

campaign toolkit

minority
rights
group
europe

Countering cyberhate against Roma:

A toolkit for civil society organizations and activists

Marco Cadena, Zsófia Farkas and Nicole Garbin



[#FreedomFromHate](#)



Bulgarian youth shooting
their video about tolerance

Authors

Marco Cadena, Digital Communication Officer of Minority Rights Group Europe.

Zsófia Farkas, Managing Director of Minority Rights Group Europe.

Nicole Garbin, Freedom from Hate Programme Coordinator of Minority Rights Group Europe.

Acknowledgements

Project: Freedom from hate: Empowering civil society to counter cyberhate against Roma. REC-RRAC-ONLINE-AG-2017.



This toolkit has been produced with the financial assistance of the European Union. The contents of this report are the sole responsibility of Minority Rights Group Europe and can under no circumstances be regarded as reflecting the views of the European Union.

Minority Rights Group Europe (MRGE) has been based in Budapest since 1996, with the aim of promoting and protecting the rights of minorities and indigenous people across Europe and Central Asia. This toolkit does not necessarily represent in every detail the collective view of MRGE or its partners. Copies of this toolkit are available online at www.minorityrights.org. The toolkit was developed based on the training materials co-designed by MRGE and Media Diversity Institute for the training of project partners at the beginning of the programme. We are very thankful to Media Diversity Institute for their support.

A draft of this toolkit was reviewed by Anna-Laura Schreilechner, *Zivilcourage & Anti-Rassismus-Arbeit (ZARA)* Advanced Monitoring trainer. We are extremely grateful to her for her valuable feedback.

© Minority Rights Group Europe 2020

All rights reserved.

Material from this publication may be reproduced for teaching or for other non-commercial purposes. No part of it may be reproduced in any form for commercial purposes without the prior express permission of the copyright holders. For further information please contact MRG. A CIP catalogue record of this publication is available from the British Library. **ISBN:** 978-615-81504-0-8 / 978-615-81504-1-5 (pdf) **Published:** January 2020.

Countering cyberhate against Roma:

A toolkit for civil society organizations and activists

Introduction	1
1 What is hate speech?	3
1.1 A brief introduction	3
1.2 International legal standards	5
1.3 Hate speech in the Council of Europe and the European Union legal frameworks	8
1.4 Online hate speech	11
1.5 IT companies and online hate speech	12
1.6 Responses to hate speech	15
2 Counter hate speech strategies	17
2.1 Counter-speech and counter-narratives	17
2.2 Campaign planning	20
2.3 Content creation	26
2.4 Safety measures	28
3 How to develop a training	31
3.1 Planning the training	31
3.2 Training design	33
3.3 Training methods	34
3.4 Evaluation	37
4 Resources and tools	39
Useful resources	39
Tool 1: Counter-narrative strategy template	41
Tool 2: Campaign plan template	42
Tool 3: Pre- and post-training checklist	45
Tool 4: Training design worksheet	46
Tool 5: Online hate speech training agenda sample	47
Tool 6: Evaluation form sample	48
Notes	50

Introduction

Freedom from Hate is a two-year programme co-funded by the European Union (EU). It tests and evaluates effective counter-narrative campaigns targeting online hate speech against Roma communities in Bulgaria, Croatia, Czech Republic, Hungary and Slovakia, and across Europe.

With far-right groups becoming increasingly prominent and xenophobic discourse on the rise, in recent years there has been a resurgence of hate speech against Roma. Bulgaria, Croatia, Czech Republic, Hungary and Slovakia have some of the largest Roma communities in Europe proportionate to their populations and have seen extensive online hate speech against them. Counter-narrative campaigns are therefore urgently needed, and the public participation of Roma themselves, in particular through social media platforms, is fundamental.

This programme works with civil society organizations (CSOs) in the target countries to develop effective campaigns to challenge online hate speech. Our partners include:

- *Amalipe Center za mezhdruetnicheski dialog i tolerantnost* (Centre for Interethnic Dialogue and Tolerance), based in Bulgaria, is a leading Roma organization working for the equal participation of Roma. The organization plays a central role organizing the Roma civil movement and advocates for the inclusion of Roma in government institutions.
- *Institut ľudských prav* (Human Rights Institute), based in Slovakia, works to promote human rights through online and offline campaigns. The organization has significant experience in tackling online hate speech, including an online campaign on housing issues for marginalized Roma communities.
- *Romedia Alapítvány* (Romedia Foundation), in Hungary, is a regional media organization managed by Roma. Romedia runs media campaigns to promote awareness and understanding of the Roma community. A recent project involved delivering filmmaking and journalism training for Roma women.
- *Romsko nacionalno vijeće* (Roma National Council, RNV), based in Croatia, is an umbrella organization for the protection and promotion of Roma minority rights. The RNV supports inclusion of Roma while preserving their historical and cultural heritage and strengthening their identity.
- *Fórum pro lidská práva, z. s.* (Forum for Human Rights), based in the Czech Republic, focuses on international human rights litigation and advocacy in

Countering cyberhate against Roma

Central Europe. Its lawyers have represented a number of strategic cases in the Czech Republic aimed at fighting discrimination against Roma.

This toolkit has been developed by Minority Rights Group Europe (MRGE) as a guide to identify hate speech and how to respond to it with counter-narratives. The toolkit is structured as follows:

- In *Chapter 1*, the international and European legal frameworks are outlined with a focus on the existing definitions of hate speech and current freedom of expression safeguards, the challenges posed by online hate speech, the role of IT companies in addressing these issues, and possible ways to respond to it.
- In *Chapter 2*, strategies to counter hate speech are explored and guidance on how to plan, design and run an online campaign is provided, with a section focusing on safety measures to ensure personal security and wellbeing.
- *Chapter 3* of the toolkit provides information and tips concerning the organization and development of a training event, covering various aspects from logistics to evaluation.
- *Chapter 4* includes references to useful resources as well as tools developed by MRGE that offer a starting point for developing an online campaign or delivering a training session on countering hate speech online.

1 What is hate speech?

1.1 A brief introduction

Hate speech is a **complex** and **controversial** term that might mean different things to different people. It may seem self-explanatory to most, but people tend to offer very disparate descriptions when asked. Its understanding may depend on various factors, including the identities of the victims of hate speech and their previous personal experiences, the context in which hate speech has been deployed and the identity of the speaker. Indeed, hate speech definitions often tend to be very broad and can include lawful forms of expression. At times, hate speech regulation may even be used as an instrument of oppression, rather than protection, to silence already marginalized voices – those critical of a government’s policies, for example, or those who hold unpopular views.

What do you think ‘hate speech’ is?

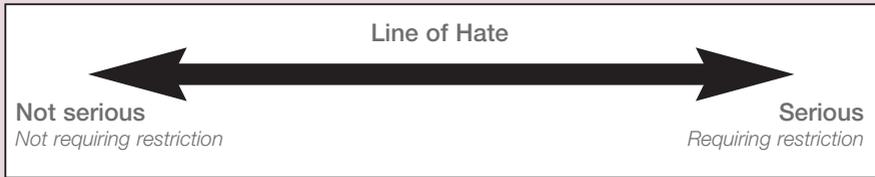
Before exploring the legal boundaries and implications of hate speech, ask participants to think about an example of hate speech and identify what elements could be considered to qualify the expression as hate speech (for example, its content, tone, targets or consequences), as well as their relevance in determining the seriousness of the examples given.

Below are some examples of questions that may be of help when guiding the discussion:

- Is who says it important?
- Can anyone be a target of hate speech (for example, members of majority or dominant groups)?
- Does context – for instance, whether something is said in private at home or on Twitter – make a difference?
- Does hate speech have to incite certain actions or cause an emotional response in the target?
- Does hate speech include expressions fostering stereotypes and intolerance?
- Is the tone important in defining hate speech?

Countering cyberhate against Roma

After having analysed a few examples, ask participants to locate those examples along the 'line of hate':



This exercise can provide the group with an opportunity to explore some of the challenges around hate speech from the start of the training, and also allow the trainer to understand the level of comprehension of these issues among participants.

However, it is recommended that before starting the training that participants are asked to be as vigilant as possible when describing examples of hate speech to avoid offending others or to create a situation where any participants may feel unsafe.

Hate speech and Roma

Hate speech, if left unchecked, can have far-reaching effects on its victims and on society as a whole, creating a cycle that may expose those targeted to further discrimination, harassment and violence. Over time, hate speech can force minorities to withdraw altogether from public life, reinforcing their marginalization. In order to fully understand its specific consequences regarding Roma, and to respond to it effectively, attention must be paid to **antigypsyism**, its deep historical roots in our societies and its systemic nature, both with regard to its mode of operation and effects.

Antigypsyism

The Alliance against Antigypsyism defines antigypsyism as:

'A historically constructed, persistent, complex of customary racism against social groups identified under the stigma "gypsy" or other related terms, and incorporates:

- 1. A homogenizing and essentializing perception and description of these groups;*
- 2. The attribution of specific characteristics to them;*
- 3. Discriminating social structures and violent practices that emerge against that background, which have a degrading and ostracizing effect and which reproduce structural disadvantages.'*¹¹

Antigypsyism has a long history and is still being perpetuated and institutionalized in the 'real' world.² However, today most anti-Roma rhetoric takes place online. The most common narratives can be categorized within three underlying themes: criminalization, welfare chauvinism and de-humanization, Roma generally being portrayed as outsiders, inferior citizens, and an imposed 'threat and burden' for majority society.

