

## 1) UNDERSTANDING HATE SPEECH



**“** *Hate speech, threats and attacks can make it effectively impossible for people being attacked to have their voices heard or make the environment so toxic that they would not wish to be involved. A free-for-all can create so much noise that it is all but impossible to hear much of what is said. As for so many areas . . . freedom of speech is not a simple subject at all.*

*Paul Bernal, The Internet, Warts and All: Free Speech, Privacy and Truth*

## UNDERSTANDING HATE SPEECH

### What is hate speech?

Everyone's perception of what constitutes online hate, what is permitted, and what is too extreme, is different. It does not help that there is no clear-cut definition of hate speech in international law. If we consider not just hate speech alone, but also hateful content in general, the issue becomes even more complicated.

What hate speech and hateful content have in common, though, are the targets: a group, or members of a group, who share a particular characteristic, such as race, gender, or opinion. In the case of migrants and refugees, this may be their national origin, religion, or the fact that they are not citizens of a specific country.

Hate seems to require deeming some individuals as *other*—as essentially different or inferior in some way. It is built upon negative stereotypes derived from this judgment.

“Hate speech is *“all forms of expression which spread, incite, promote, or justify racial hatred, xenophobia, anti-semitism, or other forms of hatred based on intolerance, including: intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants, and people of immigrant origin.”*

*Committee of Ministers of the Council of Europe, 1997*

### How much do we hate?

To understand how widespread online hate speech is, we have to begin by looking at how frequently it is reported.

However, many people feel that instances of online hate speech are vastly underreported.

And, indeed, there are many reasons why victims or witnesses of hate speech on social media may not report it. They may think reporting is pointless. They may have had previous negative experiences with reporting. They may lack confidence in the justice system. The police themselves may not be well equipped to deal with online hate speech.

The lack of reliable data makes it very difficult to quantify how much hate speech exists online, which in turn makes it hard to address it effectively. However, studies that have attempted to quantify online hate speech, have generally found it to be on the rise.



**Annegret:** The problem with social media is that people get more extreme just by talking with like-minded people, so countering that is very tedious, slow work.



## WHAT IS YOUR EXPERIENCE WITH REPORTING HATEFUL COMMENTS?



**Dóra:** I have never reported any comments to the social media platforms. Several times, however, a comment disappeared from the platform, so it is possible that it was reported by another reader and was subsequently deleted.

I have also never reported any comment to the police. In one case, I received a very real and personal rape threat from someone who claimed to know me. That was of course very distressing. I replied to this person saying that what they had published could be reported to the police and punished by law. The person deleted their comment and never posted anything after that, so I can only speculate, and hope that they learned a lesson.



**Anna and Karol:** We have mixed experience with reporting hateful content to the social media companies. At times, we reported content which we clearly deemed as violating the platform community standards, but it was not always taken down.

It also happened that the reaction from the platform was very slow, which made us think it could have been quicker to deal with the comment ourselves. However, reporting can be a teachable moment. If the person who wrote a hateful comment receives a notification from Facebook saying that it violates their community standards, we can hope that they learn something.



**Annegret:** I have reported hateful comments to the platforms, but the results have been mixed at best. I do not remember seeing any effect on Facebook, while on Twitter it has been a bit better. It is quite unpredictable, but sometimes they say that a specific content was violating their rules.

When I am in Germany, I have not yet dared to report anyone because you get a message that looks a bit frightening when you want to report something – it is different from how it works in Switzerland. So I was too afraid that an important working tool will stop working properly to engage in reporting in Germany (For background on the strict law in Germany on hate speech online and the obligations of social media companies, see the case study on the German NetzDG law on page 13).

### TIP

### WHEN SHOULD I REPORT?

Social media platforms have “community standards” (though they may have slightly different names) and if you feel an account or a comment violates these standards you can report it to the platform.

To report a post or person on Facebook, click on the three dots to the right of their name or on the post itself, and choose “report post”.

If you are not sure that a comment is violating the standards, or feel unsure about reporting, you can consider these other options:

- Respond to the post
- Hide the post (this will delete it from your own timeline but will still be available in your friends’ newsfeeds)
- Unfollow the person
- Block the person

## Why do we hate online?

Understanding why online hate is so common and widespread helps us to counter it. Many studies have explored the differences between interactions in the real world (offline) and on the internet (online). While this is a complex topic, some of the facts below may help to explain why hate, bigotry, and vulgarity are so pervasive online.

