

## 2) RESPONDING TO HATEFUL CONTENT ONLINE



**“** *While technical access is a factor, there is more: we must succeed in establishing ground rules for communication in the digital public sphere that enable minorities and vulnerable groups to exchange views and make themselves heard.*

*Dr Ellen Ueberschär, Co-president of the Heinrich-Böll Foundation, speaking at WACC's 50th Anniversary Symposium, May 2018*

## RESPONDING TO HATEFUL CONTENT ONLINE

How can we deal with hate online? There are, broadly speaking, three main areas of action:

- a) Advocating for appropriate legislation and voluntary codes of conduct
- b) Supporting education and media literacy efforts
- c) Engaging in counter-speech and other counter-actions

While all three approaches are necessary, the key is to adapt the strategy depending on the different manifestations of hate speech and the changing nature of digital technology. Familiarize yourself with these approaches and adapt them for use in the situation at hand.

### Legislation and voluntary codes of conduct

Binding legislation and voluntary codes of conduct are based on the belief that social media companies should be held responsible for content posted and shared on their platforms. Hate speech legislation is, however, a very challenging matter, as demonstrated by the German NetzDG law case study (see page 13).

Deciding what role social media companies should play in moderating content and hate speech is complicated. When they are expected to set their own standards and police their own users, they may lean towards removing flagged content out of pressure to comply with national laws, to avoid fines, or simply to seem worried about the spreading of hateful content on their platforms.

This in turn may lead to inadequate protection of freedom of expression as the 'private censorship' of these platforms may be more restrictive than that imposed by international human rights law. ARTICLE 19, a civil society organisation specialising in freedom of expression, analysed the community standards of Facebook and Twitter, and found that both fall below international standards on freedom of expression, especially in regards to hate speech. Lack of transparency and accountability in the removal process and lack of appeal provisions for users whose content is deleted are also concerns.

In May 2016, the European Commission worked with Facebook, Microsoft, Twitter, and YouTube on a voluntary Code of Conduct on Countering Illegal Hate Speech Online, which was later additionally signed by Google, Instagram and Snapchat. The companies agreed to evaluate notifications by users within 24 hours and to remove content deemed illegal. Progress on the code is monitored through periodic exercises and evaluations. The most recent evaluation, published in February 2019 by the European Commission Directorate-General for Justice and Consumers, states that "IT companies are now assessing 89% of flagged content within 24 hours and 72% of the content deemed to be illegal hate speech is removed, compared to 40% and 28% respectively when the Code was first launched in 2016. However, companies need to improve their feedback to users."

Tackling hate speech with legislation and guidelines, whether binding or voluntary, is one way to attempt to deter the problem. However, the very nature of hate speech on the internet, its volume, its reach, and its transnational nature, are considerable complications. Legislation on hate speech is one avenue that can be pursued. At the same time, other measures aimed at addressing the issue are equally important.

### Education and media literacy

Education is key to counter hate speech and hateful online content. As a preventive tool, education and awareness raising are fundamental in increasing our understanding of how hateful content spreads on the internet, and how we can double-check the information we find. The more we are able to do so, the less hate speech will have a free pass.

Media literacy is particularly important in addressing and countering hateful online content. Essentially, it is about developing critical thinking and "critical clicking". It is a conscious use of social media, which allows individuals to identify and question hateful content, to understand the prejudices underneath it, and to develop arguments to confront it.



As the nonprofit European Association for Viewers Interests (EAVI) puts it, media literacy is not about how to technically use media and social media platforms, but about how to “critically evaluate and analyse numerous sources of information simultaneously. This skill requires traditional literacy, reasoning, social injunction, and the ability to decipher symbolic and cultural codes and conventions.”

One of the key projects in this regard is carried out by the Council of Europe No Hate Speech Movement, a youth campaign to combat hate speech and promote human rights online. Also, in 2019, the European Commission initiated the EU Media Literacy Week to highlight the importance of media and information literacy as a key factor enabling digital citizens to take informed decisions, online and offline.

## Counter-speech and other counter-actions

*Counter-speech* is a term that includes all activities aimed at responding to hateful content online. If you see a violent or vulgar comment and engage with it, you are doing counter-speech.