Moreover, hate speech may use different narratives and affect different people differently. Hate speech can have a **disproportionate impact** on women, members of minorities, LGBTQ+ and people with disabilities who can face further isolation and distress. Roma women are often confronted with discrimination or violence in every area of their lives because of both their gender and their belonging to a stigmatized community. This combination of gender-based and racist violence, well captured by the concept of **intersectional discrimination**,³ may expose Romani women to specific and aggravated forms of sexist, misogynist and racist hate speech targeting them on the basis of their multiple identities that will require added efforts to understand and address.

1.2 International legal standards

Hate speech has been defined and understood in different ways at the national and international levels, and the exact boundary between critical expressions and those of real hatred is difficult to establish.

Under international and European human rights law, there is **no universally accepted definition of hate speech**. Moreover, regulation of hate speech calls for a **reconciliation** between **freedom of expression** on the one side and the **right to equality** and **freedom from discrimination** on the other. Numerous international and European regional legal instruments regulate hate speech and freedom of expression. In this section the key binding and non-binding instruments dealing with hate speech are briefly examined.

Hate speech in international law

Despite the fact that hate speech is not defined in international law as such, several international instruments contain provisions focusing on various kinds of **expression that would typically be considered as constituting hate speech and must be prohibited:**

- **Article III(c) of the Convention on the Prevention and Punishment of the Crime of Genocide**

Direct and public incitement to commit genocide

Key elements

Direct: the expression must be sufficiently specific as a call for action, showing a close relationship between the expression and the danger of an act of genocide occurring. However, direct does not mean explicit, as implicit expression may also directly incite genocide if in its linguistic and cultural context it is sufficiently clear to its audience.⁴

Public: communication in a public place, or to the public or a section of the public, for example through mass media and digital technologies.⁵

Intent: the speaker must specifically intend to incite genocide and for it to occur.⁶

- **Article 20(2) International Covenant on Civil and Political Rights (ICCPR)**

Any advocacy of national, racial or religious hatred that constitutes **incitement to discrimination, hostility or violence.**

Key elements

Advocacy: 'intention to promote hatred publicly towards the target group'.⁷

Hatred: a state of mind, characterized by the 'intense and irrational emotions of opprobrium, enmity and detestation' towards a target group on the basis of a protected characteristic.⁸

Incitement: while the proscribed outcome need not in fact occur, the term 'incitement' strongly implies the advocacy of hatred must create 'an imminent risk of discrimination, hostility or violence against persons belonging to [the target group]'.⁹

'Advocacy' and 'incitement' both entail the intent of the speaker to incite others to commit acts of discrimination, hostility or violence. This should exclude from this conduct mere negligence or recklessness.

● Article 4 International Covenant on the Elimination of All Forms of Racial Discrimination (ICERD)

All dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin, and also the provision of any assistance to racist activities, including the financing thereof.

A clear and strict understanding of these definitions is fundamental to draw a line between speech that despite raising concerns in terms of intolerance and discrimination cannot and should not be restricted, and expressions that can and must be restricted.

Hate speech and freedom of expression

The definition of exact boundaries between speech that must be restricted¹⁰ and speech that cannot be limited¹¹ is paramount to ensure both freedom of expression and protection against hate speech, particularly in a context where states may use hate speech restrictions as a pretext to silence opposing views. Broad or vague definitions of hate speech and related crimes risk jeopardizing **freedom of speech**, one of the essential foundations of democratic societies.

This is why it is critical to assess hate speech in light of legal standards protecting freedom of expression. At the international level, **Article 19 of the ICCPR** sets those standards. While the right to hold opinions is an absolute right, **the right to express those opinions is not absolute**. However, **restrictions must be exceptional**: Article 19(3) ICCPR recognizes the possibility for states to restrict freedom of expression only where there is legal provision to do so (*legality*) and when it is necessary (*necessity and proportionality*) to protect the rights or reputations of others, national security or public order, or public health or morals (*legitimacy*).

Prohibitions referred to in Article 4 of the ICERD and Article 20 of the ICCPR need to meet the conditions set out in Article 19 ICCPR.¹² The scope, the content and the relationship among these provisions have been clarified by the work of the UN committees and bodies. The Rabat Plan of Action,¹³ for example, recommends a clear distinction between (a) expression constituting a criminal offence, (b) expression that is not criminally punishable, but may justify a civil suit or administrative sanctions, and (c) expression not giving rise to any of these sanctions, but still raising concern in terms of tolerance, civility

and respect for the rights of others. Also, it proposes a six-part threshold test for expressions to be considered as criminal offences, namely, one concerned with the particular **context, speaker, intent, content and form, extent of the speech act and likelihood** (including imminence), and clarifies that criminalization should be used as a last resort to combat the most serious kind of incitement under Article 20(2) of the ICCPR.

States increasingly rely on the digital industry to control, censor or monitor expression online. However, they should not use internet companies and hate speech frameworks as tools to limit expression that they themselves would be precluded from limiting under international human rights law. Internet shutdowns and the criminalization of online political dissent or criticism of government are among the examples of practices clearly in breach of Article 19(3) of the ICCPR.¹⁴

1.3 Hate speech in the Council of Europe and the EU legal frameworks

Council of Europe

At the Council of Europe level, there is no binding definition of hate speech and the European Court of Human Rights (ECtHR) does not provide for a definition of hate speech in its case law.

Despite the lack of a binding definition of hate speech, the following **non-binding definitions of hate speech** provide useful guidance:

- **Committee of Ministers Recommendation No. R 97(20)**¹⁵

All forms of expression which spread, incite, promote or justify racial hatred, xenophobia, antisemitism or other forms of hatred based on intolerance, including: intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants and people of immigrant origin.

- **European Commission against Racism and Intolerance (ECRI) General Policy Recommendation No. 15**¹⁶

The use of one or more particular forms of expression – namely, the advocacy, promotion or incitement of the denigration, hatred or vilification of a person or group of persons, as well any harassment, insult, negative stereotyping, stigmatization or threat of such person or persons and any justification of all these forms of expression – **that is based on a non-exhaustive list of personal characteristics or status** that includes ‘race’, colour, language, religion or belief, nationality or national or ethnic origin, as well as descent, age, disability, sex, gender, gender identity and sexual orientation.

Comparing the two definitions it is clear how the understanding of hate speech has evolved, expanding to include other protected categories not originally foreseen in international law, but which are now among the main targets of hate speech.

Hate speech restrictions and protection of free speech

The ECtHR does not provide for a definition of hate speech. However, in its jurisprudence, the Court has made clear that principles of equality and non-discrimination on one side, and freedom of expression on the other, both play a role in defining hate speech that cannot be tolerated.

‘Tolerance and respect for the equal dignity of all human beings constitute the foundations of a democratic, pluralistic society. That being so, as a matter of principle it may be considered necessary in certain democratic societies to sanction or even prevent all forms of expression which spread, incite, promote or justify hatred based on intolerance.’¹⁷

‘Freedom of expression constitutes one of the essential foundations of [a democratic] society, one of the basic conditions for its progress and for the development of every man ... It is applicable not only to “information” or “ideas” that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population.’¹⁸

In a case-by-case analysis the Court, not always consistently, uses the two following approaches:

● Article 10 European Convention of Human Rights Setting restrictions on the protection

The speech, although it is hate speech, **is not apt to destroy** the fundamental values of the Convention. If an interference with freedom of expression exists, this must be prescribed by law, pursue a legitimate aim, and be necessary in a democratic society.

Example → *Vejdeland and others v Sweden*¹⁹

The applicants were convicted for distributing leaflets which contained extremely offensive material about gay people, including in or on pupils’ lockers. They claimed their purpose was to start a debate. The Court found no violation of Article 10 by the Swedish authorities on the ground that the statements had constituted ‘serious and prejudicial allegations,’ even if they had not been a direct call for hateful acts.

Countering cyberhate against Roma

● Article 17 European Convention on Human Rights (ECHR) Exclusion from the protection

The speech amounts to hate speech and **negates** the fundamental values of the Convention. For instance, incitement to violence or racial hatred may be excluded from the protection of the Convention.

Example → *Seurot v. France*²⁰

A teacher was sanctioned for an article he wrote that was published in a school bulletin. In the article, the author deplored – as he put it – the overrunning of France by ‘hordes of Muslims’ from North Africa. The Court found that this sanction did not violate the applicant’s rights under Article 10 of the ECHR, because of the undeniably racist tone of the article and the duties and responsibilities of the applicant in his capacity as a teacher.

European Union

In the EU, it is the Council Framework Decision on combating certain forms and expressions of racism and xenophobia by means of criminal law which regulates hate speech.²¹ It indicates that the **following intentional conduct** is punishable:

- ‘**publicly inciting to violence or hatred directed against a group of persons** or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin’.
- ‘**publicly condoning, denying or grossly trivialising crimes of genocide, crimes against humanity and war crimes** as defined in Articles 6, 7 and 8 of the Statute of the International Criminal Court’ and in ‘Article 6 of the Charter of the International Military Tribunal appended to the London Agreement of 8 August 1945, directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin when the conduct is carried out in a manner likely to incite to violence or hatred against such a group or a member of such a group’.

