforms, have resulted in additional obstacles for and threat to a shrinking civil society and public media space. Indeed, in BiH social media platforms have become powerful forces that have reconfigured public debate and participation – the cornerstones of democracy – yet show little interest in resolving these issues in this part of the world.[107]

Among the general population, the Internet is not perceived as a hybrid space,[108] but rather as a tool for entertainment, communication, and a platform for receiving information, including news content. This misconception, among other factors, has created an atmosphere of impunity for those involved in information manipulation and hatred, which, in turn, chills speech and public debate, as everyone is a potential target of hatred and threats. Yet, in BiH, "the Internet is the most free territory we have"[109] and, for this reason, "content moderation is important, because it took years for the civil sector to convince people that they should express themselves, that no one would send them to prison. But now some new kind of fear exists (not from the authorities) that someone (invisible) online is watching us and knows everything about us."[110]

ARTICLE¹⁹

Given the prevailing ethno-political narratives and social media infrastructure, the space for an independent critical response is closing, and the number of viable actors for change, dwindling. Existing actors do not have strength in numbers and are often from under-represented communities and lack the experience, avenues, and competencies to bring about change. It is crucial, therefore, to increase the capacities of the existing actors about the issues covered in this report through the lens of local dynamics, and hope that their increased capacity will increase the interest in this topic on a societal level. For this reason, the proposed coalition could become an essential mechanism to strengthening these voices and addressing the analysed local-sensitive issues of content moderation through ensuring transparency and availability of internal documents, and also as a redress mechanism in the local language to give a clear understanding of the content governance regimes, thus improving the effectiveness of problem resolution mechanisms, and providing support for independent media and effective channels of collaboration with social media platforms (e.g. focal points).

There is a lot standing on the agenda of such a coalition; therefore, it is crucial that it is not designed and resourced as merely a project-based exercise, but rather a long-term and sustainable collaborative platform, altering the current course and creating a path to genuine civic participation.

Recommendations

In the context analysed in this report, a **Coalition on Freedom of Expression and Content Moderation** could become a critical mechanism to bring together a range of stakeholders to address the complexities of content moderation. At the same time, a capacity-building process should ensure that all coalition members join the platform on the same footing and can contribute to the definition of its agenda. To this end, a participatory and inclusive approach is essential to increase trust and collective ownership of progress on the part of all members.

To facilitate this process, the researcher suggested that a three-party coordination team could be established to initiate and manage the coalition.

During the course of the research process and interviews, three organisations surfaced as potential leaders of a coordination team:

1. **Institute for Social Science Research**, established at the University of Sarajevo's Faculty of Political Science;
2. **Press Council**, a self-regulatory media authority;
3. **Zašto ne?**, a well-established civil society organisation that is currently implementing a fact-checking project, Raskrinkavanje.

The **Institute for Social Science Research** is an important actor in this field with a high level of individual and multi-sectoral expertise. It has been involved in similar research and capacity-building processes, including cooperation with UNESCO offices in Sarajevo, and state actors to establish new and strengthen existing media literacy interventions and programmes. At the individual level, the Institute's experts are well-respected scholars and university professors with various knowledge backgrounds, including security and human rights, library and information science, communication and journalism, and political communication, to name just a few. The Institute regularly engages with scholars from across the country (the University of Istocno Sarajevo, the University of Mostar, and the University of Banja Luka), as well as with state institutions, thus creating a non-formal multidisciplinary network of experts. For this reason, the Institute and the multidisciplinary

and multi-ethnic group of scholars and experts has been identified as the best-positioned actor for coalition leadership, including agenda-setting processes and facilitation of work.

The **Press Council**, a voluntary self-regulated body, is a well-respected actor, both by media outlets and other stakeholders. Given that their membership includes a large number of media outlets, they serve as a sort of gatekeeper for communication with and amongst media outlets – electronic, online, and print. They also act as a mediator when problems arise between media outlets or other stakeholders. The Press Council would be a logical choice and have expressed willingness to be included in the coalition.

Established twenty years ago, **Zašto ne?** has been recognised as a key driver of knowledge and change in the field of engaged citizenship and democratic deliberation, but also for freedom of expression online and media freedom. They have established a publicly available platform for political engagement (Javna rasprava) and the inclusion of citizens in legislative and policy processes. In addition, they generate knowledge surrounding issues related to Internet freedom and the use of technology to improve democratic processes and facilitate the work of Raskrinkavanje.

While these three organisations have emerged as potential leaders during the research, membership of the coalition could extend to a number of actors to ensure an effective representation of the whole diversity of the country (an initial, open-ended list is identified in Annex B).

- Representatives of social media companies should be invited to take part in regular meetings with the coalition.

- The UNESCO Antenna Office in Sarajevo of the Regional Bureau for Science and Culture in Europe could act as a focal point to offer logistical and organisational support for coalition cohesion, and ensure that the coalition is able to work together with key state actors.

- Given the sensitive and fragmented political situation in the country, state authorities should be involved in consultations on the matters that fall within their scope of competence, particularly during the coalition's early work.

The goals of the coalition would be to:

- **Identify a common understanding** of the most pressing issues related to content moderation;

- **Propose solutions to resolve these issues** in conformity with international standards on freedom of expression and other fundamental rights;

- **Engage in sustainable and transparent collaboration** with social media platforms to address identified issues and implement solutions.

An initial step would be to ensure that members have adequate knowledge of content moderation processes and, through a participatory process, contribute to a governance framework and agenda. This requires **tailored training programmes** for a wide range of stakeholders to be delivered by future coalition members (via joint educational platforms for capacity-building on content moderation, for example).[111] The underlying idea behind these capacity-building processes is to increase the level of understanding on the topics covered in this report (e.g. content governance, social media power and logic, regulatory approaches, and multi-governance regimes, etc) and the enabling potential of social media shifting the common understanding of the Internet as a tool to the **Internet as a hybrid space for public debate, communication, and interaction**. In parallel, this training should also identify and explore further potential and the impact of locally-led responses and projects targeting these issues, in addition to sharing knowledge and experience of the regional and international organisations from this field. In this way, coalition members will gain important advocacy skills and knowledge enabling them to establish and maintain cooperation with social media platforms, state authorities, and other global and regional stakeholders.