OFFLINE	ONLINE
<p>If we say something hateful or vulgar about someone else, we often do it in the face of the person we are targeting. Most likely there are others around. This can increase our sense of responsibility, and consequently, restraint.</p>	<p>On the internet, we can be anonymous. There seem to be no consequences for any message that we send out in the online world. This lowers our constraints, and makes it easier to say things we would not say to someone's face.</p>
<p>When we talk with someone face-to-face, we see their immediate reactions, spoken and unspoken. Nonverbal communication—body language, facial expressions, tone of voice—is a very large part of communication activity.</p>	<p>The whole nonverbal element of communication is entirely lost to us online. We cannot see the other person, which has huge potential for loss of understanding. Because we cannot see the reactions of the other person, it is easier to use language that we would not use in a face-to-face interaction.</p>
<p>In non-digital mass communication, considerable effort and access to specific technology—such as printing presses, cameras, radio, and television broadcasts—is required to write, produce video, and disseminate information to large numbers of people. In the past, this limited hateful messaging by individuals, although it was still used as a tool for propaganda by those with power.</p>	<p>Posting or sharing a hateful comment or content on social media is a quick, impulsive, and generally effortless decision, which people may find extremely satisfying. On the other hand, engaging in counter-speech requires a conscious decision and involves considerable effort. This may explain why there is so much hateful content online and relatively little counter-speech.</p>
<p>The role played by media and politicians is key. The way media covers specific events, and which events it covers, has an impact on the way the audience perceives them. Likewise, the way in which politicians talk—or do not talk—about specific groups or events has a strong impact on how the general public perceives them. Studies show that when certain behaviours are sanctioned by authorities, people will act on their prejudices in the worst ways.</p>	<p>The more we see and hear hateful content, the more we become desensitized to it. Rather than shocking us, hate becomes normalised, a feature of everyday life.</p>



## FROM HATE SPEECH TO HOPE SPEECH

The Protestant Academy of Berlin developed workshops on responding to hate speech from a Christian perspective. Timo, the project coordinator, explains how they work.



**Timo:** One of the first actions we took in relation to hate speech was to carry out an analysis of hate speech and toxic narratives on Facebook within a Christian framework. We received excellent feedback, but those who read the report found it insufficiently concrete for people to actually put in practice. So we decided to create workshops as a way for people to reflect on the issue of hate speech in their own personal circumstances. The workshops take place offline, in order to facilitate interaction among participants.

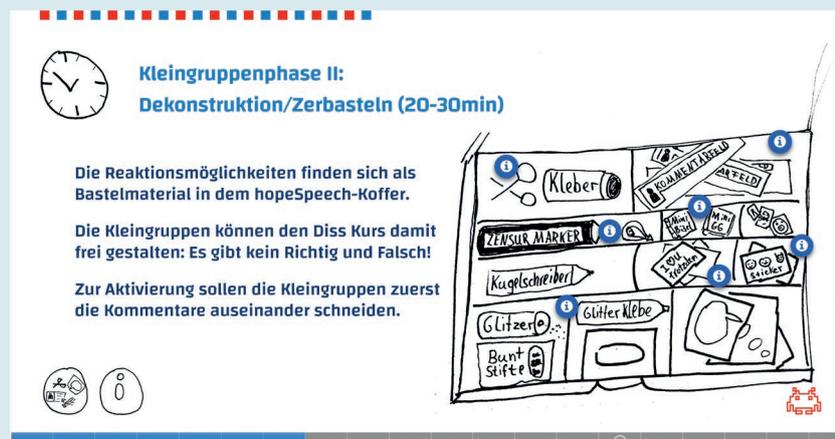
For the workshops, we use a fake social media platform (which we called *Diss Kurs*) and we display real comments and posts from Facebook. We facilitate the ensuing discussion, as participants reflect on the comments and share their own personal experiences with similar content, online or offline. They also share their strategies for dealing with this content. Together we reflect on the contributions brought by everyone and add some suggestions.

We then invite a broader reflection on social media and the way we communicate. We want to move from an attitude of "There is nothing that can be done about it" to finding new ways, new possibilities for replying to hateful content. We also discuss the audience people are trying to communicate with: the authors of hateful content themselves, or the broader audience, the silent bystanders?

The workshops are targeted at multipliers, people who can replicate the experience in their own communities. The programmes are easily scalable, and can accommodate from 8 to more than 30 people. We've had workshops with pastors who can replicate them in their parishes, and also with teenagers, who can do them again in their own communities. This way, so far, we have been able to reach about 1,000 people. An example of a half-day workshop structure can be found on page 36.