Counter-speech can also expand to become a counter-narrative or campaign. These are especially useful if you are working in a group or for an organisation. Both counter-narratives and campaigns are larger-scale activities, and require more planning, time, and resources. The suggestions for individual action also apply to those handling a group’s social media accounts. If you are interested in developing a counter-campaign, you will find more information in the resources section of this report.

## WHAT DO YOU DO WHEN YOU RECEIVE A HATEFUL COMMENT ON SOCIAL MEDIA?



**Dóra:** When I receive a hateful comment on social media, I ask myself two questions: Do I have time to deal with this now? And, is there an actual question or comment to reply to?

If it is a pointless attack that is not leading to a conversation, I ignore it. This is also the case when the comment is made from a fake account.

On the other hand, if I can find some sort of contribution underneath the vulgarity, I usually engage. It may be useless for the person who wrote the comment, but it may have an impact on those who read it.



**Anna:** When I see a vulgar or hateful content or comment, I read it first, with attention. If the comment has a potential for discussion, I reply, even if I do not agree with the views expressed.

On the other hand, if there is no potential at all for discussion, I delete the comment and/or block the user. I do not want to encourage hate, and some people are just looking for a reaction of any kind; this is when I delete and disregard the comment entirely. But it is always very important to be aware of our own biases: I will not delete a comment simply because it expresses views that I am opposed to.

If I see that someone is using a fake account, I block them. I run my website under my real name and my personal social media accounts, and I do not think it is fair to me or to the other participants to have fake accounts in the discussion.



## WHAT ADVICE DO YOU HAVE FOR PEOPLE WHO ARE NEW TO COUNTERING HATE SPEECH, BUT WOULD LIKE TO START?



### Annegret:

- Stay calm and friendly when correcting inaccurate information.
- Make your message short, and include a link to the correct information.
- If you are not an expert in dealing with hateful content, it is a good idea to take a break and then reread your message one more time before you publish it. This will help to ensure that it is informative rather than heated.
- Try not to get angry; your matter-of-fact, factual comment will be much more effective.

## #IAMHERE. CREATING A CIVILISED SPACE FOR DISCUSSION ON SOCIAL MEDIA.

Swedish journalist Mina Dennert wanted to improve the tone of comments on Facebook and to establish a respectful dialogue on the platform. In 2016, she created the Facebook group **#jagärhär**, **#IamHere**, which currently has almost 75,000 members.

When a group member encounters hateful content in the comments section on Facebook, they respond and call others into the conversation with the hashtag **#IamHere**. The aim is to insert facts and reasonable viewpoints in the conversation, so that other social media users will see a balance of opinions.

**#IamHere** is not about changing the political views of society. Indeed, it welcomes members from the most diverse political views. It is, rather, about changing the way we debate on the internet. Its intention is to promote respect and civility to stop hate and disinformation, improve the debate, and ensure that more people are able to express their opinions without fear, so long as these views are not violent or threatening.

The **#IamHere** network has spread all over the world and has thousands of volunteers. You can find more information about your country group here:

<https://www.jagarhar.se/kolumnen/the-iamhere-network/>

The **#IamHere** network has also created a list of tips for engaging in counter-speech on Facebook to create a civilised dialogue with those we do not agree with. Tips include the following:

- Assume the other person means well, and listen to what they have to say.
- Try to find a common ground, and to understand if words mean the same for both of you—maybe they do not!
- Do not attack, and keep your feelings under control.

Read the full list at:

<https://www.jagarhar.se/kolumnen/best-practices-counter-speakers/>

## IF YOU ARE TAKING ACTION AS AN INDIVIDUAL:

Your decision to engage in counter-speech activities, and how much time you dedicate to this work, is **entirely up to you**. Countering hateful content can be difficult, so it is important to be mindful of the tips below, to ensure that you remain as safe and healthy as possible:

- **Do not work on counter-speech activities alone.** Make sure that you are in a supportive environment, whether online, offline, or both.
- **Do it for only a limited amount of time every day/week.** Limit the amount of time you engage in countering hate speech and take breaks.
- **Your mental health is the most important parameter.** If you feel that you have had enough, or that you cannot take it anymore, stop. Get up and do something pleasant. Only come back to counter-speech activities when you feel safe and grounded.