Only some of the possible potentially vulnerable categories are indicated in the definition, leaving out other equally relevant ones. For this reason, the European Parliament highlighted the need for a revision of the Framework Decision, so as to include also manifestations of antisemitism, religious intolerance, antigypsyism, homophobia and transphobia.²²

1.4 Online hate speech

‘Online hate is no less harmful because it is online ... To the contrary, online hate, with the speed and reach of its dissemination, can incite grave offline harm and nearly always aims to silence others. The question is not whether to address such abuse. It is how to do so in a way that respects the rights everyone enjoys.’²³

With the advent and growth of the internet, we are witnessing profound changes in the way we communicate, with an ever-larger volume of information and perspectives readily available to a wide audience. At its best, the internet has a central role to play in enhancing public debate and engagement in democratic societies.²⁴

Principles and legal standards relevant to online hate speech:

- **Offline standards apply to online contents:** the Council of Europe’s definition of hate speech covers ‘all forms of expression’, including hate speech conducted online through the use of new technologies and platforms. Article 19(2) of the ICCPR covers all forms of expression and the means of their dissemination, including electronic and internet-based modes of expression.²⁵
- **Providers and platforms have a responsibility to vet their user content for hate speech:** internet news portals which, for commercial and professional purposes, provide a platform for user-generated comments, assume the ‘duties and responsibilities’ associated with freedom of expression in accordance with Article 10(2) of the ECHR, when users disseminate hate speech or comments amounting to direct incitement to violence.²⁶
- **Restrictions of online contents must respect international standards on freedom of expression:** ‘Any restrictions on the operation of websites, blogs or any other internet-based, electronic or other such information dissemination system, including systems to support such communication, such as internet service providers or search engines, are only permissible to the extent that they are compatible with [Article 19] paragraph 3. ... It is also inconsistent with paragraph 3 to prohibit a site or an information dissemination system from publishing material solely on the basis that it may be critical of the government or the political social system espoused by the government.’²⁷
- **States are responsible for preventing dissemination of hate speech, including online:** The Committee on the Elimination of Racial Discrimination (CERD) recommends that the states parties ‘take strict measures against any incitement to discrimination or violence against the communities, including through the Internet’.²⁸

Countering cyberhate against Roma

Hate speech online is not intrinsically different from similar expressions found offline. However, its **versatility**, the **anonymity** it allows, its **potential to reach out to large audiences quickly** and the relatively **low entry barriers** characterizing its accessibility, raise peculiar challenges unique to online content and regulation.²⁹

As of today, the only binding international instrument related to online hate speech is the **Additional Protocol concerning the criminalization of acts of a racist and xenophobic nature committed through computer systems**³⁰ of the **Convention on Cybercrime of the Council of Europe**, also known as the **Budapest Convention**.³¹ The Budapest Convention provides for the minimum legal standards necessary for the fight against crimes committed over the internet, serves as a guideline for any country developing comprehensive national legislation against cybercrime, and as a framework for international cooperation between state parties to this treaty. The Additional protocol entails an extension of its scope so as to also cover offences of racist or xenophobic propaganda.

● Article 2 of the Additional Protocol provides that:

“racist and xenophobic material” means **any written material, any image or any other representation** of ideas or theories, which **advocates, promotes or incites hatred, discrimination or violence, against any individual or group of individuals**, based on **race, colour, descent or national or ethnic origin**, as well as **religion** if used as a **pretext** for any of these factors.’

Key elements

‘Advocates’ refers to a plea in favour of hatred, discrimination or violence, ‘promotes’ refers to an encouragement to or advancing hatred, discrimination or violence and ‘incites’ refers to urging others to hatred, discrimination or violence.

1.5 IT companies and online hate speech

When dealing with hate speech online, internet intermediaries such as social networking platforms, internet service providers or search engines tend to develop their own definitions of hate speech and measures to respond to it, stipulating in their terms of service how they may intervene in allowing, restricting or channelling the creation and dissemination of specific content. Given the harm caused by the spread of illegal content online, as well as the technological capacity needed to handle online services, governments have been placing increasing pressure on companies to act as adjudicators of hate speech.

Many online service providers have taken on those responsibilities and progress has been made through, for instance, the adoption by Facebook, Microsoft, Twitter and YouTube of the ‘Code of conduct on countering illegal hate speech online’, helping users to report incidents of illegal hate speech on social platforms.

The European Commission’s Code of Conduct for major IT companies

In an effort to respond to the proliferation of racist and xenophobic hate speech online, the European Commission launched its Code of Conduct in May 2016 together with four major IT companies (Facebook, Microsoft, Twitter and YouTube).

The key commitments related to the removal of illegal hate speech online include the following:

- ‘The IT Companies to have in place clear and effective processes to review notifications regarding illegal hate speech on their services so they can remove or disable access to such content.’
- ‘The IT companies to have in place Rules or Community Guidelines clarifying that they prohibit the promotion of incitement to violence and hateful conduct.’
- ‘Upon receipt of a valid removal notification, the IT Companies to review such requests against their rules and community guidelines and where necessary national laws transposing the Framework Decision 2008/913/JHA, with dedicated teams reviewing requests.’
- ‘The IT Companies to review the majority of valid notifications for removal of illegal hate speech in less than 24 hours and remove or disable access to such content, if necessary.’

Today, nine companies adhered to the code, notably Facebook, YouTube, Twitter, Microsoft, Instagram, Google+,³² Dailymotion, Snapchat and Webedia.

One year on, it was reported that on average IT companies responded to 59 per cent of reports of illegal hate speech by removing the content, more than double the response rate (28 per cent) six months before. The volume of flagged incidents assessed within 24 hours rose from 40 per cent to 51 per cent in the same six-month period.³³ There has been noticeable progress in the years since: by early 2019, IT companies were responding to 89 per cent of reported incidents within 24 hours and removing 72 per cent of reported incidents of illegal hate speech – a relatively high proportion, given that in some cases content highlighted by users may not in fact be illegal.³⁴

Countering cyberhate against Roma

Despite the enthusiasm around the adoption of measures such as the Code of Conduct, there are also **concerns** regarding the risks of applying this pressure on private companies, especially with regard to immediate takedown of content through automated AI (artificial intelligence) filters. **Internet intermediaries are expected to respect human rights.**³⁵ However, censorship without judicial overview, silencing minorities, lack of transparency when implementing community standards, demonization of unpopular expressions with the support or upon request of authoritarian governments are among the most concerning challenges:

*'Under protest, in silent acquiescence or as willing participants, they [private companies] are often essential to State censorship and surveillance. ... What governments demand of private actors, and how those actors respond, can cripple the exchange of information; limit journalists' capacity to investigate securely; deter whistle-blowers and human rights defenders. Private actors may also restrict freedom of expression on their own initiative. They may assign priority to Internet content or applications in exchange for payment or other commercial benefits, altering how users engage with information online. Companies that offer filtering services may influence the scope of content accessible to their subscribers.'*³⁶

1.6 Responses to hate speech

Responses to hate speech need to recognize that 'hate' falls along a spectrum: although all expressions of hate may be bad to some extent, one case can still be worse than another: for example, it may be more offensive, it may affect larger numbers of people, be more inflammatory, potentially more damaging, and so on. The tripartite approach delineated in the Rabat Plan of Action provides guidance in drawing a distinction among the different forms of hate speech and approaches to respond to it.

Strategies to counter hate speech are various and can be regrouped in the following three categories:

1. **Submit a formal complaint:** Depending on the country and the applicable law, there may be the possibility to:
 - lodge a formal criminal complaint with the police;
 - a civil complaint before the national courts under anti-discrimination law; and/or
 - a complaint before the equality body and/or national human rights institution.

For your training

- Carry out research in order to determine the applicable law defining hate speech and available ways to challenge it in court.
- Consider the advantages (for example, its deterrent effect) as well as the risks (the length of the procedure, the possibility of the perpetrator using the procedure to gain more visibility and resonance for his/her messages) of submitting a complaint.
- Consider risks like victimization or barriers in access to justice for members of disadvantaged and vulnerable groups.
- Verify the mandate and powers of the competent equality body and/or national human rights institution and the conditions for lodging a complaint before it.

2. **Report:** Social platforms community guidelines, standards and terms of service provide users with the opportunity to report illegal hate speech. You can report to IT companies directly using channels available to general users, and for IT companies adhering to the Code of Conduct also to trusted flaggers/reporters which benefit from specific channels available only to them.

For your training

- explore in advance the terms of service/community standards of the most popular IT companies among your target audience and set a time during the training to explore with the participants the reporting tools.
- search for organizations which have the trusted flagger status and/or are active on relevant issues which you think participants should be aware of for reporting and getting support.

3. Provide an alternative narrative – counter-speech

Techniques and strategies on how to counter hate speech online and develop alternative narrative campaigns online are developed in Chapter 2 of the Toolkit.

If you do witness hate speech online, do not leave those directly targeted alone.