For more information on the project, read here:

<http://www.wacceurope.org/projects/social-media-divide/hope-not-hate/>



The "Netzteufel" (NetDevil) initiative focuses on helping people in civil society and church networks to engage with the web and develop skills, know-how and stories to become positive "web devils". The workshops are based on the understanding that what takes place in the virtual world is real and needs to be taken seriously by both civil society and the churches.

## Hate speech and hate crimes

Hate speech is an act of intolerance, which, if not addressed, can provoke hate crimes—acts of conflict and violence.

The latest migration monitoring report released by the Fundamental Rights Agency of the European Union provides an overview of hate crimes against refugees and migrants in the member states. These crimes are often spurred on and accompanied by hate speech.

Hate speech on the internet can and does have effects in real life. A 2018 study analysed the correlation between Facebook usage and violent crimes against refugees in Germany. Examining anti-refugee sentiments expressed on the Facebook page of the extreme right *Alternative für Deutschland* (AfD) party, the study found that crimes against refugees increased disproportionately in areas in Germany with a high use of Facebook and during times of strong online anti-refugee sentiments. Thus, according to the authors, social media was a medium of propagation of anti-refugee sentiments, which could and did lead to actual violent crimes against refugees.

Outside of Europe, in the case of violence against the Rohingya population in Myanmar, similar correlations between increased spreading of hateful messages on Facebook and real-life violence have been documented. The same was true for the 2019 attack against Muslims in Christchurch, New Zealand, and the attack against Jews in Pittsburgh, United States, in 2018.

These and other similar cases make visible the connection between hate speech on social media and hate-motivated crimes. Notably, while not all instances of hate speech lead to hate crimes, all hate crimes involve previous instances of hate speech.

In the worst scenarios, when hate speech and hate crimes are left unchecked, situations can degenerate still further and lead to crimes against humanity and genocide. That was the case in Nazi Germany, in the former Yugoslavia, and in Rwanda: in these places, the target groups were first vilified, denigrated, and then dehumanised in the press, by politicians, and by the broader society.

## UNITED NATIONS AND HATE SPEECH

Even though there are difficulties in defining and quantifying hate speech, the consensus about its danger for society is wide. In a Strategy and Plan of Action on Hate Speech released in 2019, United Nations Secretary General António Guterres called it “a menace to democratic values, social stability, and peace.” He further stated:

“Addressing hate speech does not mean limiting or prohibiting freedom of speech. It means keeping hate speech from escalating into something more dangerous, particularly incitement to discrimination, hostility, and violence.”

The United Nations has repeatedly called for attention to this issue. In September 2019, 26 experts signed an open letter expressing grave concern at the rise of hate in the world against migrants and other minorities, highlighting the connection between hate speech and hate crimes and urging public officials, politicians, and the media to promote tolerant and inclusive societies.

### TIP

When writing something on social media, consider whether you would be happy to say it to the person directly or see it printed publicly with your name attached to it.

The Pyramid of Hate is a visual depiction of how hate speech can degenerate into hate crimes.

While all hate speech is potentially dangerous, not all hate speech is equally dangerous. It is important to recognise its diverse forms and degrees, as well as to consider its broader societal context. For example, a hateful comment posted by a user in a small closed group will not have the same impact as a comment shared by a politician with thousands of followers.

Understanding the different degrees of hate speech is critical for identifying the strategy that will work best in a given situation. In some cases, education and awareness raising are key. In others, reporting to social media platforms or even the police may be in order.

## The Pyramid of Hate

*Pyramid of Hate, Anti-Defamation league, ADL.org. Published with permission.*



## Freedom of expression

“Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive, and impart information and ideas through any media and regardless of frontiers.

*United Nations Universal Declaration of Human Rights, Article 19*

Freedom of expression is a widely recognised right, enshrined in the United Nations Universal Declaration of Human Rights and in further treaties, among which are the International Covenant on Civil and Political Rights and the European Convention on Human Rights. Freedom of expression is a pillar of democracy, vital to prevent censorship and promote accountability, and indispensable for an effective and free media.

However, freedom of expression is often invoked by those spreading hateful content as the right which makes it permissible to engage in hateful discourse. It may seem that freedom of expression implies that hate speech must be tolerated.

However, one person's freedom of expression should not stifle another's. Freedom of speech is a right, but speech that incites hatred or violence against a person or a community is subject to legal sanction.