Although counter-speech is often advocated as the best way to deal with hate speech, it is important to be aware of the power imbalance between those who post hateful content, and those who engage in counter-speech. Engaging in counter-speech requires much more mental and emotional effort than posting or sharing something hateful, which is why so many people remain silent in the face of hate. Furthermore, harassment of a particular group may prompt other members of that group to remain silent, for fear of facing the same harassment.

This is also why engaging in counter-speech activities—as scary and as unrewarding as that may be—is so important. Especially if you are not from one of the targeted groups, you can remind the authors of hateful messages that those who are being targeted are also human beings. In this way, you can help decrease the feeling of isolation that targeted people frequently suffer from and you can promote a more respectful online debate.

Many times, the hate aimed at one specific group or topic may actually be about something else. Migrants and refugees are often targeted because of their visibility; however, migration is

sometimes a smokescreen used to get attention, while people really want to talk about something else. In these cases, it is useful to identify the real issues behind the nastiness and vulgarity and address these issues rather than the emotions expressed through the smokescreen of migration.

**Remember:** Even if your objective in engaging in counter-speech is not achieved (for example, if the author of the hateful content does not remove it or does not apologise), there is a much broader audience of silent readers who may place great value in your counter-speech. Do not be discouraged!

If you feel ready to engage with hateful online content, the tips on the following pages may be useful for you. As every expression of online hate is different, every possible counter-action is also different. Use the suggestions to find a way to respond that works for you.

**Remember:** It is very important throughout this process to be mindful of your own biases. Not everyone who disagrees with you is a hater. Be mindful of this, and be open to different viewpoints. You may learn something new!

## IF YOU ARE THE TARGET OF HATE:

- **Evaluate** whether it is worth engaging. Is the comment nothing but hateful? Then it could be better to ignore, delete it, or even block the author. Is there an actual message or question in the comment? Then you may want to respond.
- **Remind** the author of the consequences of their words. This may be the harm their speech causes you or others. The fact that this content will be visible for an indefinite time may also affect the writer by negatively impacting their relationships or future employment opportunities.
- **Report** the hateful comment to the social media platform. This may not have the immediately desired effect, but the more hateful content is reported, the more we can measure and understand it. Plus, this experience may educate the author and have positive longer-term effects.
- **If** you know the author of a hateful comment or post, reach out to them privately and let them know that you are uncomfortable with what they wrote and why. Doing this before debating it publicly may give the author a moment to reflect rather than feeling attacked and retreating into their initial hateful position. This will give them a chance to re-think their post and maybe even edit or delete it.
- **Do** not use hateful or vulgar tones in your replies. Replying to hate with hate only generates more hate, and that may be exactly what the author wanted in the first place.
- **Use** normal language, the same as you use when speaking with friends. When we bring normality back into the discourse, we can establish a human connection and may initiate a dialogue.
- **Speak** to the underlying objective of the comment, not to the overt negative narrative.
- **Humour** may defuse the situation.
- **Use** visuals in your replies. An image or short video can sometimes go much further than a written reply.
- **Ask** for help. It is fine if you do not want to read hate anymore. You can ask someone else to go through comments for you, deleting ones that are pointless or just hateful, so you do not have to read them.
- **Even** if you are not able to change someone else's mind, remember that there is a vast audience of passive social media users. These are people who do not engage in the conversation but read the comments. Your response to a hateful comment may not be useful to the one who wrote it, but it could have an impact on others who read your reply.



**Anna and Karol:** We have a good working relationship with the police in Krakow and once, when we received a comment saying that 'migrants bring crime to Poland so they should not be accepted', we replied saying 'What are your concerns about safety and security in Krakow? The police can answer!' and tagged the police into the comment. Of course, this only works if the police actually reply.