There is also a need for training in relation to coalition building as multi-stakeholder platforms are not a commonly used instrument for collaboration.[112]

Both capacity-building initiatives – content moderation and the creation of multi-stakeholder platforms – should ensure that coalition members have genuine independence and, utilising participatory processes, create a culture of dialogue and

collaboration. In this way, the coalition could develop a real potential to propose and develop 'a collective approach to collective problems'[113] and 'to articulate local contexts and concerns in a way that are understandable to social media companies'.[114] These are necessary prerequisites for the coalition to be officially established as an independent and multi-stakeholder entity that can effectively give local stakeholders a voice in content moderation and advocate for the protection of freedom of expression online.

The following focus areas have emerged from the interview process as complementary fields of interests:[115]

- Advocate for the creation of transparent social media platform points of contact for the country;[116]

- Analyse and re-draft company policies and transparency reports, in-line with international and human rights standards;

- Launch advocacy initiatives with state authorities, international organisations, and social media companies to ensure BiH visibility in global fora;

- Develop public awareness-raising activities to increase knowledge of social media and content moderation processes in the general public.

In BiH, a country lagging behind when it comes to Internet and content regulation policies and discussions, the coalition could become a necessary platform to fill in this gap using participatory dialogue, co-creation of policy, innovative practices, and advocacy initiatives. It is essential that the coalition is built with a long-term plan and vision for future work, to ensure sustainability and longevity of its efforts.

Annex A: Risk analysis

The **Coalition on Freedom of Expression Online and Content Moderation** emerges as a unique opportunity for participation and contribution by all the actors and as a mechanism for meaningful change. The coalition offers a path to consensus on key content moderation issues – and opportunities to address them. The following table provides an overview of the potential risks related to the formation and functionality of the coalition, identified by the respondents, including potential ways to overcome and mitigate them.

Risk type*	Description of risk	Likelihood**	Impact***	Monitoring and mitigation
Finance	Sustainability and longevity of the coalition	Possible	**Severe**	• Joint funding initiatives of coalition members to support the work of the coalition • Joint funding applications and participation of coalition members in donor meetings and agenda-setting process
Political	Negative interference of state actors, and ethnic-political division with the work of the coalition	Likely	**Severe**	• State authorities should be informed about the work, scope, and members of the coalition and invited to join the session and open meetings as observers and on the consultation level. They should be regularly updated about the work of the coalition through meetings with individual authorities and joint newsletter, or similar initiatives. • Public and state authorities should be offered targeted training by the members of the coalition
Political	Legislative initiatives on different state levels (e.g. Law on Public Order and Peace) that can have a chilling effect on freedom of expression online	Likely	**Major**	• The coalition can assist and be consulted in regard to legislative initiatives, thus advocating for the values and objectives that potential future regulation in this field needs to ensure
Political	Institutional practices, court decisions, and regulatory plans (e.g. media literacy strategy that imposes a state-mandated online monitoring)	Unlikely	**Minor**	• Through their internal communication channels, the coalition should be aware of state, ethnic, and canton-led initiatives. They should issue statements or provide expertise to improve the institutional practices, upon request.

ARTICLE¹⁹

Political	Political events such as a 2022 election process, or enforcement of a ban of denial of genocide in Srebrenica that can raise societal tensions	Possible	**Moderate**	• The coalition should plan their activities and corresponding timeframes by considering a wider political landscape and use 'windows of opportunity' to make important gains
Reputation	The failure of the coalition would reflect badly on the reputation of ARTICLE 19 in the country, the EU, and UNESCO	Unlikely	**Minor**	• The coalition should be facilitated by local actors, so that the role of ARTICLE 19 and other international organisations is merely supportive, but not instructive and decisive
Stakeholder	The coalition is perceived as a project not as a platform	Possible	**Major**	• Inclusion of all relevant stakeholders from the beginning to ensure fair representation of all identified actors • Creating a sense of belonging and trust among members • Ensuring continual funding
Stakeholder	Lack of expertise and experience in this field	Unlikely	**Major**	• ARTICLE 19, regional and individual country experts are able to provide training and initial support
Stakeholder	Limited resources and capacities for cooperation of civil society organisations and media outlets due to other multi-project engagements	Possible	**Moderate**	• Strike the balance: involve a limited but critical number of experts per organisation that can take the knowledge generated in the coalition back to the organisations, but also ensure the effectiveness of the coalition
Stakeholder	Lack of previous experience in coordination of multi-stakeholder platforms	Likely	**Moderate**	• Start small and increase the agenda of the coalition after each success and ensure inclusion of all actors in an equal and transparent manner • Provide training on networking and coalition governance
Stakeholder	Potential conflicts and collision of interest, values, and viewpoints	Unlikely	**Minor**	• The coalition should be modelled and presented as a platform for dialogue that is able, through inclusion of all actors, to facilitate and resolve conflicts • Include in the process of creation conflict−resolution mechanisms
Stakeholder	Social media is not interested in being involved in discussions, and stakeholders in turn feel discouraged	Possible	**Major**	• In parallel with the coalition setting process, ARTICLE 19, UNESCO, and other actors need to engage with social media, especially Facebook whose interest and early engagement can focus the attention

ARTICLE¹⁹

				of stakeholders to be involved in the work of the coalition

Notes:

* The risk type is pre-classified in the following categories: Political, Safeguarding, Stakeholder, Finance, Compliance, Reputation, Other, Covid-19.

** The risk likelihood is presented on the scale: Unlikely, Possible, Likely, and Almost certain.

*** The risk impact is presented on the scale: Minor, Moderate, Major, and Severe.

ARTICLE¹⁹

Annex B: Potential members of the coalition

The following table provides an overview of an initial and open-ended list of organisations and actors that could take part in the creation and development of the coalition.