Show your support by:

- Liking their post.
- Show that you agree with their statement.
- Show that you disagree with haters.
- Condemn hate speech.
- Providing facts supporting victims' statements or contradicting haters' ones.
- Providing victims with information about the possibility of reporting hate speech, and how to access support organizations and bodies.

Wrap-up exercise

At the end of this session, a wrap-up exercise allowing the trainer and participants to review the key points covered is suggested.

The trainer, for example, can go back to the examples of potential hate speech discussed during the first exercise and ask participants to again situate those examples along the 'hate-speech line' in light of the information provided during the session, to see whether different conclusions are reached.

2 Counter hate speech strategies

Challenging hate speech is important, but it is not always effective – and sometimes it can even be counterproductive. It is very important to consider the context, the environment, and whether the perpetrators or their audience would be open to persuasion or are actively trying to trigger a response (a practice known as trolling).

The perpetrator could be a bot – a fake profile – and without adequate safety measures you could become a target of hate speech and threats. So, assess the risks and take steps to mitigate them before challenging the haters.

There are 3 billion social media users today, yet a small percentage of users can poison conversations with hate speech, bully and divide people, and put many people off from even engaging online.

2.1 Counter-speech and counter-narratives

Counter-speech can be a one-off comment, a way of responding to hateful messages.

Counter-narratives, on the other hand, are about a **longer-term strategy**, reframing a particular narrative online. For example, right-wing groups may attempt to tap into established misconceptions, such as the belief that Roma people do not like to work: that is of course a myth and a negative stereotype, which often fuels outright racism and hate. Challenging that narrative, however, requires a more concerted effort and tactics to reach a wider number of people and effectively transform their views.

The case studies below showcase just a few examples of the diverse range of tactics that can be applied as counter-narrative campaigns.

Reporting

If the comments are very hateful and offensive, it is also worth considering reporting either the comments or the person making them. Note, however, that it is highly unlikely that the social media platforms will go as far as taking down the perpetrator's accounts. Unfortunately, the typical response is to suggest that users block the offender themselves and only to take direct action in the most extreme

Countering cyberhate against Roma

instances or if it can be proven that the account itself is fake. On Facebook, if people manage multiple accounts that are in violation of Facebook's Community Standards, it is likely the company would take the accounts down. The same does not apply to Twitter, Instagram or YouTube.

Counter-speech and counter-narratives can be effective tools, but the following has to be considered in every case:

- **Is the person you are trying to engage with open to persuasion?** For example, a message with a very aggressive tone might imply that the person is likely to ignore or even be encouraged by an effort to change their viewpoint.
- **Is it safe to speak out?** Take into account where the conversation is taking place. Are there any commenters who are likely to agree with you or be open to your views? Consider whether your comment would help persuade the audience reading the conversation.
- **Is the comment part of a more concerted effort?** Try to look up the profile of the perpetrators. Are there any recognizable similarities? Do you suspect they are fake profiles? If so, make sure you report them immediately.

Case studies: successful counter-speech campaigns

Below are some examples of successful counter-narrative and counter-speech campaigns. The list is not exhaustive but intended to serve as inspiration and to show that there are plenty of different approaches that counter-speech campaigners can employ.

Like Attack

The aim of the campaign by **Laut Gegen Nazis (Loud Against Nazis)** was simple: flood the NDP (Germany's neo-Nazi party)'s Facebook page with a storm of likes, by a diverse group of Germans, and then flood the page with positive, funny anti-racist messages. More than 100,000 people participated in the action, generating plenty of attention and media coverage. <https://www.lautgegennazis.de>

Game of Trolls

Get the Trolls Out (GTTO) developed short videos to educate people about how to respond to trolls. The campaign had three main pieces of advice: do not be abusive, respond promptly and involve the community for help. <https://www.getthetrollsout.org>

#JegErDansk (#IAmDanish)

The campaign drew attention to a very serious issue, which is often ignored: how minority and second-generation immigrant children feel when confronted with the Danish government's anti-immigration stance and with messages that claim they are not Danish. The video received more than 15 million views and garnered multiple awards internationally.

<https://www.youtube.com/watch?v=e7mqfmZS5xM>

Face Your Brother

This viral campaign, launched for #IDAHOT2017 (International Day Against Homophobia, Transphobia and Biphobia), was designed to raise awareness around LGBTQ+ Muslims and the International Day Against Homophobia, Transphobia and Biphobia. Featuring a practising Muslim who also happens to be transgender to highlight the internal conflict for both the speaker and the listener, the result was a beautifully executed and courageous visual surprise that achieved more than half a million views and won several international awards.

https://www.youtube.com/watch?time_continue=61&v=jAT340d4eY4

Face Your Sister

This video by the Danish Women's Society was designed to reach female Muslim audiences online to engage more women with minority backgrounds. Using a story that women of all backgrounds could relate to, the video demonstrated how Muslim hijab-wearing women have the same experiences in order to promote reflection and female solidarity.

https://www.youtube.com/watch?time_continue=84&v=AU9QStJUf7I

#ichbinhier

In this model, organizations such as Ichbinhier in Germany gather a number of supporters and activists, to be able to mobilize them against online hate speech. Ichbinhier has over 37,000 people who sign up for actions, and are directed to hateful content, to engage with it in a positive, fact-based manner.

<https://www.dw.com/en/german-anti-hate-speech-group-counters-facebook-trolls/a-38358671>

2.2 Campaign planning

As the examples above demonstrate, a concerted campaign – whether large or small – with an enthusiastic, passionate group of people behind it can achieve great results online. However, it is worth bearing in mind some of the following considerations before campaigning against hate speech on a larger scale.

Audience

One of the key elements of successful campaign planning is understanding your target audience. The key questions you need to ask are:

- Who are you trying to influence?
- What do you want your audience to **do** as a result of your campaign?

Defining the audience is key for a successful campaign, and it needs time and discussion within your group to think through and debate what would be the best target audience for a planned campaign.

The list below presents some of the areas you should consider when defining your target audience.

- **Awareness-raising in a certain locality or sector:** for example, targeting students, their teachers and parents and other youth support workers in a particular town or region.
- **Reaching specific age or gender groups,** such as women aged 18–24.
- **Tailoring appropriate content:** for instance, young people who search and watch extremist online content, including those at the brink of radicalization.

Finding your audience

It is very important to understand what your audience is doing online and offline. If you understand your audience you will have a better sense of how to talk to them and what their likely reaction might be.

It is also important to involve members of your audience when developing your content. For example, if the campaign aims to build a counter-narrative around a particular issue, like school segregation, it is important you involve students, teachers and parents to help develop the content together. At the very least, when designing provisional content and messages be sure to consult a select group of people from your audience to get their feedback on your campaign and any suggestions for improvement.

Get to know your audience

Understand how your audience is interacting online – for example, what social media they use and what they are really interested in. The more information you can collect, the more effective you will be in finding out their habits and behaviour, and the better your chance of being able to produce relevant content for them.

Do not worry if you cannot research every single detail: this is more about finding out what sort of content interests them, what sort of style and language they use, what media they engage with the most. Gut feeling is also important. Listen to your instincts and test different types of content, especially if your research is not very detailed.

Reach out

Contact organizations and people with experience in counter-narrative campaigns. You should be able to find a lot of like-minded people and organizations who have already encountered similar challenges and questions, and who may be able to give you guidance. There are numerous anti-hate and counter-speech organizations, fact-checking websites and support communities – the more engaged you are with them, the more likely you are to build future relationships as well.

Message / Story

When you have an audience, you need a story to tell them. A story is a message which people can relate to and has a purpose. It is more effective if it does not try to simply lecture the audience, but offers points to reflect on and encourages the audience to arrive at a conclusion themselves. Messages and stories can take various forms, and some of the most common examples are collected here.

It is important to be thought-provoking and encourage reaction, especially in a counter-speech campaign. However, do try to avoid being over the top or pointlessly controversial as this can alienate sections of your audience.

Some different types of messages include:

- Humour and satire, undermining and ridiculing the efforts of extremist groups.
- Emotional message, highlighting the negative impact of hate speech and its impact on the victims and society.
- Fact-based messages, to debunk and discredit hateful extremist messages.
- Positive and inspiring messages from people within the audience.

The power of stories

It is increasingly evident that personal stories that are relatable and resonate with people's own experiences or lives have a much greater engagement rate and trigger a much stronger response.

It is a natural, when reading a story or a book, to relate to the protagonist, to imagine a certain scene or situation that resonates with the message of the story. The same happens when people come across campaign stories: something that speaks to the viewer's personal experiences has a better chance of triggering an emotional response.

It is not always easy to get first-hand personal experiences in a campaign, for many reasons, but when creating content you should try to imagine how your content might be perceived from the audience's point of view. Is it relatable and understandable for someone who is completely unfamiliar with the topic? Is it simple enough that people do not switch off? Does it have the potential to trigger an emotional response and is it not overdone?

These are important questions to keep asking as you go, but there will never be a perfect answer and you will only find certain things out once your content is already online, so it is better to experiment with various types of stories, and various angles of the same story, to see which one works best with your audience.

The effects of cyberhate in Hungary

Cyberhate is an ongoing problem in Hungary that further deepens negative prejudice and stigmatization of Roma in all areas of their life, for example in schools. This pernicious form of negative discrimination hinders all efforts to close the gap between Roma and non-Roma.