## IF YOU ARE WITNESSING SOMEONE ELSE BEING TARGETED BY HATE:

- **Be** supportive, both of the person or group targeted, and of counter-speakers. Engaging in counter-speech is difficult, so an alternative to replying directly is to support those who do speak out. “Like” their comment, share their post, or write some words of support. This will not only make the counter-speaker feel supported, it will increase the reach of the post/comment.
- **Report** the hateful comment to the social media platform. This may not have the immediate desired effect, but the more that hateful content is reported, the more we can measure and understand it. Plus, this experience may educate the author and have positive longer-term effects.
- **Remind** the author of the consequences of their words. This may be the harm their speech causes you or others. The fact that this content will be visible for an indefinite time may also affect the writer by negatively impacting their relationships or future employment opportunities.
- **If** you know the author of a hateful comment or post, reach out to them privately and let them know that you are uncomfortable with what they wrote and why. Doing this before debating it publicly may give the author a moment to reflect rather than feeling attacked and retreating into their initial hateful position. This will give them a chance to re-think their post and maybe even edit or delete it.
- **Change** the tone of a hateful conversation to a more empathetic one. For example, find some common ground with the writer—that may have nothing to do with the topic of the hate speech.
- **If** there are threats of violent actions, and they seem credible, inform the police.
- **Do** not use hateful or vulgar tones in your replies. Replying to hate with hate only generates more hate, and that may be exactly what the author wanted in the first place.
- **Speak** to the underlying objective of the comment, not to the overt negative narrative.
- **Humour** may defuse the situation.
- **Use** visuals in your replies. An image or short video can sometimes go much further than a written reply.
- **Join** organised counter-speech activities, such as the Facebook group #IamHere.
- **Even** if you are not able to change someone else’s mind, remember that there is a vast audience of passive social media users. These are people who do not engage in the conversation but read the comments. Your response to a hateful comment may not be useful to the one who wrote it, but it could have an impact on others who read your reply.

### WHAT IS IMPORTANT FOR YOU WHEN DEALING WITH HATEFUL CONTENT?



**Annegret:** What is important to me is to remember that people are more likely to spread hate on social media than face to face, and that it would be too much to expect to be able to change their minds. I think that what we are trying to do in responding to hateful content is for the benefit of the bystanders, the silent readers—to show them that there is a proper way of dealing with hate, and to set the right example.

Sometimes just stepping back is helpful. There is really no point in having a long discussion; if the aim is to talk to a third

## IF YOU ARE WORKING FOR AN INITIATIVE OR AN ORGANISATION:

- **Have** clear Terms and Conditions or Engagement Guidelines on your social media or website. The guidelines should clearly state what type of language is permitted and what type of content will not be tolerated. This way, you will be able to link back to the guidelines whenever you need to reprimand a user for their conduct or delete their comments altogether.
- **Develop** a strategy with ready-made actions and answers to be used on different platforms.
- **If** you delete a comment or block a user, be sure to post a standard explanation and refer to your engagement guidelines.



### WHAT IS IMPORTANT FOR YOU WHEN DEALING WITH HATEFUL CONTENT? (cont.)

reader, then it is also important to know that they will not read beyond a couple of messages anyway.

One important aspect to take into consideration is how much time you have. There is always a balancing exercise between how likely it is that other people would see the hateful comment, and would be influenced in a negative way, and how much time it would take you to deal with it. There is a need to maximise the impact while at the same time minimising the resources needed.

- **Take** a screenshot of the post and username before deleting, for your records.
- **In** case of someone posting misinformation, post a link to correct information and state that the previous information is incorrect.
- **Always** be polite and friendly.
- **Be** sure you have a clear understanding of organisational policy and practice in determining what content is allowed, and that you have sufficient human resources to be actively and responsibly engaged.
- **Work** on creating a counter-narrative: What is the positive message that your organisation/initiative wants to promote? Once you have developed your counter-narrative, you can always come back to it when you reply. This way, you are not just replying to a hateful comment, but actively promoting your own message. Promoting your own narrative may also be more effective than simply countering the hateful content with, for example, fact-checking.
- **If** you know a post will spark hate-filled responses, consider settings that will allow you to moderate the comments before they are visible.
- **Where** possible, bring in trusted partners who can support your statement, for instance, by tagging or linking to external sources of information.
- **If** you have the means, consider engaging with a well-respected personality in your context. Often, famous people can become speakers on a topic, and influence a vast public who would not otherwise pay much attention to your organisation or initiative.
- **Even** if you are not able to change someone else's mind, remember that there is a vast audience of passive social media users. These are people who do not engage in the conversation but read the comments. Your response to a hateful comment may not be useful to the one who wrote it, but it could have an impact on others who read your reply.