Organisation	Category	Portfolio
Activist/Pride BiH	Activist/civic educator	Through education, media relations, and active engagement with a group of like-minded people, she is an active citizen struggling to improve the rights of LGBTQI+, women's rights, and other marginalised communities. With a team of activists, they have organised a Pride March in BiH for the last three years. She is also a part of the InternNews team in the country.
Analiziraj.ba	Media organisation	Established in 2015 as a media watchdog and they produce high-quality media content and topic-oriented analysis.
BIRN BiH/Detektor.ba	Media organisation	Media outlet, established in 2003, that focuses on war crimes – terrorism, radicalisation, and extremism; reform of judiciary; and corruption.
BIRN Hub	Regional civil society organisation	One of the most recognised human rights and media freedom organisations that has established itself as a regional leader in this field. They also focus on digital rights and media freedom online. 'Digital rights clinic', an upcoming project, will include a series of research and capacity-building interventions to strengthen the expertise of media. Focus will be on the educative and multidimensional reportings about social justice, media freedom, and technology-related issues. Mapping digital rights violations is an ongoing monitoring project at the regional level (BIRD). BIRN regional offices also regularly published stories on the role of technology in democratic societies.
CIN BiH	Media outlet	In 2004, the first investigative journalists outlet in the Western Balkans was established. Today, they publish up to 50 investigative stories, have gained a high level of trust, and reached a wide audience.
Civil Rights Defenders	Civil society organisation	Human rights protection with a focus on freedom of expression. It operates at the regional level and provides support to media outlets (e.g. content production), regional meetups, and conferences for capacity-building.
Crvena (Red)	Civil society organisation	Established in 2010, this multidisciplinary radical feminist collective focuses on artistic, research, educational, and political practices covering topics such as social position of women, production and management of natural and social resources, political decision-making, everyday life, work and neighbourhood relations, and art and creativity.
Faculty of Philosophy, University of Banja Luka	Academia	
Faculty of Philosophy, University of East Sarajevo	Academia	

Faculty of Philosophy, University of Mostar	Academia	
Faculty of Philosophy, University of Tuzla	Academia	
Independent expert	Digital rights protection	Researcher and activist in the field of Internet politics, open tech and data, freedom of the media and information societies.
Institute for Social Science Research, Faculty of Political Science, University of Sarajevo	Academia	
Internet Society of BiH	Civil society organisation	A newly established organisation, mainly focusing on gender and digital divide, environmental protection and technology, and technology-facilitated support for people with disabilities.
Journalists' Association – BH Novinari	Media association	Grounding their work on three pillars – safety, labour laws, online and gender threat – they run a free legal aid helpline, offer journalists help through funds of solidarity, and provide educational programmes. Their advocacy strategy focuses on the changes of criminal law to increase the protection of journalists, together with the adoption of key missing media-specific laws.
Mediacentar Sarajevo	Media organisation	Supports development of independent and professional journalism through education, research, advocacy, consulting, media research, audio-visual production, and archiving work. They run the media-related news portals: Media.ba and Discrimination.ba
Post-Conflict Research Center	Civil society organisation	Peace-building and transitional justice research centre that covers an array of issue such as peace education, prevention of genocide, and post-conflict research. A large part of their work focuses on equipping young activists to engage in peace-building through multimedia and content production.
Press Council	Self-regulatory authority	See the stakeholders section.
Raskrinkavanje	Media outlet	An official third-party fact-checker for Facebook, established in 2017 as a part of Zašto ne? It focuses on public participation and political accountability of civil society organisations. So far, more than 10,000 media and other user-generated content has been debunked.
Researcher and feminist activist	Online media activism	Feminist thinking and humour as a synergy that can bring the new fresh air. It's All Witches is an interdisciplinary grassroots platform for identification and subversion of patriarchal culture and a healthy public segment that has a potential to bring change. She works as a consultant in the field of feminist thinking and political participation but is also as an activist in the field of art and intersectional lives.
Zašto ne?	Civil society organisation	See the stakeholders section.

Annex C: Interview sheet

Organisation	Category	Name	Position	Interview date
Activist/Pride BiH	Activist	Lejla Huremović	Project officer/InternNews	14 September 2021
Analiziraj.ba	Media organisation	Anonymous	n/a	30 August 2021
BIRN BiH	Media organisation	Denis Džidić	Director	27 August 2021
BIRN Hub	Regional civil society organisation	Sofija Todorović Aida Ajanović	Project manager Project manager	21 September 2021
CIN BiH	Investigative media organisation	Lejla Bičakčić	Director	27 August 2021
Civil Rights Defenders	Regional civil society organisation	Ena Bavčić	Project coordinator	2 September 2021
Communication Regulatory Authority	Independent Authority	Lea Čengić	Director of department for content and media and information literacy	15 September 2021
Crvena (Red)	Civil society organisation	Danijela Dugandžić Živanović	Director	31 August 2021
EU delegation to BiH	International organisation	H. Firuze Demir	Political adviser	6 October 2021
Faculty of Philosophy, University of Sarajevo	Academia	Mario Hibert	Associate professor	30 August 2021
Faculty of Philosophy, University of East Sarajevo	Academia	Vuk Vučetić	Professor	2 September 2021
Faculty of Political Science – Department of Communication and Journalism, University of Sarajevo	Academia	Lejla Turčilo	Professor	10 September 2021
Faculty of Political Science, University of Sarajevo	Academia	Emir Vajzović	Professor	10 September 2021
Independent expert	Activist	Fedja Kulenović	Assistant at the Faculty of Philosophy,	30 August 2021

ARTICLE¹⁹

			University of Sarajevo	
Independent expert	Adviser	Mak Kapetanović	Independent expert	29 August 2021
Internet Society of BiH	Civil society organisation	Haris Omergic	Project coordinator	29 August 2021
Mediacentar Sarajevo	Civil society organisation	Elvira Jukić-Mujkić, Anida Sokol	Editor-in chief Project coordinator	27 August 2021 2 September 2021
Post-Conflict Research Centre	Civil society organisation	Tatjana Milovanović	Director	17 September 2021
Press Council	Self-regulatory authority	Dženana Burek, Maidu Bahto Teskeredžić,	Director, and Project Coordinator	22 September 2021
Raskrinkavanje	Civil society organisation	Tijana Cvjetićanin Rašid Krupalija	Editor-in chief Editor	2 September 2021
Sve su to vještice/It's All Witches	Activist	Hana Ćurak	Feminist thinker	16 September 2021
UNESCO BiH office	International organisation	Siniša Šesun Melisa Durak	Head of office Project officer	27 August 2021

Bibliography

Legal texts

Official Gazette of Republika Srpska No. 11. Law on Public Order and Peace. Adopted on 17 February 2015.

United Nations, General Framework Agreement for Peace in Bosnia and Herzegovina (Dayton Agreement). Annex IV. Signed on 21 November 1995.

Cases

Council of Europe, European Court of Human Rights, *Sunday Times v. The United Kingdom*, 50/1990/241/312, 24 October 1991.