Romedia developed a campaign raising awareness about cyberhate and stereotypes regarding the Roma community in Hungary. The campaign aimed to challenge the negative perceptions widespread among many Hungarians towards Roma parents and children, spread in the form of cyberhate. The campaign, focusing on an imaginary family, showed how both parents and children are subjected to discrimination through the use of electronic means of communication.

This was a perfect example of a visual campaign structured around a universal theme – a family – and through their story addressing cyberbullying, an often unnoticed, yet crucial problem in society, Romedia joined the dots with images that captured the complexities of cyberhate.

Content types and platforms

Once you have a clear idea about your audience and the story you want to tell them, consider the best way to create it.

Choose the type of content and platform that your audience is using the most, and consider the language and tone. Platforms and content types vary greatly by country and context. For example, Twitter is very popular in English-speaking countries but less so in other regions. Instagram has a huge percentage of younger users and is the fastest growing platform today. Facebook, on the other hand, is getting increasingly 'old', having seen a great influx of users over 60 in recent years.

You might think there are a number of different ways you can express those messages, so it could be that you choose a variety of mediums in order to disseminate your message. Here are just the most common examples:

- **Text:** articles, letters, but also hashtags and slogans;
- **Images:** photos, infographics, memes, profile or cover pictures;
- **Video:** short movies, animations, documentaries, campaign videos, testimonies, interviews;
- **Other online materials:** posters, brochures, briefing documents; and
- **Audio:** music, podcasts, audio mixes, short clips, interviews, reports.

Extremist groups are producing a huge amount of content along these lines, so counter hate speech campaigners should do the same. Do research, however, what mediums would resonate best with your chosen audiences, to stay as effective as you can be.

If you are on a low budget, try to keep things as simple as possible. Always favour quality over quantity. A small amount of good-quality, well-thought-out materials will take you much further than a heap of low-quality, poorly considered content.

Learning to use a new type of content and platform in Slovakia

The Human Rights Institute (HRI) in Slovakia used memes (funny images with text overlay) for the first time in their counter-speech campaign in early 2019, and found that making relatable and funny content actually worked: 'We learned how to simplify complex messages.' Rather than trying to start an entirely new discourse, they tried to plug into an ongoing online discussion about fake news and this strategy proved successful in generating high levels of engagement.

HRI also experimented with a new platform (Instagram) for the first time and found that people from all age groups engaged with their content, not exclusively young people as is often presumed about Instagram. Humorous memes worked really well on Instagram, and, while they addressed a serious message, the medium was packaged perfectly for the platform's audience. Once they had built an Instagram audience, HRI started offering links to their websites and Facebook page, where people could find out more about their campaign.

Researching the haters

In most cases there are a number of elements and layers to hate speech by extremist groups and their 'aware' or 'unaware' followers. However, it is best to focus on certain aspects of the narrative you are trying to address, instead of trying to counter everything in one go.

For example, when it comes to hate speech against Roma, there may be a number of features in the hate narrative, from negative stereotypes (such as Roma not wishing to work or study) to questioning their right to belong (for example, claiming that Roma are not 'really' citizens in their country). Addressing all the elements in one go would be next to impossible, especially for a group with limited funding.

However, picking the right counter-arguments that could resonate best with your audience and tailoring the counter-messages accordingly could have a lasting impact.

Contrasting online vs real-life hate in Croatia

When the Roma National Council (RNV) started their counter-speech campaign, they had one objective: they wanted Croatians to know more about how much online hate the Roma community is exposed to on a daily basis in the country.

They spent a few days interviewing people on the street in the Croatian capital, Zagreb, and one of the questions they asked people was to read out hateful messages that were posted online. The reactions of people on the street was visible shock and confusion – since saying those abusive messages out loud in real life made people feel very uneasy. RNV used the opportunity of the International Roma Day on 8 April to pitch the videos for national television networks, and luckily some were picked up and broadcast on television nationwide.

Challenging the silent majority with powerful stories in Czech Republic

Forum for Human Rights in Czech Republic used a newly created Facebook page to publish stories about the dire housing situation the Roma community is facing in the town of Ústí nad Labem. By interviewing a number of people who were affected by discrimination and negative stereotypes, the organization was able to present these stories to a wider public in the city, exposing some of the underlying injustice and unfair treatment.

In one case, a disabled person described being evicted from a council house, finding himself not only separated from his family but being put into accommodation with no step-free access. His story triggered empathy and anger from the public, since through it people could understand easily how bureaucracy and structural discrimination could affect those who are most vulnerable.

Forum for Human Rights also busted the common popular myth that Roma residents pocket their housing benefits, by exposing how non-Roma landlords charge inflated rents for Roma tenants on benefits, in order to pocket the government's housing benefits themselves. This story also helped to challenge widespread stereotypes about the Roma community in a way that was relatable for most people.

Main aim and objectives

In order to ensure your campaign remains focused, it is best to set one overarching goal, and then a series of specific, measurable and achievable objectives that will help you on your way to achieving it.

These also help to set a benchmark to evaluate the impact of the campaign, and help you in your planning.

Some campaigns do have objectives, but often they are too broad, difficult to measure and not realistically achievable in light of the available resources.

Defining objectives is not always easy. In general it is important to consider the size of your audience, the resources you have available and your estimated costs. The aim of the campaign is not always to go viral but to reach the intended audience.

Campaign planning

Careful planning of the campaign is of utmost importance. Use the 'TOOL 1: Counter-narrative strategy template' and 'TOOL 2: Campaign planning template' at the end of this toolkit. Go through the following steps and, using the key campaign elements, elaborate your own draft campaign plan. Rely on the findings of your research about perpetrators, messages, platforms and audiences, etc. Look for inspiration in already existing campaigns and use those elements which you think would be most effective in your country. You can use the tools and tips from the <http://www.counternarratives.org> website to design your own campaign.

Funding and budgeting

While some funding is helpful, there are plenty of examples of highly successful campaigns with very small budgets, and in contrast, huge failures with enormous budgets.

So budget is not the real issue. If your group can make the most of what is out there and have the commitment and passion to see your campaign through, you are already in a very strong position.

It is important, however, to break down the costs that you would incur and the resources you have available, as well as to keep accurate records of your spending.

2.3 Content creation

Needs assessment

If you follow your planning steps – defining your audience, reflecting on the story you wish to tell and the medium best suited to deliver it – you will have some idea as to what content is needed for the campaign.

Always keep your goals and objectives in mind, as those are the benchmarks of your campaign.

Tone and style

Each campaign can have a variety of different tones. It can be provocative, cool, satirical, antagonistic, sad, regretful, and a host of other characteristics.

For each campaign, it is best to work out what would fit the campaign's goals, objectives and audience best. Would it work better with a personal story or testimonies? Would interviews with victims help? Would it need animation to describe something abstract?

Stories that create empathy and which people can relate to are likely to generate a stronger response. It is important that the topic and style resonate with the audience, perhaps by exploring a problem or an ongoing social issue that is shared by many. For instance, housing is a challenge that is experienced not only by Roma but by many others in society, so it is possible to make the topic resonate with people who are not familiar with the specific issues Roma communities face. Having someone from the target audience delivering the message can also have a strong impact and encourage the audience to engage with the message.

Some of the key questions you should think about when it comes to content design:

- What do you want people to **feel** when they see your content?
- What is the single thing you would want people to **remember** from the campaign?
- Is the language easy to **understand** and appropriate for your chosen audience?

The tone and style of the campaign is ultimately shaped by you as a person, and the group behind it. It is also good to consider data from previous campaigns or use the data from the current campaign to design the next one.

Engaging Roma youth in Bulgaria

In Bulgaria, the organization Amalipe worked with its volunteers to create short videos about online bullying and harassment. The videos were emotional and targeted mainly at Roma youth. Since the people giving the message in the videos were themselves young Roma, naturally the audience responded to the message and the video was shared widely in the Roma community.

Amalipe found that focusing on peer-to-peer communication, with content that explored issues in an open and honest fashion, helped to deliver an authentic message while also leading to higher engagement levels online.

Opening up public discussion of Roma issues in Croatia

As mentioned earlier, Croatia's Roma National Council (RNV) has managed to air its street interviews with members of the public on national television, helping to create a public debate about a topic that is rarely discussed at the national level.

Even though the videos were not professionally produced, they attracted the attention of the media due to the fact that RNV timed their release around International Roma Day. 'Video creation takes plenty of time, skills and technological resources,' according to one staff member, 'but the videos were attractive for media and others.'

Often the media prefer 'real-life' videos and experiences, and the opinion of members of the public. Therefore, with the right approach, it is possible to engage with national media on topics that are not often discussed, if the campaign is timely and coincides with some bigger international dates.

Testing content

It is often good to get a second or third opinion about your campaign or content from a focus group of specific people from your audience, as it is possible to redefine and fine-tune your content during your campaign. Data from social media platforms should also give you a good indication as to which content triggers the most effective reaction.

Ideally, you should also test your content with Roma community representatives. You do not want to reinforce stereotypes inadvertently through your campaign, and this will be a good way of making sure that you do not.

2.4 Safety measures

It is always important to stay safe online, but this is especially the case if you are engaging hate speech perpetrators.

