MONITORING METHODOLOGY
INSTRUCTIONS FOR MONITORS

ADAPTED VERSION (2022)

HATE SPEECH IN THE WESTERN BALKANS

MONITORING GOAL

The goal of this monitoring is to identify and document the most significant cases ('incidents') of hate speech and divisive discourses in the media.

Incidents are defined by the speaker regardless of where they occurred, but monitoring refers to their media coverage. They can originate in the Parliament, party event or any public occasion, not necessarily being generated by the media, but have to be reported or spread out by the media.

The media can be traditional (television, newspaper etc.) or social media (Facebook, Twitter, Instagram, TikTok etc.) or combination of both.

The purpose of the monitoring is not systemic overview of all the media in the country, overall media discourse, nor the content of selected media. The goal is to single out individual incidents of hate speech which are either reported or distributed by the media, investigate frequency of their occurrences and their major forms. Monitoring will identify who commits the incident but also how media transmit, amplify or critically counter the hate speech through their coverage.

Equally important goal is to identify groups and individuals who are targeted in those incidents, what kind of hatred, insults and threats are they exposed to and what language is publicly used against them in various media settings and environments.

Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

'INCIDENT' DEFINITION AND SIGNIFICANCE

The monitoring focuses on the most significant 'episodes' of hate speech which can be identified as 'incidents'. There are three criteria to select them:

First, they can be defined by the relevance of the person making the offence. State officials, political representatives, public personalities, people of authority or celebrities carry more weight, and their speech is more consequential for the public opinion. Therefore, the importance of the speaker determines the tone of public debate beyond the single incident.

Second, the incidents appearing across the media, either by being reported in many media or repeated though prolonged periods of time. Cases of viral hate speech, incidents that travel across media platforms or those with extended presence in certain media both have magnifying effects and therefore carry more significance than a random incident.

Third, public perception of the incident, its consequences, influence, harm, potential to cause chain reaction, be reprinted or become viral, adds to its significance even if the speaker would otherwise have been unnoticed. Due to rising importance of social media and its networking potential, certain incidents capture public attention and stir public sentiments by their inflammatory, intimidating, and discriminatory nature amplified through comments and sharing.

MAJOR FORMS OF HATE SPEECH

There are many forms of expressing hatred in the public. This monitoring distinguishes three major groups of them: (1) hateful and offensive speech (2) fake news, mal/disinformation and (3) inflammatory speech.

First, **hateful and offensive speech** also includes different forms of expressing intimidation, insult or humiliation, which is difficult to fit into a single definition. But all those statements share several common features such as: "they are targeting a group or individual as a member of a group, content of the message expresses hatred, the speech causes harm… the speaker intends a bad action beyond the speech… they are public …and the content makes violent response possible".

Therefore the monitoring focuses on the:

(1) Negative collective labelling, attributing negative qualities associated with a group by negative stereotyping, hostile language or qualifications addressing the whole group and assuming that each member of the group has the same negative attributes.

Co-funded by the European Union
B | T | D The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND
Norway
Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

(2) Discriminatory, harassing, offensive, denigrating, humiliating speech directed to a person or a group. Explicit verbal harassment and openly offensive or humiliating attacks which are causing harm

(3) Incitement to violence, open call to violence, or justification of violent action against a group or individual. In particularly polarized societies or divisive social situations this can be expressed through various metaphorically or culturally specific forms of speech.

Also specific forms of 'humour' or satire should be carefully considered. They should not be ignored because of the 'artistic' nature if blatant and negatively stereotyped to clearly humiliate individuals or groups. Common sense negative stereotyping can be used to normalize offensive and hatful speech through 'everyday jokes'.

The second group, colloquially referred to as Fake news, false or misleading information which form contemporary 'information disorder' (source) includes following varieties that often lead to hate speech:

- Mallinformation, entirely or partly accurate information but used in a malicious context or intentionally placed to harm the person, discredit or humiliate it. It can range from sex types, intimate details, private data obtained without permission etc.
- Disinformation, on the contrary is not accurate and is used with malicious, harmful purpose, intended to manipulate, disinform, mislead the public and cause harm.
- Fake news family also include misinformation, not accurate but not intentionally harmful information, honest mistakes that can be corrected and usually not resulting in offensive speech)

Inflammatory speech is singled out as a third group because of intensity and spread where its devastating nature results from the content but more so from the contextual nature of its use. It refers to repeated media offence by different actors or a prolonged hate speech by the same media or actors in a conflict, or potentially divisive situation. In such conditions or conflicting social environment, it can stimulate, incite or directly contribute to discriminatory or violent behaviour against involved actors.