References

Chinmayi, C., 'Facebook's Faces', *Harvard Law Review Forum*, vol. 135 (2022).

Cohen, J., *Between Truth and Power: The Legal Constructions of Informational Capitalism* (Oxford University Press, 2019).

Couldry, N. and Mejias, A. U., *The Costs of Connection. How Data Is Colonizing Human Life and Appropriating It for Capitalism* (Stanford University Press, Redwood City, 2019).

Cvjetićanin, T., Dezinformacije u online sferi: slučaj BiH [Disinformation in the online space: case study BiH]. Zašto ne?, Sarajevo (2019).

Cvjetićanin, T., 'Disinformation in the election process'. Zašto ne?, Sarajevo (2020).

Helberger, N., 'The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power', *Digital Journalism*, 8 (2020), p. 843.

Keller, D., *Amplification and Its Discontents. Occasional Papers: Knight First Amendments Institute* (Columbia University, New York, 2021).

ARTICLE19

Klonick, K., 'The New Governors: The People, Rules, and Processes Governing Online Speech', *Harvard Law Review,* 1598 (2018).

Krupalija, R., *Disinformation during Covid-19 Pandemic*. Friedrich-Naumann-Stiftung für die Freiheit Bosnien-Herzegowina, Sarajevo (2020).

Livančić-Milić, B., *Disinformation during Covid-19 Pandemic: Regional research*. Friedrich-Naumann-Stiftung für die Freiheit Bosnien-Herzegowina, Sarajevo (2020).

Osmančević, E., Mapiranje medijskih veb portala u Bosni i Hercegovini [Mapping online media portals in BiH]. Centar za promociju civilnog društva, Sarajevo (2021).

Perkov, B., *Digital Rights Falter Amid Political and Social Unrest*. BIRN and Share Foundation, Belgrade (2021).

Ristić, M. et al., *From Cures to Curses, Manipulation Flourishes in the Digital Environment*. BIRN, Belgrade (2020).

Sokol, A., *Hate Narratives in the Media and User-Generated Content*. SEENPM, Tirana; Peace Institute, Ljubljana; and Foundation Mediacentar Sarajevo, Sarajevo (2020).

Sokol, A., *Polarized Public Trust in the Media and Social Networks in Bosnia and Herzegovina*. South East European Network for Professionalization of Media (SEENPM) and Mediacentar Sarajevo, Sarajevo (2020).

Sokol, A., *Propaganda, Disinformation and Hate Models of Media and Communication in Bosnia and Herzegovina*. SEENPM, Tirana; Peace Institute, Ljubljana; and Foundation Mediacentar Sarajevo, Sarajevo (2020).

Swift, J., *Legal Responses to Disinformation*. International Center for Not-for-Profit Law, Washington DC (2020).

Turčilo, L. and Buljubašić, B., Uništavanje reputacije na bh.nacin:govor mržnje botova u online prostoru kao sredstvo sužavanja javnog prostora za alternativna mišljenja u Bosni i Hercegovini [Destruction of the reputation in the BiH way: hate speech of both in the online

ARTICLE¹⁹

space as a mean to shrink public space for alternative opinions in BiH], presented at the conference Komentari, govor mržnje, dezinformacija i regulacija javne komunikacije, Agencija za elektronicke medije i Medijska istraživanja, Zagreb (2022) (forthcoming).

Zulejhić, E., Preko margine: dezinformacije o marginalizovaim grupama u bh.medijima [Over the margin: disinformation about marginalized groups in BiH media content]. Zašto ne?, Sarajevo (2020).

Reports

ARTICLE 19, *Side-stepping Rights: Regulating Speech by Contract, Policy Brief*, 2018.

ARTICLE 19, *Social Media Councils: One Piece in the Puzzle of Content Moderation*, 2021.

Bertelsmann Stiftung's Transformation Index (BTI), *BTI Country Report: Bosnia and Herzegovina* (2020).

Council of Europe and Communication Regulatory Authority, *Annual Report* (Sarajevo, 2020).

Council of Europe and Communication Regulatory Authority, *Media Habits of Adults in BiH* (Sarajevo, 2019).

European Commission, *Commission Staff Working Document, Bosnia and Herzegovina 2020 Report*. Accompanying the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions (SWD(2020) 350 final) (Brussels, 2020).

European Commission, *Commission Staff Working Document*, Bosnia and Herzegovina 2021 Report. Accompanying the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions (SWD(2021) 291 final) (Strasbourg, 2021).

IREX, *Media Sustainability Index 2019 Report: Bosnia and Herzegovina* (2019).

Javna Rasprava.ba, *Šta bi za građane i građanke FBiH značilo proglašavanje interneta javnim mjestom?* [What it would mean for citizens of BiH if Internet becomes a public space]. Zašto ne?, Sarajevo.

Media Plan Institute, *Istraživanje: Medijske navike, afiniteti i percepcija medija* [Research on media habits, preferences and mediaperception] (Sarajevo, 2019).

Policy Department for External Relations, *Mapping Fake News and Disinformation in the Western Balkans and Identifying Ways to Effectively Counter Them* (European Commission, Brussels, 2021).

World Bank, *Western Balkans Regular Economic Report No.19: Subdued Recovery* (World Bank, 2021).

Websites/blogs/news articles

BIRD, 'Bosnia trying to censor information about pandemic, journalists say', 9 April 2020, accessed 10 October 2020.

BIRN, 'SEE Digital Rights Network established', 14 August 2020, accessed 21 October 2021.

Blagovčanin, P., 'Stranački botovi i kontaminacija online prostora dezinformacijama' [Poltical parties bots and contamination of online public spaces with disinformation]. Mediacentar Online (25 November 2019), accessed 12 October 2021.

Civil Rights Defenders, 'Journalists' days to focus on Western Balkans media freedoms in light of Covid-19', 8 October 2020, accessed 21 October 2021.

Dojčinović, S., 'Hey, Mark Zuckerberg: my democracy isn't your laboratory', *The New York Times* (15 November 2017), accessed 18 January 2022.

Đorđević, N., 'Serbia's reality problem', Emerging Europe (14 November 2020), accessed 10 January 2022.

Jeremić, I. and Stojanovic, M., 'Facebook, Twitter struggling in fight against Balkan Content violations', Balkan Insight (16 February 2021), accessed 21 October 2021.