## IF YOU ARE TIRED OF SEEING SO MUCH HATE ONLINE AND WOULD LIKE A NICER, SAFER ONLINE ENVIRONMENT FOR ALL:

- **“Like”** and share positive news and resources that promote inclusive communities, tolerance, respect, and dignity.
- **Ensure** that your own online posts are models of good practice. Be considerate, respect privacy, and share news and information only from trusted and/or verified sources.
- **Be** supportive of people who become targets of hateful content. “Like” their original comment; share their post; write some words of support. This will not only make the person targeted feel supported, it will increase the reach of their post/comment.
- **Develop** some standard replies to potential hateful content so that you can respond more quickly and without as much drain on your energy and emotions.
- **Try** to find common ground, both online and offline, between people or groups of people who are divided. The common ground may have nothing at all do with the topic you are discussing. For example, you may be discussing the NGO rescue boats in the Mediterranean with someone whose views are completely opposed to yours. But what if you both have a keen interest in gardening? Finding the common ground may help establish a basic level of trust, which can be used to move the dialogue forward.

## HOW DO YOU HANDLE HATE ON SOCIAL MEDIA?



**Annegret:** The WCC has developed a strategy with some ready-made answers for certain cases. This has evolved over the years. First, we started dealing with hate on a case-by-case basis, then when the same issue would come up, we would deal with it in the same way, and at some point, this became a document so that other people could use it. The pre-made answers are also a way to be transparent about the way we operate.

So it depends on whether there is a question being asked, or an inaccurate statement about what WCC is doing, or a clear case of a hateful message. In the case of a clear hateful message without any other content, I block the person, delete the comment and write a new comment saying that we deleted the content and explaining why.

This is the message that I would post in this case:

*“Note: some comments had to be removed from this thread due to the WCC’s policy to remove comments containing hate speech or inciting violence. The WCC welcomes comments; however, it reserves the right to delete comments that are vulgar, defamatory, clearly spam (including self-promotion), or in general, not contributing to the ongoing discussion.”*

I want people who saw the negative comment to know that they do not have to pay attention to it because it is just hate. The idea is also to educate people on how to behave correctly, and on what constitutes problematic behaviour on social media. I also always take a screenshot of the comment and username when I block someone and keep it in the records.



## HOW DO YOU HANDLE HATE ON SOCIAL MEDIA? (cont.)

On Twitter you cannot really delete a comment, so I try to reply with the ready-made answer and/or block the person. Sometimes, however, it feels that by replying to a comment, I am actually giving it more visibility, and a platform. That happens especially on Twitter, where if you reply, you make a comment more visible. So in many cases I do not reply from the WCC main account, but instead use my personal account which has fewer followers.

As a consequence, this has also made me the target of hate. At that point, it's a personal discipline not to look at the hate in the wrong moments, to preserve some rest time. You can mute a person, so you do not see their comments for a certain amount of time.

In the case of someone posting inaccurate information, I provide a link to the correct reading material and state that the previous information is incorrect. (We have a pre-formulated answer for this case also.) In this case, often people do follow up, mostly because they do not believe that they are wrong. The discussion can then progress, though if there is hateful content in the follow-up, I would block and delete.

Sometimes it is not so easy to decide when to delete or to respond to a comment, so we have to discuss it case by case.

If someone makes multiple comments and they tend to be hateful, then it is better to block the person, because it saves time. I have to say that banning is quite an effective way of reducing the hate, of not being a platform for hate. It has happened that people come back under different accounts, but only rarely.

In the beginning I also wrote messages to people directly, but then it became impossible because of the workload. Now I only do that if there is a starting point that I can use and I think there is a possibility to educate the

person. I hope that by engaging this way there is a little seed planted in the brain, but I have not checked the response, so I could not say if this works or not.