SPEAKERS AND MEDIA IDENTIFICATION

The monitoring is identifying who commits the incident. Identifying speakers by their public roles will help understand who generates the hate speech, what kind of public actors are mostly behind this kind of verbally aggressive and harmful public narratives. It is one of the major monitoring goals to see the role of public officials, political actors or celebrates in the production of harmful public speech compared to other participants in the public life.

The monitoring goal is also to identify where the incidents appear and how they spread throughout the networked media environment. When the hate speech is covered as news or later reported in other media its trajectory over different media platforms will reveal how it is shared between old and new media,

Co-funded by the European Union
The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND
Norway
Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

between online and offline environments. If or when the incidents become viral, the monitoring will trace its original appearance/ original quote, check out the tracking of the page and follow the reactions thread.

Following traditional media web pages, Facebook or Twitter accounts when they cause reaction and stimulate traffic, comments or likes connected to the incident helps understand the movement and amplification of hate speech incidents in the networked media environment.

The monitoring will include certain number of mainstream media for regular observation in each country. They need to reflect the media range in terms of media types, ownership, editorial policy and audience reach. For most countries it will be sufficient to regularly monitor six mainstream outlets: two TV channels, two major print media and two online media. TV channels will be a public service broadcaster and the major commercial TV and their news content will be followed major news programmes and current affairs political shows.

Focusing on the incident- actor- media trajectory it will help develop additional alert system for hate speech cases and bring the balance between incident oriented and regular full-time monitoring into a more harmonious overview.

INCIDENT REPORT FORM

Overall the Incident report form is structured to provide three types of data.

The first group are general identifiers and include reference number, date of publication, country of origin, URL location and monitor's identification.

The second group of data is about the content of hate speech incidents, and it includes types of hate speech which is reported on, what group or individuals it is against, what kind of hatful language is used against them including quotes to illustrate that, and who commits those incidents.

The third group of data refers to the media treatment of the incident and includes the medium where incident is covered, headline, brief description and the context of the covered incident including the reactions it potentially triggers in the larger media environment.

Clean data from the Incident report form will be inserted in Excel data sheet for further processing.

Based on monitoring findings regular yearly reporting will offer description of each country situation, including significant transformations and comparative regional insights.

Co-funded by the European Union
The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND
Norway
Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

HATE SPEECH IN THE WESTERN BALKANS

INCIDENT REPORT FORM

REFERENCE NUMBER

DATE OF PUBLICATON

COUNTRY OF THE INCIDENT

- Albania
- Bosnia
- Kosovo
- Montenegro
- North Macedonia
- Serbia

WHAT KIND OF HATE SPEECH ARE YOU REPORTING ON

- Against Religion (Anti-Semitism, Islamophobia, Anti-Christian)
- Against Gender (Sexism, Sexual harassment, Misogyny)
- Against Sexual Minority (Homophobia)
- Against Ethnicity (Ethnic discrimination, Racism, Xenophobia)
- Against Migrants / Refugees
- Against people with Disabilities or Illnesses
- Against Journalists
- Against Political / Ideological opponents
- Against certain Professions
- Other (Physical appearance, Victims of war)

WHAT GROUP OR INDIVIDUAL WAS THE INCIDENT AGAINST

PERSONALISATION (how was the group or individual named)

WHAT TYPE OF FIGURE COMMITED THE INCIDENT

- Politician, political party, state official
- CSO, NGO or other civil society organization
- Journalist, media personnel, media writer/ analyst

Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

Celebrity, Artist, Popular Culture person
Other type of public figure, Professor, Intellectual
Influencer, blogger, Social media activist
Private person
Other

WHAT TYPE OF CONTENT YOU ARE REPORTING ON

Negative group labelling, stereotyping, hostility
Insult (personal, denigrating, humiliating)
Spreading of harmful lies, misinformation, disinformation
Misuse of personal data, half- truths, leaked information from state records
Threat, Statements potentially threatening to safety
Incitement to violence
Inflammatory speech (conflict situation, repeated messages from different actors, prolonged by the same media)

QUOTE

SENTIMENT ANALYSIS

1 – Disagreement
2 - Negative actions
3 - Negative character
4 - Demonizing and dehumanizing
5 – Violence
6 – Death
X - It is impossible to determine

URL

HEADLINE (original language first, then translation)

BRIEF DESCRIPTION

CONTEXT

Co-funded by the European Union
The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND
Norway
Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

WHAT TYPE OF MEDIA WAS THE INCIDENT IDENTIFIED IN

- Radio
- Television
- Newspaper
- Other traditional media
- Info Portal
- Facebook page
- Twitter
- Tiktok
- Instagram
- Other social media
- Other

WHAT WAS THE REACTION TO THE INCIDENT

MONITOR'S NOTE

Co-funded by the European Union

The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND

Norway

Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

INSTRUCTION FOR MONITORS

REFERENCE NUMBER

Each incident is identified by a number. Numeration starts with three digits (or two if enough) to leave sufficient room for all incidents in the monitoring cycle. The number is preceded with coder's initials (e.g. SM001… MP01). This is easy for data processing and for individual access to each incident.