Jukić, E. M., 'Bosnian Serbs adopt 'Net Censorship' Law', Balkan Insight (6 February 2015), accessed 1 September 2020.

Karic, H., 'What Bosnian's new genocide denial ban means', Analysis (27 July 2021), accessed 8 September 2021.

Karović, A., 'Zastrašujući izostanak reakcije i neistine Avaza o prijetnjama Raskrinkavanju', Raskrinkavanje.ba (12 June 2021), accessed 21 October 2021.

Klix.ba, 'SDA osniva brigadu botova za "Internet rat" uoči izbora, najboljima obećali uhljebljenje' [SDA is creating a bot brigade to enter 'Internet rat' before the elections, the most successful bots will be granted a job], 27 April 2018, accessed 8 October 2021.

Kovačević, D., 'Bosnian Serbs to boycott state institutions over genocide denial ban', Balkan Insight (27 July 2021), accessed 8 September 2021.

Latal, S., 'Stakes could not be higher in Bosnian electoral reform bid', Balkan Insight (11 March 2020), accessed 5 September 2021.

Latal, S., 'Twenty-five years on: Bosnia remains caught between fear and hope', Balkan Insight (19 November 2020), accessed 5 September 2021.

O'Neil, C., 'Facebook's VIP "whitelist" reveals two big problems', *Bloomberg* (15 September 2021), accessed 18 October 2021.

Radio Slobodna Evropa, 'Nezaposlenost u BiH porasla za 2,96 posto' [Unemployment increased for 2.96 percent], 7 January 2021, accessed 29 September 2021.

Radio Slobodna Evropa, 'Svjetska banka: BiH će zahvatiti najveća recesija u posljednjih 25 godina' [World Bank: BiH is going to be caught by the highest recession in the last 25 years], 22 October 2020, accessed 29 September 2021.

ARTICLE¹⁹

Raskrinkavanje, 'Djelovanje stranačkih "botova" u komentarima na "Klixu"' [The mechanisms of political bots in the comment section on "Klix"] (5 August 2018), accessed 8 October 2021.

Ristić M. and Krivokapic, D., 'Digital rights falter amid political and social unrest', Balkan Insight (1 July 2021), accessed 10 October 2021.

Safejournalists.net, 'SJ: The Brutal Campaign against Lejla Turčilo and Borka Rudic from Sarajevo Must Stop', 15 September 2021, accessed 21 October 2021.

Safejournalists.net, 'Threats against the lives and limbs of journalists, Srđan Puhalo and Dragan Bursać, 03.11.2020', 16 December 2020, accessed 21 October 2021.

Share Foundation, 'Političko-informaciono ratovanje: kratko uputstvo' [Political-informational war: brief instruction book], 31 July 2017, accessed 21 September 2021.

Toe, R., 'Bosnians launch petition against Serbian reality show', Balkan Insight (7 December 2015), accessed 15 January 2022.

Turčilo, L., 'Komentari na novo javno mnjenje ili novo bojno polje' [Comment sections: new public debate or new mind field]. Analiziraj.ba (28 October 2020), accessed 13 October 2021.

Turčilo, L., 'U susret lokalnim izborima: Portali za diskreditaciju neistomišljenika i radikalizaciju birača' [Approaching new elections: Online portals for discretions of opponents and radicalisation of voters]. Analiziraj.ba (19 June 2020), accessed 12 October 2021.

Žuvela, M., 'Govor mržnje na stotinama portala u BiH niko ne kontroliše' [No one controls hate speech on hundreds of media portals], Radio Slobodna Evropa (15 June 2021), accessed 21 October 2021.

ARTICLE¹⁹

Other resources

Agency for Statistics of Bosnia and Herzegovina, 'Press Release: Demographic and Social Statistics No.7', accessed 20 September 2021.

Balkan Discourse Platforms, 'Ordinary Heroes', accessed 20 October 2021.

BH Novinari, accessed 15 January 2022.

Crvena, Umjetnost, accessed 21 October 2021.

Data Reportal, 'Digital 2020: Bosnia & Herzegovina', accessed 1 October 2021.

Edwards, K., 'European Ecommerce Overview: Bosnia and Herzegovina', Ecommerce Germany News, accessed 29 September 2021.

Facebook, 'Introducing journalists registration on Facebook', blog (30 June 2020), accessed 15 October 2021.

Facebook, 'Working to stop misinformation and false news', accessed 15 October 2021.

Facebook Community Standards, 'Hate Speech', accessed 15 October 2021.

Facebook Transparency Center, 'Bosnia and Herzegovina', accessed 21 October 2021.

Google, 'Google News Initiative', accessed 15 October 2021.

Istinomer, 'Bosna i Hercegovina', accessed 21 October 2021.

Journalists' Association, https://bhnovinari.ba/bs/category/mrezanovinarkiubih/, accessed 21 October 2021.

Office of the High Representative, Official website, accessed 5 September 2021.

Point Conference, Official website, accessed 21 October 2021.

Press Council, 'Stop hate speech', accessed 21 October 2021.

Press Council, JUFREX 2, accessed 21 October 2021.

Raskrinkavanje, 'Methodology', accessed 21 October 2021.

South East European Network for Professionalization of Media (SEENPM), 'Resilience News', accessed 21 October 2021.

Council of Europe, South-East Europe Freedom of Expression (JUFREX), accessed 21 October 2021.

Sve su to vještice, Official website, accessed 20 October 2021.

Twitter, 'Transparency report, Bosnia and Herzegovina', accessed 21 October 2021.

Twitter Community Standards, 'Hateful Conduct Policy', accessed 15 October 2021.

United Nations International Tribunal for the former Yugoslavia (ICTY), 'The Conflicts', accessed 15 January 2022.

YouTube Community Standards, 'Hate Speech Policy', accessed 15 October 2021.

ARTICLE 19

Endnotes

1 General Framework Agreement for Peace (Annex IV), also known as Dayton Agreement, signed on 21 November 1995.

2 Interview, Leila Bičakčić, Director of CIN-investigative media organisation, 27 August 2021; Denis Džidić, Director of BIRN media outlet – detektor.ba, 27 August 2021; Fedja Kulenović, independent expert and activist, 30 August 2021.