Personal information

When challenging hate speech online, it is better to keep your personal information hidden from public view. On all platforms you can choose what information to give and what information to display as public information. Do check on every platform, as in most cases the default settings are very open, meaning anything you put down as personal information will be public by default, which could include your address, phone number and other personal information.

Key points to consider, when planning an online campaign:

- Do you want the campaign to link back to you or your organization?
- Do you want the organization to be visible online?
- Are others in your organization also aware of the safety measures they need to take?
- Are all your passwords up to date and strong?
- Do your computers have up-to-date software?

Comments and backlash

It is always worth considering the risk of a potential backlash to your campaign and, while planning the campaign, also prepare a strategy as to how to react to negative comments, as once your campaign gets some traction, it will probably generate many comments.

There are different ways to manage comments, and there is no single solution, as different actions can trigger different reactions. For example, deleting hateful comments could lead people to react against ‘censorship’, but it can also help root out vicious and offensive comments. Responding to comments with your own or your organization’s account could also have both positive or negative effects. Ignoring a comment may contribute to it being quickly forgotten, but if you are faced with coordinated trolling by many users, it can also let the discussion get out of hand, as other people will also be reinforcing the initial hate speech. If you think a comment goes against the terms of the particular platform, you should also consider reporting this and blocking the person entirely.

What you need to bear in mind is that whatever you decide to do, you should be on top of whatever comments are being made about your campaign, rather than respond reactively without considering the possible effects. Be prepared logistically to be able to interact quickly so that the discussion does not escalate.

Also, if you do decide to engage, you may consider trying to deal with the issue privately. There are several manuals and toolkits that already provide some practical tips on how to deal with negative comments and trolls that you may want to study while preparing your campaign. It is worth conducting research online for more information, and then decide with your group as to what is the best way to deal with comments for your particular campaign.

Personal wellbeing

When campaigning on difficult topics such as hate speech or human rights, it is also essential that people take care of their own wellbeing. In these campaigns you may be exposed to malicious messages, abuse and threats. It is therefore vital for every campaigner to understand and monitor their own tolerance levels, and take the necessary measures to protect themselves from burnout. Some of these might help:

- Take regular breaks for a couple of days, and always talk with your team or colleagues if you feel the need. Do not allow bad feelings to linger, find a way to talk through them and clear your head.
- Build or join existing online communities where you can talk through issues and concerns.
- Do get professional help or counselling early, if you feel the work is affecting your mood or you develop signs of depression or anxiety. It is important to learn about symptoms and act early if you feel the need.

3 How to develop a training

Coordinating a training course requires a variety of steps, tasks and skills. Although a lot of training coordination takes place during the design phase, coordination is very important during all phases of the training process. Coordination begins at the moment training is first proposed and continues even after it is delivered and participants leave.

The various 'hats' a training coordinator wears include communications manager, materials producer, facilities and equipment supervisor, problem-solver, and even entertainment director. In short, coordinating training requires endless management of many details and people.

3.1 Planning the training

When you organize a training event, your tasks will be divided between designing the training, developing and preparing training materials, preparing logistics and evaluation. The '*Tool 3: Pre- and post-training checklist*' will help ensure you do not forget anything. Be sure to devise a timeline for each stage of the preparations; it is crucial that you book the venue, contact the speakers and invite the participants early enough to avoid problems later.

The tasks

Training design:

- Determine the audience for the training and their knowledge and skills needs.
- Identify appropriate training methods for the target audience.
- Identify trainers and facilitators.

Participant and training logistics:

- Send invitations and circulate information about the training. Be sure to formulate clear aims around the training to avoid raising expectations that will not be met.
- Registration.
- Check travel, accommodation and other arrangements.
- Identify venue.

Countering cyberhate against Roma

Training material / curriculum development:

- Develop training material content.
- Prepare material for dissemination.
- Write training material text.
- Produce visual aspects of training material.

Training materials:

- Produce curricula, ordering of materials and printing.
- Obtain all training materials, e.g. binders, pens, nametags, resource lists.
- Collate and assemble training materials, e.g. binders.

Evaluations:

- Conduct training evaluation.

Accessibility

In the early stage of planning, you should think about potential trainees with special needs. Make sure that the venue fulfils all criteria of accessibility. Provide sign language translators if needed.

Needs assessment

When you begin planning a training event, the next four questions must be answered. Those answers will help you to assess who is your audience and what gaps in their knowledge your training will address.

- **Audience:** Who is the target audience for the training?
- **Current roles:** What do members of this target audience presently do in their roles?
- **Knowledge gaps:** What gaps exist between what these providers know how to do, and what they need to know to carry out their roles successfully?
- **Outcome:** Will training help fill this gap?

These questions form the foundation of a training needs assessment. You may already have previous experience of working with the target group. First, make sure you use any relevant experience and data from within your organization to support the training. Second, talk to your target group well in advance about their needs and interests. Use the following steps to determine your participants' needs:

- Draw from your past experience with similar groups.
- Gather information from informal discussions among potential trainees.

- Conduct surveys.
- Observe participants' work practices.
- Interview participants.
- Learn about critical incidents of hate speech in the country.
- Use data and research on your target group.

3.2 Training design

The findings of your needs assessment inform the training design components. For example, if you know the level of understanding among your target audience of hate speech legal frameworks and what they need to learn, you can then develop learning outcome statements accordingly. A needs assessment will also help you determine who you will need as experts for the training, and how long a course should be. When you develop your training design use the *'Tool 4: Training design worksheet'* at the end of this chapter. Four elements are necessary in order to design a successful training:

- What are the **learning objectives**? What do you expect the trainees will know and apply in their work after the training?
- What kind of **expert or trainer** will you need?
- What kind of **training methods** should you use?
- **Where, when and how long** will the training be?

Once you have set your objectives and clarified the target audience, as a next step you should determine the length of the training and the structure.

For example, if you organize a training for activists from one region or city, you may choose a one-day training module. If participants travel from other regions or countries, your training will be 2–2.5 days long, leaving trainees sufficient time for travel. If you plan a training event that aims to build various skills among the trainees, you need at least 4–5 days training that enables you to give lectures, work on case studies and include practical exercises in the design.

If the topics of the training require a particular expertise, for example someone specializing in social media and communication or legal issues, you may choose a facilitator for the overall management of the training event and invite specialists for only one or two sessions. In many cases, the speakers are chosen to share their own practices or experiences in another country. Those sections might be the highlights of the training, especially if the presentation involves sharing materials, films, campaign materials, videos or publications. In both cases, always leave time for questions and personal encounters after the session.

'Tool 5: Online hate speech training agenda sample', at the end of this chapter, includes a design for a one-day training session on counter hate speech campaigns. This training programme assumed that the trainees had already worked on human rights violations and minority issues, even if they had a limited knowledge of hate speech and how to challenge it through their campaigns.

About speakers or trainers

Participants often like to know who will be leading the training in advance. If recognized experts in a particular field will be participating, that can also be a way of generating interest ahead of the event. When you organize a training event, ask prospective speakers to provide information on the following:

- their educational background
- learning objectives for their presentations
- a list of audio-visual equipment and handouts they will use during their presentations
- a biographical sketch.

3.3 Training methods

When you design a training programme, choosing the most appropriate training method for your audience and the selected topic is essential. You may also vary the methods to keep attention levels high and mix passive and active participation.

Lecture: this provides concise information on a topic. Key speakers or recognized experts are expected to give an introduction to a topic.

Panel discussion: with multiple speakers at the panel, this offers a range of different perspectives. The panel discussions are usually facilitated by a moderator.

Presentation with slides: this is a commonly used method, featuring illustrations, graphs, figures and tables, with handouts to ensure further dissemination of the information.

Film screening: screening short educational films, case studies or longer documentaries can be a very appealing element to include in trainings. Providing discussion points in advance increases engagement with the material.

Case studies: real-life examples of the issues covered in your training can be very useful, particularly when drawn from the experiences of your organization or other groups engaged in the same field. The format may vary from short overviews to illustrated step-by-step case studies.

Group work: this format enables direct engagement with the topics under discussion and gives space for individual input. Careful planning is key to group work. Clear instructions about the aim and deliverables of the exercise, good timing, small groups (ideally no more than five people in each) and careful facilitation of their work together is essential.

Buzz group: short, time-limited discussion on a given subject, usually in pairs or smaller ad hoc groups.

Brainstorming: very open group discussion to gather initial ideas on a topic. Be aware of the size of the group and ensure there is adequate facilitation to give space to all.

Role play: as participants are actively involved, this method ensures deeper understanding of an issue and improved skills. It is time consuming and needs careful preparation of roles and tasks.

Simulation games: these can be used as an exercise to change attitudes and practices, using games sourced by the trainer or developed by your own organization.

Recording with playback: this method is commonly used in communication and presentation skills development trainings. It can serve as a good model for building advocacy skills: for example, how to present an issue to a local decision maker or other stakeholders.

Energizers

You may use energizers whenever people look sleepy or tired or to create a natural break between activities. Try to choose games that are appropriate for the age and local context, thinking carefully about games that involve touch or talking about religion or other personal issues.