DATE OF PUBLICATON

This refers to the date of the incident's first appearance in the media, or the original appearance of the incident when it gets viral or has multiple coverage. (dd. mm. yyyy)

COUNTRY OF THE INCIDENT

The country is coded for each incident by assigned name.

WHAT KIND OF HATE SPEECH ARE YOU REPORTING ON

The incident will be classified in one of five major categories based upon the nature and attributes of the offensive speech which are the most common in the Western Balkans media.

The incident can consist of offensive treatment or statements against religion, gender, sexual minorities, ethnicity, refuges or migrants. Within each of these large groupings there are variations for monitors to recognize. But the purpose of this large classification is to diagnose the frequency and incidence of them and to map out social groups (or phenomena in general e.g. religion, gender, sexual orientation etc.) that are most often targeted in harmful and hateful incidents.

In cases with multiple kinds of hatred expressed, monitors should identify the central intent or the 'main target' of the speaker or the message and classify it as such. Other significant forms or hatred, or their particular combination of them, should be indicated in the 'Monitor's note'.

If none of these applies the incident can be classified as 'other' again with explanation in the 'Monitor's note'. If numerous, these exceptions can later be categorized and added to the Incident report from or further explained in final project reporting.

Co-funded by the European Union

The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND

Norway

Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

WHAT GROUP OR INDIVIDUAL WAS THE INCIDENT AGAINST

Previously identified kind of hate speech can also be expressed as general negative framing of particular phenomenon (e.g. gender equality, migration, political ideology etc.) without identifying a particular group. Monitors need to note that the "Other" option in the total sample should not exceed five percent. If that percentage is exceeded, subsequent classification is necessary.

But if the group or individual representing the group is mentioned the monitor should identify it by using its conventional name (ethnic group, migrants, all women, gay community etc.) not the name used in the incident.

This is important to capture for further analysis of hate speech against various groups.

PERSONALISATION

Personalization further specifies whether the incident is against certain individual who is selected to personify the 'target' of the hate speech or a group in general.

The monitor needs to repeat exact words which the speaker used for labelling (naming) in his statement. Repeating exact wording of the speaker (e.g. syntagm used to name the group, discriminatory labelling of the person etc.) only refers to the explicit description or naming of the actor, not to the contextual information.

WHAT TYPE OF FIGURE COMMITED THE INCIDENT

Person who committed the incident should be possible to identify within those listed social roles. If none of the above applies it can be coded under 'other' and additionally explained in the Monitor's notes.

WHAT TYPE OF CONTENT YOU ARE REPORTING ON

The type of content needs to be classified in one of the seven offered varieties. If there are more than one, the most offensive or the one which is causing most public reactions should be coded and others indicated in the Monitor's notes.

If the type of content cannot be recognised within the offered varieties it should be listed as 'other' but in the Monitor's note classified in one of the three broad categories (hateful or offensive speech, fake news, inflammatory speech).

Co-funded by the European Union
The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND
Norway
Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

QUOTE

The most significant quote by the person who committed the incident which clearly illustrates the kind of hate speech or the type of content that is being reported.

SENTIMENT ANALYSIS

Within sentiment analysis a score for each case should be given, except in situations when it is not feasible. The score ranges from 1 to 6, as shown in the table, according to the methodology of George Washington University. Pay attention: Not only can units of vocabulary have a hateful and non-hateful context, but language can be structured to communicate hateful context using sarcasm, double entendre, innuendo, euphemism, metaphor, and other forms of rhetorical obfuscation.

1	Disagreement	Rhetoric includes disagreeing at the idea/mental level. Challenging groups claims, ideas, beliefs, or trying to change them.	False, incorrect, wrong.
2	Negative actions	Rhetoric includes negative nonviolent actions associated with the group.	Threaten, stop, outrageous behaviour, poor treatment, alienate, hope for their defeat
3	Negative character	Rhetoric includes non-violent characterizations and insults.	Stupid, aggressor, fake, crazy
4	Demonizing and dehumanizing	Rhetoric includes sub-human and superhuman characteristics.	Alien, demon, monkey, Nazi, cancer, monster, germ
5	Violence	Rhetoric implies infliction of physical harm or metaphoric/ aspirational physical harm. Responses include literal violence or metaphoric/ aspirational physical harm.	Hurt, rape, starve, torturing, mugging
6	Death	Rhetoric implies literal killing. Responses include the literal death/elimination of a group.	Kill, annihilate, destroy

URL

HEADLINE

Original language first, then translation of the headline.