On Facebook, because we already know that some topics are going to provoke incendiary reactions, the comments are hidden by default according to a long list of keywords. I then read the comments one by one and decide if they can be published. It is very time-consuming work, but we opted for this rather than receiving a lot of hate and having to delete it. We regularly post this comment to make this practice transparent:

*"Please note: The WCC welcomes discussion; however, it reserves the right to delete comments that are vulgar, defamatory, clearly spam (including self-promotion), or in general not contributing to the on-going discussion. As some topics tend to attract hateful comments, and the WCC doesn't wish to be used as a platform for spreading hate, the settings of this page are such that some comments will be hidden based on keywords until they can be reviewed by a moderator. This thread has attracted a high number of comments, some of them not very clear or unlikely to advance the discussion in a positive way. In the interest of maintaining a focus on positive contributions, we regret not to be able to react to such comments. Thank you for your understanding."*



## HOW DO YOU HANDLE HATE ON SOCIAL MEDIA?



**Anna and Karol:** We moderate every single comment on Islamista, which means we are able to see them before they are published. Sometimes we hide an aggressive or vulgar comment and write a personal message to the author letting them know that if they want their post published, they need to edit it to comply with our standards of use.

Sometimes this strategy works, and the author edits the comment and then we publish it; sometimes they delete the comment themselves. Other times, they do not change anything, so we delete the comment.

On the blog we can also track the authors of the comments, so if the comment is particularly hateful, we may even write to the author something like: "Hate speech on the internet is trackable, I can go to the police with your IP address."

Moderating the comments and explaining why certain comments are not tolerated has also been an educational experience for the audience. The group of readers is in a way "educating" itself, is understanding what hate speech is, what type of language to use, and is increasingly able to have a civilised discussion.



**Dóra:** When I was writing for a collaborative blog, we adopted a supportive strategy. Instead of reading the comments under our own articles, we read the comments under our colleagues' articles. That way, we could filter out all the comments that were simply hate or personal attacks, and leave the ones which were genuine contributions or questions for the author to reply to. We

adopted this strategy after realising it was too disturbing for one person to handle it alone. Feeling supported and not isolated is really important when dealing with hateful comments and content.

It is also important to have a public engagement policy to explain what type of content will be tolerated and most importantly, which content will not—such as personal attacks, vulgar language, and hate speech. It is important to have this policy because then you can always refer to it. If you decide to delete or hide a comment, you can link the policy below, so the author of the comment can understand why the comment was removed.

If I know personally the person who wrote a hateful comment, I reply directly to them—just as I would in an offline interaction with an acquaintance or friend.

I also always try to bring the discourse back to reality. Migration issues get a lot of attention, but there are actually very few migrants in Hungary. So I try to ask commenters about other issues where they live: in a country where millions live in poverty, surely there are bigger problems than migrants?

I do not like to ban people from the platforms. I believe that we all live together as a society, off- and online, and I treat the people who post hate online as people who are looking for answers or help. I try to see the people behind the comment, and banning people does not bring anything positive. Our society is already so divided; we should try to build bridges instead.