3 A most striking example of this deeply fragmented political system is a phenomenon called 'two schools under one roof' which refers to children who come from different entities, go to the same school, but attend separate classes and learn different school curricula, depending on their ethnic background. See Organization for Security and Co-operation in Europe (OSCE), '"Two Schools Under One Roof" – The Most Visible Example of Discrimination in Education in Bosnia and Herzegovina' (2018); S. Latal, 'Twenty-five years on: Bosnia remains caught between fear and hope', Balkan Insight (2020).

4 In practice, this means that the Criminal Code of Bosnia and Herzegovina will be amended to include this new provision which, in fact, prohibits anyone to publicly condone, deny, trivialize, or try to justify genocide, crimes against humanity, and other crimes in line with the international humanitarian law. For an extensive discussion, see H. Karic, 'What Bosnian's new genocide denial ban means', Analysis (2021).

5 Interview, activist, Lejla Huremović, 14 September 2021.

6 Interview, director of investigative journalistic outlet, Leila Bicakcic, 27 September 2021.

7 Agency for Statistics of Bosnia and Herzegovina, 'Press Release: Demographic and Social Statistics No. 7'.

8 Radio Slobodna Evropa, Nezaposlenost u BiH porasla za 2,96 posto (2021), accessed 29 September 2021. See also Bertelsmann Stiftung's Transformation Index (BTI) *BTI Country Report: Bosnia and Herzegovina* (BTI, 2020), p.20.

9 Interview, Leila Bičakčić, 27 August 2021.

10 Interview, Denis Džidić, 27 August 2021.

11 Council of Europe and Communication Regulatory Authority, *Annual Report* (Sarajevo, 2020), p.19; Agency for Statistics of Bosnia and Herzegovina.

12 Council of Europe and Communication Regulatory Authority, *Annual Report*, p.23.

13 Agency for Statistics of Bosnia and Herzegovina, Bosna i Hercegovina u brojevima (2020), p.106, 107.

[14] This data is calculated based on the potential number of people that marketers can reach using adverts on Facebook and Instagram. See Data Reportal, 'Digital Report 2020: Bosnia and Herzegovina'; and A data-aggregator Statista.com reports similar data: around 1.8 million Facebook users, and around 1 million Instagram users.

[15] Interview, Leila Bičakčić, 27 August 2021; Denis Džidić, 27 August 2021; Dženana Burek, executive director of the Press Council, 22 September 2021.

[16] See, for example, R. Toe, 'Bosnians launch petition against Serbian reality show', Balkan Insight (2015); N. Đorđević, 'Serbia's reality problem', Emerging Europe (2020), https://emerging-europe.com/after-hours/serbias-reality-problem/.

[17] Communication Regulatory Authority, Rezultati istraživanja o adekvatnoj informisanosti mladih u Bosni i Hercegovini o situaciji vezanoj za Covid-19 (Sarajevo, 2020), p.13.

[18] Another survey reported that 64% of all citizens in BiH think that politicians and political parties have the largest influence over the media and that the media are fragmented along 'ethno-national lines', See Sokol, *Polarized Public Trust in the Media*, p.4, 11; Interviews, Tijana Cvjetićanin, 2 September 2021; Interview, Leila Bičakčić, 27 August 2021; Interview Ena Bavčić, 2 September 2021.

[19] Interview, Leila Bičakčić, 27 August 2021.

[20] Interview, Ena Bavčić, Civil Rights Defenders Representative for BiH, 2 September 2021.

[21] This has been repeatedly mentioned by the respondents and it also became clear in the early stage of the research process that there is no one focal organisation that addresses Internet freedom and digital rights broadly.

[22] Interview, Tijana Cvjetićanin, 2 September 2021.

[23] Lejla Huremović, an LGBTQI+ and feminist activist, also mentioned that there is a rather limited number of people who engage around this or any other social justice and intersectional issues (interview, 14 September 2021).

[24] Interviews, Leila Bičakčić, 27 August 2021; Tijana Cvjetićanin, 2 September 2021; Elvira Jukić-Mujkić, 27 August 2021; Sokol, *Polarized Public Trust in the Media*, p.19.

[25] Official Gazette of Republika Srpska No. 11, Law on Public Order and Peace, adopted on 17 February 2015, art.7.

[26] Javna Rasprava.ba, Šta bi za građane i građanke FBiH značilo proglašavanje interneta javnim mjestom?, Zašto ne?, Sarajevo.

27 Interview, Sofija Todorović, project manager at BIRN HuB, 21 September 2021.

28 In 2017, Facebook experimented with their content curation rules in several selected countries: Guatemala, Slovakia, Bolivia, Cambodia, and Serbia. This experiment promoted the visibility of content shared by friends over content and posts made by 'pages', including media organisations. In practice, it meant that these posts had been removed from the regular News Feed and that the content was available in a separate section called Explore Feed that users have to select before they can access. As a result, the visits to the pages of media outlets immensely decreased. In his response to Facebook, the editor-in-chief of the investigative media outlet Crime and corruption network, Stevan Dojcinovic, wrote: "By picking small countries with shaky democratic institutions to be experimental subjects, it is showing a cynical lack of concern for how its decisions affect the most vulnerable." See S. Dojcinovic, 'Hey, Mark Zuckerberg: My Democracy Isn't Your Laboratory', *New York Times*, 15 November 2017.

29 Interview, Fedja Kulenović, 30 September 2021; also mentioned by an executive director of the Press Council, Dženana Burek, 22 September 2021.

30 Interview, Tijana Cvjetićanin, 2 September 2021.

31 Facebook's Community standards prohibit hate-speech that 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, see Facebook's latest Community Standards. Similarly, see YouTube's Hate Speech Policy, and Twitter's Hateful Conduct Policy.

32 The reports base their findings completely on the survey results so they do not present official social media companies' data and analysis.

33 Interview, Tijana Cvjetićanin, 2 September 2021. See also Jeremić et al., Facebook, Twitter Struggling.

34 Interview, Tijana Cvjetićanin, 2 September 2021.

35 Interviews, Tijana Cvjetićanin, 2 September 2021; Leila Bičakčić, 27 August 2021.

36 Facebook enabled an option to download the country report but the presented data is not human-readable.

37 Interview, Tijana Cvjetićanin, 2 September 2021.

38 Interview, Maida Bahto Kestendžić, project coordinator at the Press Council, 22 September 2021.

39 This is discussed extensively in the context of Facebook in A. Chinmayi, 'Facebook's Faces', *Harvard Law Review Forum,* vol. 135 (2022), p.8, 9. See also C. O'Neil, 'Facebook's vip "whitelist" reveals two big problems', *Bloomberg* (2021).