Things in common

The facilitator calls out a particular characteristic, such as 'speaks more than one language fluently', and people have to move accordingly into different parts of the

Countering cyberhate against Roma

space: for example, all those who speak more than one language fluently should move to a corner of the room. The facilitator calls out more characteristics, such as 'likes football', and people move to the indicated areas. You may call out issues that apply to work, such as 'gave presentation at conference in the last three months'.

Simon says

The facilitator tells the group that they should follow instructions when the facilitator starts the instruction by saying 'Simon says ...' If the facilitator does not begin the instructions with the words 'Simon says', then the group should not follow the instructions! The facilitator begins by saying something like 'Simon says clap your hands' while clapping their hands. The participants follow. The facilitator speeds up, always saying 'Simon says' first. After a short while, the 'Simon says' is omitted. Those participants who do follow the instructions anyway are 'out' of the game. The game can be continued for as long as it remains fun.

Introduction activities

The first moments of the training are crucial to set the mood and create a friendly environment. You may skip the conventional approach of 'let us introduce ourselves in a circle' and use games to get to know each other.

True or false

Everyone writes their name, along with four pieces of information (one of which is not true!) about themselves on a sheet of paper. You may ask them to write professional statements. For example, 'Lilian likes working with children, is a lawyer, has five years of experience working for Coca-Cola and enjoys paintball in her free time'. Participants then circulate with their sheets of paper. They meet in pairs, show their paper to each other, and find out which of the 'facts' is a lie.

Name and adjective

Participants think of an adjective to describe how they are feeling or how they are. The adjective must start with the same letter as their name, for instance, 'I'm Henry and I'm happy' or 'I'm Cecil and I'm creative'.

Role of the facilitator / trainer:

- Sets ground rules.
- Creates a learning-friendly environment.
- Facilitates discussion.
- Teaches the curriculum so that participants gain knowledge and skills.
- Provides opportunities for feedback.
- Adapts to any necessary changes onsite.
- Provides feedback in appropriate, open conversation.
- Leads energizers, role plays, small group discussions and other activities.

3.4 Evaluation

The most obvious and frequent kind of evaluation takes place after training; participants fill out an evaluation form before leaving the site. You can find a sample evaluation form (*Tool 6: Evaluation form sample*) at the end of this chapter. The evaluation aims to find out how useful and relevant the training was for participants and how they will apply the knowledge in their work. In your evaluation form you may pose various questions, which can be both quantitative and qualitative. Evaluation can also take place at the end of each day of training through quick ‘How did it go?’ discussions.

You can also evaluate your own work after the training through a staff meeting with the speakers and facilitators to assess the positive and the negative aspects of the training. This is an excellent opportunity to give feedback to your staff or trainers.

Longer-term, follow-up evaluations conducted three months to a year or more after training are also a possibility.

4 Resources and tools

Useful resources

- Alliance against Antigypsyism, *Antigypsyism – A Reference Paper* (2017): a background resource with a working definition of antigypsyism.
<https://www.getthetrollsout.org>
- ARTICLE 19, *'Hate Speech' Explained: A Toolkit* (2015): in this toolkit, ARTICLE 19 provides a guide to identifying 'hate speech' and how to effectively counter it, while protecting the rights to freedom of expression and equality.
<https://www.article19.org/data/files/medialibrary/38231/'Hate-Speech'-Explained--A-Toolkit-%282015-Edition%29.pdf>
- ARTICLE 19, *The Camden Principles on Freedom of Expression and Equality* (2009): a set of principles and recommendations to promote greater consensus about the proper relationship between respect for freedom of expression and the promotion of equality.
<https://www.article19.org/data/files/pdfs/standards/the-camden-principles-on-freedom-of-expression-and-equality.pdf>
- Committee of Ministers of the Council of Europe, *Guide to Human Rights for Internet Users* (2014): the guide is a tool for internet users to learn about human rights online, their possible limitations and available remedies, with information on what rights and freedoms mean in practice in the context of the internet.
<https://rm.coe.int/16804d5b31>
- Council of Europe, *No Hate Speech Movement*: this youth campaign focuses on engaging young people to challenge hate speech and support human rights online. The initiative has produced a large volume of resources to develop alternative narratives to hate speech.
<https://www.coe.int/en/web/no-hate-campaign/no-hate-speech-movement>

Countering cyberhate against Roma

- Counternarratives.org website, ‘*Counter-Narrative Toolkit*’: this website provides freely available resources for use by any individual or organization looking to create counter-narratives. It is intended as a basic guide for those with little to no previous experience of counter-narrative campaigning, and is by no means comprehensive.
- Dangerous Speech Project, *Dangerous Speech – A Practical Guide* (2018): this guide explains the concept of dangerous speech and outlines how to determine what content fits within this category, with a view to reducing violence as a result.
<https://dangerousspeech.org/guide>
- European Court of Human Rights, *Factsheet on Hate Speech* (2019): a compilation of judgments of the ECtHR related to freedom of expression and hate speech.
https://www.echr.coe.int/Documents/FS_Hate_speech_ENG.pdf
- UN, *Rabat Plan of Action*: this important document provides authoritative guidance to states on implementing their obligations under Article 20(2) of the ICCPR to prohibit ‘any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.’
<https://www.ohchr.org/EN/Issues/FreedomReligion/Pages/RabatPlanOfAction.aspx>
- UN, *Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression* (2019): in this report, the Special Rapporteur evaluates the human rights law that applies to the regulation of online ‘hate speech.’
https://www.ohchr.org/Documents/Issues/Opinion/A_74_486.pdf
- UNESCO, *Countering Online Hate Speech* (2015): this provides a global overview of the dynamics characterizing hate speech online and some of the measures that have been adopted to counteract and mitigate it, highlighting good practices that have emerged at the local and global levels.
<https://unesdoc.unesco.org/ark:/48223/pf0000233231>

Tool 1: Counter-narrative strategy template



STEP 1 Choose a narrative you would like to counter.

STEP 2 Analyse the narrative that you chose.

● Why can this message be considered abusive or hateful for the Roma? ● Does it attempt to tap into a wider anxiety about a particular social issue, for example lack of access to housing? ● Is the message a simple expression of opinion or does it incite hatred against the Roma? ● Does the message have the potential to reach a big audience?

STEP 3 Based on your analysis, elaborate your own counter-narrative strategy.

● What would be the most efficient way to counter the identified hate speech? ● Why? ● How could you discredit and demystify the hateful message? ● How would you include in a message what you are standing for? ● Can you bolster your messages with facts (statistics, research findings, etc.), humour, emotions or positive examples? ● Could you rephrase the message to target a different audience?

You can use other tools and tips from the <http://www.counternarratives.org> website to design your own campaign.

Tool 2: Campaign plan template



STEP 1 Based on the findings of your research using Tool 1, define who would be the target group of your campaign (be as specific as possible).

STEP 2 What are you trying to achieve? Define the goals and the objectives of your campaign. Set something that is achievable and for which you can evaluate the impact. (Goals are the desired end-result and can be quite broad: e.g. to persuade people not to join far-right extremist groups. Objectives are the means to the end, and are specific and tangible: e.g. reach 1,000 people online at risk of radicalization with counter-speech content, and conduct two workshops on countering extremist propaganda.)

STEP 3 Based on the counter-narrative strategy which you drafted, define the most important messages that you would like to send to the target audience. What do you think would influence them most (for example, facts, emotions, positivity or satire/humour)?

STEP 4 Based on the messages and the target audience that you have identified, who is the messenger/voice in your campaign? The messenger delivering the messages to your target audience should be someone whom the audience is likely to trust, be inspired by, or listen to.

STEP 5 Choose the platforms you would like to use for your campaign. Which one is most often used by your target audience? Which one offers you the best possibilities to reach out to your target audience or to disseminate your message?

STEP 6 Based on the target audience and your dissemination strategy, identify the format of your content. What tools would you like to use in your campaign (text, audio, infographics, videos, images, cartoons, etc.)? Which one would be best suited to reach out to the target audience?

Countering cyberhate against Roma

STEP 7 Make a risk assessment before you launch the campaign. Could your campaign backfire? Would your campaign bring about negative effects alongside positive ones? If so, how do they balance? What are the critical factors which could endanger the implementation of your campaign?

STEP 8 Elaborate a timeline for the campaign. Try to identify important milestones and moments where your campaign could get more attention.

STEP 9 Draft your monitoring and evaluation plan. What data are you planning to collect for the monitoring and evaluation of your campaign (for example, how many people have seen your material, has there been any positive reaction to the campaign from the target or other audience, etc.)? The impact is what matters; think of the many ways you can use to assess it.

You can use other tools and tips from the <http://www.counternarratives.org> website to design your own campaign.

Tool 3: Pre- and post-training checklist

minority
rights
group
europe

Course name:

Course date:

Course venue:

Pre-course checklist:

- Reserve room
- Reserve AV (audio-visual) equipment
- Confirm with speakers by email/phone
- Send speakers confirmation letter and agenda
- Schedule caterer
- Send participants initial confirmation letters and pre-course documents
- Order binders and other materials
- Order certificates
- Gather copies of pre- and post-tests and evaluation forms
- Create participant list
- Create evaluation forms
- Order copies of all documents for manuals
- Check AV equipment, microphone
- Make nametags
- Create attendance list to sign (one for each day)
- Set up room
- Put out sign on day of training

Post-course checklist:

- Pass out certificates at end of course
- Collect evaluations and post-tests
- Return room to original set-up
- Summarize evaluation

Tool 4: Training design worksheet

Complete this training worksheet to help you begin designing your training.