## CAN YOU SHARE A CASE IN WHICH YOU WERE ABLE TO TURN THE HATEFUL COMMENTS AROUND?



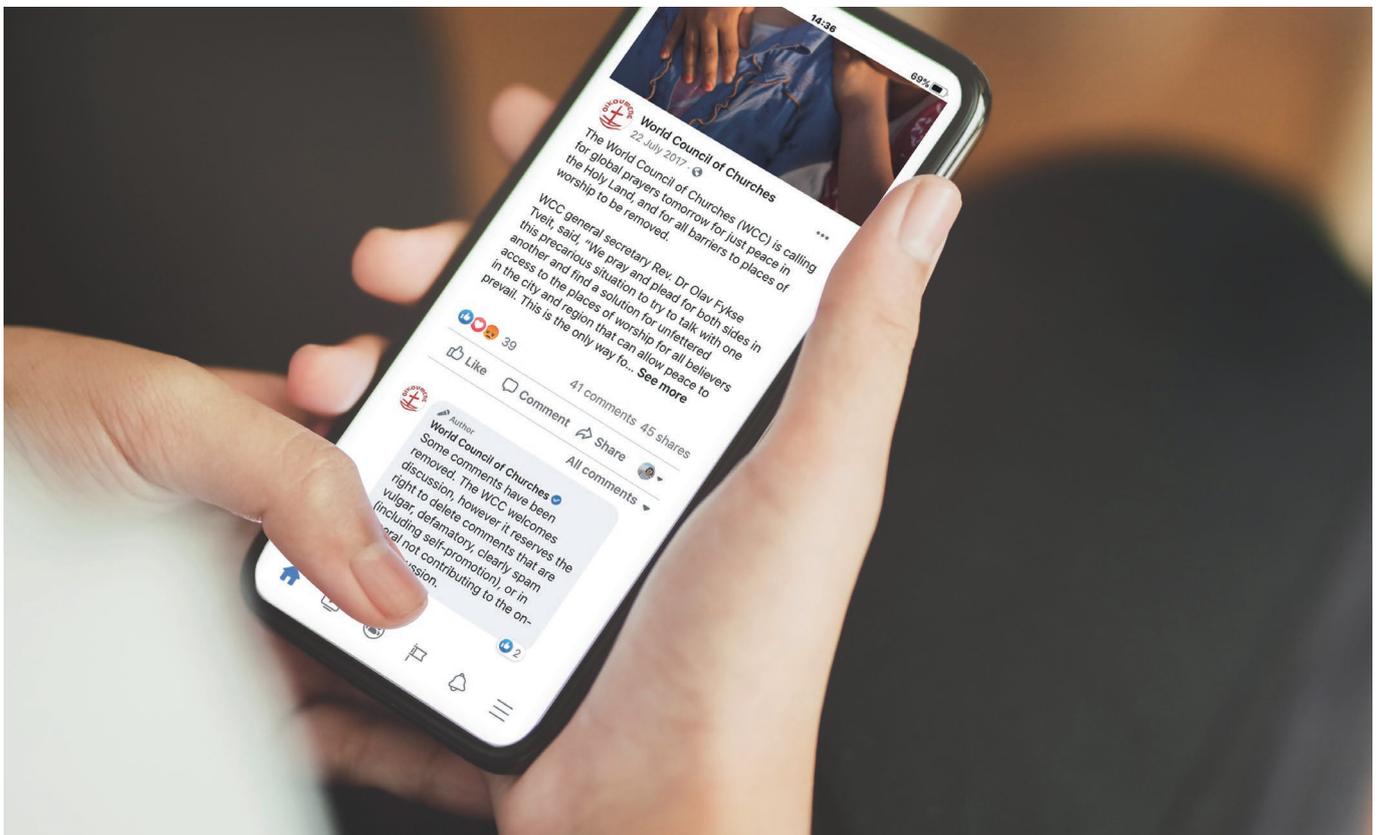
**Dóra:** I once shared an article from my personal Facebook account about the fact that my hometown in eastern Hungary was founded by Slovak migrants, and about the value of migrants and migration for our society. The original settlers of the town are very proud of their Slovak origins. I did not write the story, I simply shared it, but that did not seem to matter.

I received many comments, especially from acquaintances from the same town, disbelieving the article and questioning it. People were writing, "We can see on our public service media what migrants are, what they are doing in Western Europe, so how can we possibly be the same? How dare you claim this?"

I replied by sending private messages to the commenters, because I knew them personally. To one of them, I sent the definition of the word *migrant* from Wikipedia. After the commenter read it, they got back to me saying "You are right. I am a migrant too."



**Annegret:** Sometimes it actually happens that a person says "Hey, I really didn't know that, thanks." In any case, even if that does not happen, when I reply, it is mostly for the benefit of others who read. And I think it is always important to stay polite and friendly and always assume the best of the other person.



## Counter-narrative campaigns

For those who are developing strategic counter-narrative campaigns, the Radicalisation Action Network (RAN) Centre for Excellence has developed very useful guidelines for monitoring and evaluation, including tips like the following:

- Make an evaluation plan in advance of the campaign.
- Use realistic indicators.
- Monitor the campaign and adjust as necessary.
- At the end of the campaign, evaluate your success in reaching your goal.

The RAN Centre also created a checklist for planning a counter-narrative campaign, according to its GAMMMA+ Model. The key elements are: Goal, Audience, Message, Messenger, Medium, and Action, plus Monitoring and Evaluation. Below are some of the essential points. (For full list see [https://bit.ly/RAN\\_GAMMMA](https://bit.ly/RAN_GAMMMA))

- Effective communication campaigns have goals that are clear, realistic, and measurable.
- The promoted messages are relevant and the target audience considers the messengers credible.
- The campaign works with the target audience's preferred medium or online platforms, and is also present when the audience communicates offline.
- Narrative campaigns in the form of monologues are unlikely to meet the needs of an audience that wants to talk, or is upset or outraged about a real or perceived injustice.
- Campaigns should offer a call to action for those wishing to become involved in the issue at hand, which will facilitate monitoring and evaluation.
- Campaigns aiming to change minds and behaviours offer opportunity for sustained dialogue (both online and offline) with those in their audience who wish to talk.
- Campaigns which ensure they have monitoring and evaluation components in place from the start can then adjust ongoing activities if needed, and once completed, can learn whether they had the desired impact.