Co-funded by the European Union

The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND

Norway

Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

BRIEF DESCRIPTION

Essential description of the media report of the incident. The description briefly explains media coverage - placement, length, framing, frequency - or whatever is relevant to indicate how media treated the incident.

CONTEXT

Include crucial information needed to understand the incident. The Context section should give enough information for people not familiar with the country context to understand the incident. This is particularly important if the understanding of the incident depends on understanding the broader country or region related situation. Also, in case the incidents database is open for review by the individuals outside of the team, the information should be easily understandable for diverse audiences.

WHAT TYPE OF MEDIA WAS THE INCIDENT IDENTIFIED IN

If possible, the medium of the incident's first appearance should be identified. If the hate speech is news-worthy, i.e. later reported in other media, shared or retweeted the coding will start at the original page/medium. Even if the incident becomes viral, the quote should come from the original media source. It could be traced back by tracking of the page and follow the reactions thread. If the incident occurred within event reported by many media (Parliamentary session, formal occasion) the Monitor's note should indicate how widely it was reported. Also, if a different source significantly contributed to the visibility of the incident (i.e. original source is a social media page with small number of followers, but was picked up by the influencer or the media outlet which significantly contributed to the visibility) then this should be noted in the Monitor's note because it can significantly contribute to the reaction to the incident.

WHAT WAS THE REACTION TO THE INCIDENT

For incidents with significant online presence the media trail (number of likes, shares, retweets, covered in other media) should be reconstructed and documented. For cases recorded on websites, Crowdtangle app should be used. It provides free and easy way to see how many times a link has been shared and who shared it. It shows the aggregate share counts, as well as the specific Facebook Page posts, Tweets and Subreddits that shared a URL. In order to use this app you should download the Chrome extension. In this regard, please note that this application only works in Google Chrome, and not in other browsers. When you install the application, a small icon will appear in the upper right corner of the browser. Once you've selected content published by the online media / portals, go to each content and click the "Crowdtangle" icon. The drop-down menu will list the total number of interactions, as well as the names of the pages that shared the content. To view the results, you must have your Facebook and Twitter account. From a user perspective, you have no reason to worry, as your friends on these social networks will not see the Crowdtangle activities.

MONITOR'S NOTE

If possible, it is advisable to explain (sub)narrative type related to the target group in the analysed example. A narrative can be defined as a logical, internally coherent report and interpretation of

Co-funded by the European Union
B | T | D The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND
Norway
Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.

connected events and characters or pieces of information that makes sense to the reader/listener. Examples of sub-narratives related to migrants:

- Migrants are a threat to public health. Migrants are contagious (infected) and unclean. They bring infectious diseases with them (in the case of COVID-19 they represent the biggest threat for spreading the pandemic) and endanger the health of the population in the countries they are located in / arriving to (countries of destination / transit countries; countries that they travel through / countries they stay to reside in);
- Migrants are a threat to the core values of the society to which they are arriving. Migrants come from countries that do not respect fundamental human rights;
- Migrants represent a threat to the economic system of the society. The countries should take care of their unemployed citizens instead of allocating public funds to cover the expenses of handling of migration. Migrants are a cause of worsening of the economic position of the local population ("stealing their jobs");
- Migrants are a threat to the social security (welfare) system of the state of arrival. Migrants are lazy, they do not want to work and they only come to western countries to exploit the welfare system;
- Migrants represent a threat to cultural values of the society. Migrants come from countries with completely different, alien cultural values, incompatible with those of the society of arrival. Migrants are barbarians (under-developed / backwards), coming to the civilized western world;
- Migrants are potential terrorists;
- Migrants are a threat to population growth. Migrant families have many children. In the long term this natality policy will cause for the white people to become a minority;
- Migrants are poor and uneducated and can not contribute to the society;
- Migrants are not prepared to adjust to the environment of their arrival. They enforce their traditions, culture and values upon the local population;
- Being uncivilized, migrants are aggressive, they attack the police and local population and this is a reason for placing barbed wire on the borders, upgrading protective military equipment etc.;
- Migrants are ungrateful – when they (self)organize for their rights, they never get enough. We help them and it is still not enough;
- Migrants do not respect women. Since it is mostly men arriving, their negative attitude towards women makes them prone to harassment and rape;
- Advocates for the rights of migrants are well paid and employed by various non-governmental organizations financed by individuals wishing to destabilize the western society. If they support migrants so much, they should welcome them in their homes.

Co-funded by the European Union
The Balkan Trust for Democracy
A PROJECT OF THE GERMAN MARSHALL FUND
Norway
Co-funded by the European Union, Balkan Trust for Democracy, a project of the German Marshall Fund of the United States, and the Norwegian Ministry of Foreign Affairs.