That said, this becomes more complicated as states attempt to legislate the matter. The fact that most social media companies are headquartered in the United States presents an additional dilemma; the US has a much more lenient approach to free speech than Europe, and it is very difficult to enforce standards prohibiting hate speech across different jurisdictions.

## GERMANY AND THE NetzDG LAW

Germany was one of the first countries to take a strong legal stance against hate speech. In 2018, it adopted the Network Enforcement Act, known as NetzDG. Broadly speaking, the law obliges social media companies to remove illegal content in less than 24 hours, or face potential fines of up to 50 million euros. The deadline can be extended for cases requiring further investigation.

The law is not, however, without criticism. Organisations such as Human Rights Watch and Reporters Without Borders have raised concerns. Human Rights Watch argues that the law places an undue burden on social media companies, which are private companies and not fit to be the judges of whether content is illegal. According to the organisation, the law effectively encourages companies to suppress speech, even if it is not necessarily illegal, to avoid large fines.

Human Rights Watch also criticises the lack of appeal provisions in the law, which means that users whose published comment or content has been blocked cannot ask the social media companies to review the decision. The lack of accountability and oversight is a worrying aspect of the NetzDG law.

Perhaps still more concerning is the example that the law sets for other countries with less stable democracies. Russia has already passed a law explicitly referencing the German law. Passage of such a law in a non-democratic society clearly has different consequences. Other countries, including some which notoriously infringe on freedom of expression, such as Venezuela and the Republic of the Philippines, have also hailed the German law as a positive example.

## News you can trust

“*Political polarisation has encouraged the growth of partisan agendas online, which together with clickbait and various forms of misinformation is helping to further undermine trust in media—raising new questions about how to deliver balanced and fair reporting in the digital age.*

*Nic Newman, Digital News Report 2019,  
Reuters Institute*

Traditional media (newspapers, TV, and radio) were for a long time the only sources for news. With the advent of the internet, the landscape completely changed: information from many sources, in multiple formats, became available 24 hours a day to those with an internet connection.

Social media brought further changes. Not only did even more sources of information become readily available, social media offered the opportunity to create information easily. Barriers that previously existed to access and to produce information have almost disappeared. Almost anyone can set up social media accounts, websites, or blogs, and immediately start communicating with an audience.

These changes in how we find and communicate information have contributed greatly to the democratisation of speech and freedom of expression. Producing and disseminating content to the public is no longer only the domain of those with power, access, or expertise. This means, however, that there is also little or no control over the veracity of the information shared on the internet. Hate speech and other forms of dangerous or hateful content can propagate easily across platforms, users, and voices. Platforms such as Facebook, Instagram, and YouTube also use algorithms to reward posts that receive a lot of engagement by placing them at the top of feeds. This works to the advantage of highly entertaining or extreme content.

Conversely, the broadening of the communications landscape has also triggered a diminishing trust in the media. According to a 2019 global study by the Reuters Institute, only around 42% of the people surveyed said they trust the news, including those sources they themselves use. Trust in social media is even lower, at around 23%.

A Pew Research Center study analysed how people get their news in a number of European countries. While the most established news outlets were mentioned, many people also named Google and Facebook as their sources for news. In the countries surveyed, a consistent number of people reported regularly getting their news from social media. Where that was the case, Facebook was mentioned as the most used social media source of information, with Twitter a distant follower. Younger users (18-29 years) are even more likely to use Facebook as a source of information, according to the findings.

When looking for news on social media, a high number of those surveyed also mentioned that they do not pay attention to the source of the news items shared on the platform. This is a worrying sign in light of the disinformation campaigns and false news spreading all over the internet and particularly on social media.

### TIP

Support the news you trust.  
Financially, if you can.

## Fake news and disinformation campaigns

Charges of fake news and exposure of strategic disinformation campaigns make it all the more imperative that we critically evaluate the news that we read and see, and share only what we trust, or have verified through a fact-checking service such as [Full Fact](#).

Incorrect information should be corrected. However, it may be even more important, especially in the case of deliberate disinformation, to consider what its purpose is. In many cases, those posting do not intend people to believe the information. Their intention is, rather, to sow mistrust in the news in general or in the authorities, and to direct public behaviour in a particular way. Hence, our responses should promote behaviour and processes that respect facts, people's rights, and democratic societies.

### TIP THINK BEFORE YOU SHARE

- Is this item from a trusted news source?
- Will it promote positive action?
- Make sure YOU can be a trusted source.