- Campaigns that produce a constant stream of content for their target audience to interact with increase their chances of having an impact. Authenticity and quantity are more relevant than technical quality.
- Alternative narratives promote positive alternative perspectives, courses of action, and role models, and foster critical thinking. Counter-narratives, which aim at debunking extremist propaganda, should only be directed at a well-researched and understood audience which is already engaged with extremist content.
- Prepare for success and remember to take into consideration all security risks for your organisation and partners.

More information on monitoring & evaluating counter- and alternative narrative campaigns is available at [https://bit.ly/RAN\\_MEcampaigns](https://bit.ly/RAN_MEcampaigns)



## THE DIFFICULTY OF EVALUATION



**Timo:** Evaluating the impact of the workshops is very challenging. How can you say that a workshop has

been successful? Is it about having people replying more frequently to hateful comments? Is it about the quality of the replies? How can this quality be judged?

At the end of the workshops, we always ask participants if they will do anything differently in the future. The response we usually receive is mixed. Some still find it very difficult to impossible to reply to hate online. Others say that they will engage more, not so much for the haters, but for all the passive communicators present online.

I do not see a major problem in being unable to evaluate the workshops' impact. What we are trying to do is not so much to educate people to do something right, as to empower people to have conversations online—to engage with others. It is not just about hate speech, it is about how we see society around us. Why do we ignore certain issues? This is something I would like to be able to focus on more in the future: our positions of privilege and how we relate to others in society.

## Evaluating the impact of counter-strategies

Evaluating the impact of strategies to counter hateful online content is a challenging enterprise. As social media constantly changes and evolves, it is difficult not only to keep track of it, but also to evaluate the effectiveness of counter-strategies.

How can you evaluate the effect of your actions on social media? Many times, it may feel like you cannot. However, it is also possible that you will see an immediate result: someone deleting their hateful comment after you have responded to it for example. If you keep engaging, especially if you are working in the context of an organisation, you may feel that the climate of comments is improving in the long term, and that the readers are understanding why certain words and expression are hateful and should not be used. And that may be a success in itself!



## THE IMPACT WE DESIRE



**Annegret:** We engage in counter-speech also because we want to educate people on how to behave correctly online and help them understand what constitutes problematic behaviour on social media.

## CONCLUSION



## CONCLUSION

The presence of hate on the internet, and its increasing volume and reach, are facts of our everyday life. It may feel like we have little control over this; however, how we choose to deal with it is entirely up to us.

In a world that is increasingly divided, where people retreat into their filter bubbles and refuse to have conversations with those who do not share their views, there is a strong and urgent need to engage. We need to break down the divides we see on social media and in life, and talk with each other. The risks involved in ignoring division and hatred are extremely high. Consequences manifesting across the world include populist leaders taking charge and spreading hateful messages against demonised communities.

Migrants and refugees are one such community, omnipresent in European politics and news, and consequently, easily targeted on social media. The experiences of migrants and refugees as targets of online hate is also transferable to many other marginalised communities.

The fact is that we are all human beings—those who spread hate and those who are the targets of that hate. Reflecting on our commonalities, looking for what unites us, allows us to start having conversations with those with views diametrically opposed to ours.

Our living together depends on our ability to respect each other, and to be able to disagree with each other without using hateful, vulgar, or threatening language. There is always a person on the other side of a hateful comment. We most likely would not purposefully hurt that person in real life. Why should we do that on the internet?

This report has shown that there is not one simple way to hate on the internet. Each case is specific, and can be addressed in many different ways. However you choose to respond, your engagement in the matter is important. The number of haters out there is small in absolute terms, but they are very vocal. Using our voices to support the causes we believe in, and the targeted groups we work with, helps to demonstrate that haters are a minority.

This is how we move away from being silent bystanders. This is how we confront online hate. This is how we bring respect and civility into the dialogue and break down the social media divides.