40 For instance, a Pakistani blasphemy law is 'infused' in the content moderation system of Facebook in Pakistan, but outside Pakistan this content is still available. See Chinmayi, 'Facebook's Faces'.

41 Interviews, Tijana Cvjetićanin, 1 September 2021; Ena Bavčic, 2 September 2021.

42 Interview, Lejla Turčilo, 10 September 2021.

43 L. Turčilo et al., 'Uništavanje reputacije na bh.nacin:govor mržnje botova u online prostoru kao sredstvo sužavanja javnog prostora za alternativna mišljenja u Bosni i Hercegovini [Destruction of the reputation on bh way: hate speech of both in the online space as a mean to shrink public space for alternative opinions in BiH]', presented at the conference Komentari, govor mržnje, dezinformacija i regulacija javne komunikacije, Agencija za elektronicke medije i Medijska istraživanja, Zagreb (2022) (forthcoming), p.20. See also L. Turčilo, Komentari na novo javno mnjenje ili novo bojno polje, Analiziraj.ba (2020).

44 Raskrinkavanje, 'Djelovanje stranačkih "botova" u komentarima na "Klixu"' (2018). See also P. Blagovčanin, Stranački botovi i kontaminacija online prostora dezinformacijama, Mediacentar Online (2019).

45 Cvjetićanin, Dezinformacije u online sferi, p.7. See also: A. Sokol, *Hate Narratives in the Media and User-Generated Content* (SEENPM, Tirana; Peace Institute, Ljubljana; and Foundation Mediacentar Sarajevo, 2020), p.20.

46 Online media in the context of this report are stand-alone websites and platforms that produce news and similar content. *Note*: there is no information about the exact number of online media portals. According to the study *Mapping of media web portals in Bosnia and Herzegovina*, there are 306 online media portals, of which 44% do not have impressum available on their websites; 41% have active comment sections; and 95% have pages on social media.

47 See also IREX, *Media Sustainability Index 2019*, p.4.

48 'The actors targeted with disinformation are those persons or entities presented in a negative light, in a misleading or manipulative way; the beneficiaries of disinformation are those who are portrayed in a positive light in those same media manipulations.' Cvjetićanin, Dezinformacije u online sferi, p.40. See also IREX, *Media Sustainability Index 2019*, p.3.

49 T. Cvjetićanin et al., 'Disinformation in the election process', Zaštno ne?, Sarajevo (2020), p.31.

50 Research team, *Istraživanje*, p.33–55, 73, 77. See also A. Sokol, *Propaganda, Disinformation and Hate Models of Media and Communication in Bosnia and Herzegovina* (SEENPM, Tirana, Peace Institute, Ljubljana and Foundation Mediacentar Sarajevo, 2020), p.7; L. Turčilo, U susret lokalnim izborima: portali za diskreditaciju neistomisljenika i radikalizaciju birača, Analiziraj.ba (2020).

[51] Turčilo et al., Uništavanje reputacije. In this article, authors conducted interviews with 17 media actors (journalists, editors, media experts, academics) that confirmed the existence of a large number of bots online, fake profiles, frequent assaults in the comment sections, explicit hate speech, or that groups of people were singled out by those in control bots as those whose online activism can threaten their current political positions and the reputation of political parties they are affiliated with (p.19).

[52] Ibid., p.5.

[53] Interview, Lejla Turčilo, 10 September 2021. See Safejournalists.net, 'SJ: The brutal campaign against Lejla Turčilo and Borka Rudic from Sarajevo must stop', 15 September 2021. See also S. Hodžić et al., Surfanje po tankom ledu:mladi, mediji i problematican sadrzaja (Mediacentar Sarajevo, 2019), p.22.

[54] Interview, Dženana Burek, 22 September 2021.

[55] Safejournalists.net, 'Threats against the lives and limbs of journalists, Srđan Puhalo and Dragan Bursać' Safejournalists.net, 3 November 2020, accessed 21 October 2021. See also Jeremić et al., Facebook, Twitter Struggling.

[56] Interview, Lejla Turčilo, 10 October 2021. See also Turčilo et al., Uništavanje reputacije, p.13.

[57] Interview Fedja Kulenović, 30 August 2021. See also IREX, Media Sustainability Index 2019, p.7.

[58] See, for example, the case study on a hate campaign on a journalist of the Croat media (Sokol, Hate Narratives), p.12–14.

[59] In this report, disinformation refers to 'false or manipulated information that is knowingly shared to cause harm or is made with reckless disregard of likely harm.' Jonathan Swift, Legal Responses to Disinformation. (International Center for Not-for-Profit Law, Washington DC, 2020), p.3.

[60] Ibid., p.52.

[61] Interview, Tijana Cvjetićanin, 2 September 2021.

[62] R. Krupalija, Disinformation during Covid-19 Pandemic (Friedrich-Naumann-Stiftung für die Freiheit Bosnien-Herzegowina, Sarajevo, 2020), p.20. See also Policy Department for External Politics, Mapping Fake News and Disinformation, p.23.

[63] Over the course of nine months in 2020, a total of 2,662 pieces of media content was assessed and published by 773 media outlets and on three social media platforms compared to 2,420 articles and 752 media over the course of one year (2018). See Krupalija Disinformation during Covid-19 Pandemic, p.20

[64] Interview, Tijana Cvjetićanin, 2 September 2021.

ARTICLE[19]

[65] Krupalija, *Disinformation during Covid-19 Pandemic*, p. 20. See also: European Commission, *Policy Department for External Politics, Mapping Fake News and Disinformation in the Western Balkans and Identifying Ways to Effectively Counter Them*, Brussels (2021), p.25, 46.

[66] Krupalija, *Disinformation during Covid-19 Pandemic,* p.48.

[67] As described in the study: 'An unnamed man from Serbia, who claimed that the cure for COVID-19 had existed in Russia for years, was a viral hit in the region, collecting over 175,000 views on just one Facebook page (Glas Srbije) and 300,000 more on YouTube channels Da Best and Ne damo svetinje: it has been shared more than 100,000 times and republished by dozens of online portals throughout the region.' A woman named Aneta Krstović (nicknamed 'LaMadrina'), who claimed that salt was a cure for COVID-19, had her video shared on Facebook more than 13,000 times. Ibid., p.25, 47.