1 General topic: In general, what knowledge and skill areas will be the focus of this training?

2 Goals and objectives: What do you want participants to learn during the training?

3 Participant activities: How will participants achieve curriculum objectives? (E.g. group work, lectures, role play)

4 Resources: What resources might the trainer use to help participants accomplish curriculum objectives? (E.g. current research, guest speakers, YouTube videos, discussions, learning from peers, examples from other countries)

5 Evaluation: How will you evaluate the quality and usefulness of the training as well as its implementation?

Tool 5: Online hate speech training agenda sample



Training for civil society organizations on counter hate speech campaigning

Project Name:

Organization name:

Date:

Venue:

9.00–9.30	Introduction and setting the house rules
9.30–11.00	Introduction to hate speech definition and legal framework / Presentation (speaker: name)
11.00	Break
11.20–13.00	Best practices of anti-hate speech campaigns: sharing successful counter hate speech actions / film screening and case study (speaker: name)
12.30–13.00	Plenary discussion about the presented campaigns
13.00	Lunch break
14.00–15.30	Develop your own campaign! / Group work to design and plan one campaign.
15.30	Break
16.00–17.00	Plenary discussion: each group presents their campaign idea and participants discuss them
17.00	Wrap-up and evaluation

Tool 6: Evaluation form sample



Training title:

Project title:

Date and venue:

Question 1: How useful have you found the information provided to you during this training event in your activities? Assess on this scale.

<i>Not useful at all</i>	<i>Somewhat useful</i>	<i>Highly relevant and helpful</i>
←	→	→
<input type="checkbox"/>		

Question 2: Please assess how useful you found the listed parts of the workshop in your activities.

<i>Categories</i>	<i>Not useful at all</i>	<i>Some of it was useful</i>	<i>Very useful</i>	<i>Highly relevant</i>
<i>Add training section name</i>				
<i>Add training section name</i>				
<i>Add training section name</i>				
<i>Add training section name</i>				

Question 3: Please assess the following statements according to your experiences at the training.

<i>Categories</i>	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
<i>I found new opportunities to work with other organizations and colleagues.</i>				
<i>I was satisfied with the meeting environment.</i>				
<i>I will disseminate the information I learned here to my colleagues.</i>				

Question 3: continued...

Categories	Strongly disagree	Disagree	Agree	Strongly agree
<i>I was satisfied with the facilitators and speakers.</i>				
<i>I can apply the information in my work.</i>				
<i>The meeting was a great opportunity to design new campaigns.</i>				
<i>I am more familiar with the legal framework of hate speech in Europe.</i>				
<i>As a result of this training, I feel more confident in my capacity to develop counter hate speech campaigns.</i>				

Question 4: What are the three most important things you learned during this training?

1	
2	
3	

Question 5: Do you think you will have the opportunity to utilize the training skills you have practised during this workshop within the next six months? If so, please describe how.

Question 6: Any recommendations for the next training?

Notes

- 1 Alliance against Antigypsyism, *Antigypsyism – A Reference Paper*, 2017.
- 2 Hamelmann, M., *Antigypsyism on the Internet*, Scan project, 2018.
- 3 Theorized for the first time by Kimberlé Crenshaw in the late 1980s, intersectional discrimination recognizes the distinct experiences of victims who face discrimination as a result of the intersection of multiple grounds, such as gender and ethnicity. These areas of discrimination operate and interact with each other in such a way as to be inseparable.
- 4 International Criminal Tribunal for Rwanda (ICTR), *Prosecutor v. Akayesu*, Case No. ICTR-96-4-T, Judgment (Trial Chamber), 2 September 1998, para. 557.
- 5 ICTR, *Prosecutor v. Nahimana, Barayagwiz and Ngeze*, 3 December 2003, ICTR-99-52-T (Trial Chamber).
- 6 International Criminal Tribunal for the former Yugoslavia (ICTY), *Prosecutor v. Naletili and Martinovi*, Case No. IT-98-34-T, Judgment (Trial Chamber), 31 March 2003, para. 60.
- 7 Article 19, *The Camden Principles on Freedom of expression and Equality*, April 2009, Principle 12.1.ii.
- 8 *Ibid.*, Principle 12.1.i.
- 9 *Ibid.*, Principle 12.1.iii.
- 10 UN, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, A/74/486, 9 October 2019. The Rapporteur illustrates the risks linked to unregulated online hate by citing Facebook's failure to address incitement against Rohingya Muslim community in Myanmar.
- 11 UN, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, A/HRC/35/22, 30 March 2017. The report lists several examples of practices considered incompatible with international standards protecting freedom of expression as internet shutdowns in their various forms (see paras. 8–16) or collusion with governments complying with censorship or surveillance requests (paras. 30–33).
- 12 Human Rights Council (HRC), General Comment No. 34: Article 19 Freedoms of opinion and expression, CCPR/C/GC/34, 12 September 2011, paras 50–52; Committee on the Elimination of Racial Discrimination (CERD), *General Recommendation No. 35: Combating Racist Hate Speech*, CERD/C/GC/35, 26 September 2013, para. 35.
- 13 HRC, *Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred*, A/HRC/22/17/Add.4, 11 January 2013.
- 14 UN, 30 March 2017, *op. cit.*; UN, 9 October 2019, *op. cit.*
- 15 Council of Europe, *Recommendation No. R (97) 20 of the Committee of Ministers to Member States on 'Hate Speech'*, adopted 30 October 1997.
- 16 ECRI, *General Policy Recommendation No. 15 On Combating Hate Speech*, 8 December 2015.
- 17 ECtHR, *Erbakan v. Turkey*, 6 July 2006, para. 56.
- 18 ECtHR, *Handyside v. the United Kingdom*, 7 December 1976, para. 49.
- 19 ECtHR, *Vejdeland and others v. Sweden*, 9 February 2012.
- 20 ECtHR, *Seurot v. France*, 18 May 2004.
- 21 EU, Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law.

- 22 European Parliament, *European Parliament Resolution of 14 March 2013 on Strengthening the Fight against Racism, Xenophobia and Hate Crime* (2013/2543(RSP)), para. 5.
- 23 Special Rapporteur on freedom of opinion and expression, Press release, 21 October 2019.
- 24 See e.g. ECtHR, *Ahmet Yildirim v. Turkey*, 18 December 2012.
- 25 HRC, *General Comment No. 34, Article 19: Freedoms of Opinion and Expression*, CCPR/C/GC/34, 12 September 2011, para. 43.
- 26 European Court of Human Rights, Factsheet on Hate Speech.
- 27 HRC, *General Comment No. 34*, op. cit.
- 28 CERD, *General Recommendation No. 29, On Article 1, para. 1 of the Convention (Descent)*, 2002.
- 29 UNESCO, *Countering Online Hate Speech*, 2015, available at <https://en.unesco.org/news/unesco-launches-countering-online-hate-speech-publication>
- 30 Council of Europe, *Additional Protocol Concerning the Criminalization of Acts of a Racist and Xenophobic Nature Committed through Computer Systems*, European Treaty System No. 189, Strasbourg, 28 January 2003.
- 31 Convention on Cybercrime, ETS No. 185, Budapest, 23 November 2001.
- 32 As from April 2019, Google+ is no longer available for consumer (personal) and brand accounts.
- 33 European Commission, 'Countering online hate speech – Commission initiative with social media platforms and civil society shows progress', Press release, 1 June 2017.
- 34 European Commission, 'Code of Conduct on countering illegal hate speech online', Factsheet, February 2019.
- 35 OHCHR, *Guiding Principles on Business and Human Rights*, Principle 11, 2011.
- 36 UN, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, A/HRC/35/22, 30 March 2017, para. 1.

working to secure the rights of
minorities and indigenous peoples

Countering cyberhate against Roma: A toolkit for civil society organizations and activists

With far-right groups becoming increasingly prominent and xenophobic discourse on the rise, in recent years there has been a resurgence of hate speech against Roma. Across Europe, Roma communities continue to suffer extensive online hate speech against them. Counter-narrative campaigns are therefore urgently needed, and the public participation of Roma themselves, in particular through social media platforms, is fundamental.

Countering cyberhate against Roma: A toolkit for civil society organizations and activists seeks to support these efforts by providing activists, community members and trainers with a valuable resource to guide them through every step of the process. Beginning with a discussion of the international and European legal frameworks on hate speech, it also presents strategies on how to develop counter-narratives and guidance on how to organize training events.

ISBN: 978-615-81504-0-8 / 978-615-81504-1-5 (pdf)

Minority Rights Group Europe c/o NEKI, Ulloi Utca 68. 11.em. 15. H-1082 Budapest

Tel +44 (0)20 7422 4200 **Fax** +44 (0)20 7422 4201 **Email** minority.rights@mrgmail.org

Website www.minorityrights.org

 [www.twitter.com/minorityrights](https://twitter.com/minorityrights)  www.facebook.com/minorityrights

Visit the **Minority Voices Newsroom** for stories from minorities and indigenous peoples around the world www.minorityvoices.org

#FreedomFromHate