[68] Especially prominent was 'Semir Osmanagić, best known for his pseudo-archaeological project of the "pyramids" in central Bosnia. Osmanagić was active on various social networks before the pandemic. In February 2020, Osmanagić started talking about the pandemic, using sensational titles in his videos and repeating a number of already circulating falsehoods and conspiracy theories. His YouTube channel has 43,900 subscribers and 6,950,460 views, most of which came from pandemic-related videos', Ibid., p.47.

[69] Interview, Tijana Cvjetićanin, 2 September 2021.

[70] M. Ristić et al., *From Cures to Curses, Manipulation Flourishes in the Digital Environment* (BIRN, Belgrade, 2020), p.14.

[71] Policy Department for External Politics, *Mapping Fake News and Disinformation*, p.23, Similarly the authors conclude that 'Where a country's ability to govern – or, indeed, its ability to maintain sovereignty – are bound up in geopolitical conflicts between larger forces, domestic political agendas are especially vulnerable to distortion and/or capture by transnational disinformation narratives, [...] particularly in BiH', p.49.

[72] Interview, Danijela Dugandžic Živanović, 31 August 2021.

[73] E. Zulejhić et al., Preko margine: dezinformacije o marginalizovaim grupama u bh.medijima, Zašto ne?, Sarajevo (2020), p.40, 53–54, 56–57. See Raskrinkavamje's other fact-checking analysis, accessed 21 October 2021.

[74] Zulejhić et al., Preko margine, p.38.

[75] Interview, Lejla Huremović, 14 September 2021.

[76] The author of this report would like to thank these young people for sharing their experience by email.

[77] Interview, Leila Bičakčić, 27 August 2021.

ARTICLE¹⁹

⁷⁸ Interview, Hana Ćurak, 16 September 2021.

⁷⁹ Ibid.

⁸⁰ Ibid.

⁸¹ Ibid.

⁸² Interview, Lejla Turčilo, 10 October 2021.

⁸³ Interview, Hana Ćurak, 16 September 2021.

⁸⁴ Interview, Tatjana Milovanović, 17 September 2021.

⁸⁵ Ibid.

⁸⁶ Interview, Tatjana Milovanović, 17 September 2021.

⁸⁷ Ibid.

⁸⁸ Ibid.

⁸⁹ Policy Department for External Politics, *Mapping Fake News and Disinformation*, p.23. See also IREX, *Media Sustainability Index 2019*, p.9.

⁹⁰ The Republika Srpska public broadcaster published the largest number of political disinformation according to the study conducting by Raskrinkavanje platforms in 2018. See Cvjeticanin, Dezinformacije u online sferi, p.32; Sokol, *Polarized Public Trust in the Media*, p.13, 15; and IREX, *Media Sustainability Index 2019*, p.5.

⁹¹ Interview, Ena Bavčić, 2 September 2021; IREX, *Media Sustainability Index 2019*, p.8.

⁹² Confirmed by all respondents with media expertise. See also Hodzic et al., Surfanje po tankom ledu, p.26, 38; Sokol, *Propaganda, Disinformation and Hate Models*, p.8.

⁹³ Council of Europe: European Court of Human Rights, *Sunday Times v. The United Kingdom*, 50/1990/241/312, 24 October 1991, para.51.

⁹⁴ Interview, Leila Bičakčić, 27 August 2021.

⁹⁵ Interview, Denis Džidić, 27 August 2021.

⁹⁶ Interview, Anonymous, 30 August 2021.

⁹⁷ Interview, Anonymous, 30 August 2021.

⁹⁸ Interview, Denis Džidić, 27 August 2021.

[99] Interview, Dženana Burek, executive director of Press Council, 26 September 2021.

[100] Interviews, Leila Bičakčić, 27 August 2021; Elvira Jukić-Mujkić, 27 August 2021. Sokol, *Polarized Public Trust in the Media*, p.19.

[101] Interview, Anonymous, 30 August 2021.

[102] See, for example, D. Keller, *Amplification and Its Discontents, Occasional Papers* (Knight First Amendments Institute, Columbia University, 2021).

[103] ARTICLE 19's views on the regulation of platforms are presented in 'Side-stepping rights'; 'Watching the watchmen: Content moderation, governance and freedom of expression' (2021); and 'Taming big tech' (2021). See also the Santa Clara Principles (version 2.0).

[104] During the elaboration of this report, the researcher took part in online meetings and stakeholders workshops with Facebook organised by UNESCO within the context of the Social Media 4 Peace project.

[105] There is another important feminist network: CURE. Several attempts were made to establish connection with them but without success.

[106] Interview, Tijana Cvjetićanin, 2 September 2021.

[107] Interview, Denis Džidić, 27 September 2021.

[108] Social media platforms and networked information technologies have transformed information flows, media content distribution logic, and capabilities for participation in public debate, social and political life. Social media platforms have elicited hybrid space for individual interaction and collective deliberation, but they have also reinforced technological convergence and platformisation of the Internet and networked technologies. For more on this, see J. Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (Oxford University Press, 2019), p.75. For more on human autonomy in the context of platformisation, see N. Couldry and A. U. Mejias, *The Costs of Connection. How Data Is Colonizing Human Life and Appropriating It for Capitalism* (Stanford University Press, Redwood City, 2019), p.153.

[109] Interview, Leila Bičakčić, 27 August 2021.

[110] Interview, Elvira Jukić-Mujkić, 27 August 2021.

[111] Interview, Denis Džidić, 27 August 2021.

[112] At least to the knowledge of research, the only similar initiative from this domain was an Internet forum conference, 'One World', which brought together all key actors, including state actors; however, this platform ceased to exist. See One World and Global Information Society Watch, accessed 12 October 2021. Attempts were made to establish contacts with this organisation but without success.

ARTICLE 19

---

113 Interview, Elvira Jukić-Mujkić, 27 August 2021.

114 Interview, Anonymous, 30 August 2021.

115 These focus areas are mentioned by responses in the interview process.

116 As mentioned above, ARTICLE 19 has warned against possible risks linked to a national establishment of social media companies; however, a stable online point of contact for local stakeholders could contribute to a sustainable and more transparent engagement of local civil society organisations with social media companies.