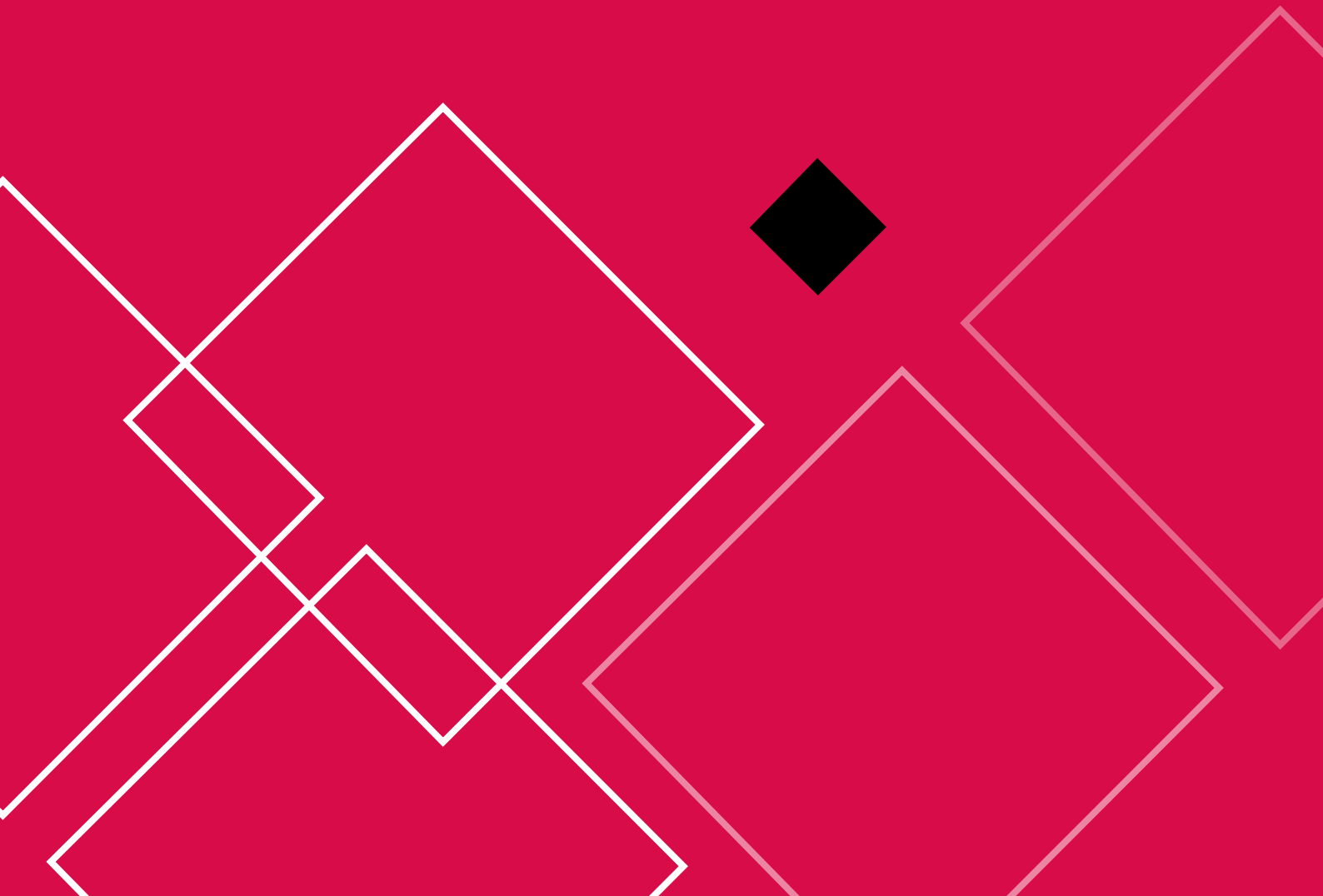




Educational TOOLKIT

to understand, prevent and react
to hate speech and disinformation



Hatedemics. Educational Toolkit

Written and edited by: Jan Dąbkowski (The Center for Citizenship Education, Poland),
Dariusz Grzemny, Vivian Rangel (Maldita, Spain)

Design and layout: Kinga Kurach

Copyright © 2025 HATEDEMICS Consortium Partners. All rights reserved.

This Education Toolkit was created as part of the HATEDEMICS project — ‘Hampering hate speech and disinformation through AI-based technologies to prevent and combat polarisation and the spread of racist, xenophobic, and intolerant speech and conspiracy theories.’

This project aims to strengthen both preventive and reactive measures against online hate speech and disinformation. It empowers NGOs, CSOs, fact-checkers, public authorities, and young activists to effectively prevent and combat polarisation, the spread of racist, xenophobic, and intolerant speech, as well as conspiracy theories. You can read more on: www.Hatedemics.eu

HATEDEMICS is a Citizens, Equality, Rights and Values Programme (CERV) project supported by the European Union under Grant Agreement (GA) no. 101143249. You are permitted to copy and distribute verbatim copies of this document, containing this copyright notice, but modifying this document is not allowed.

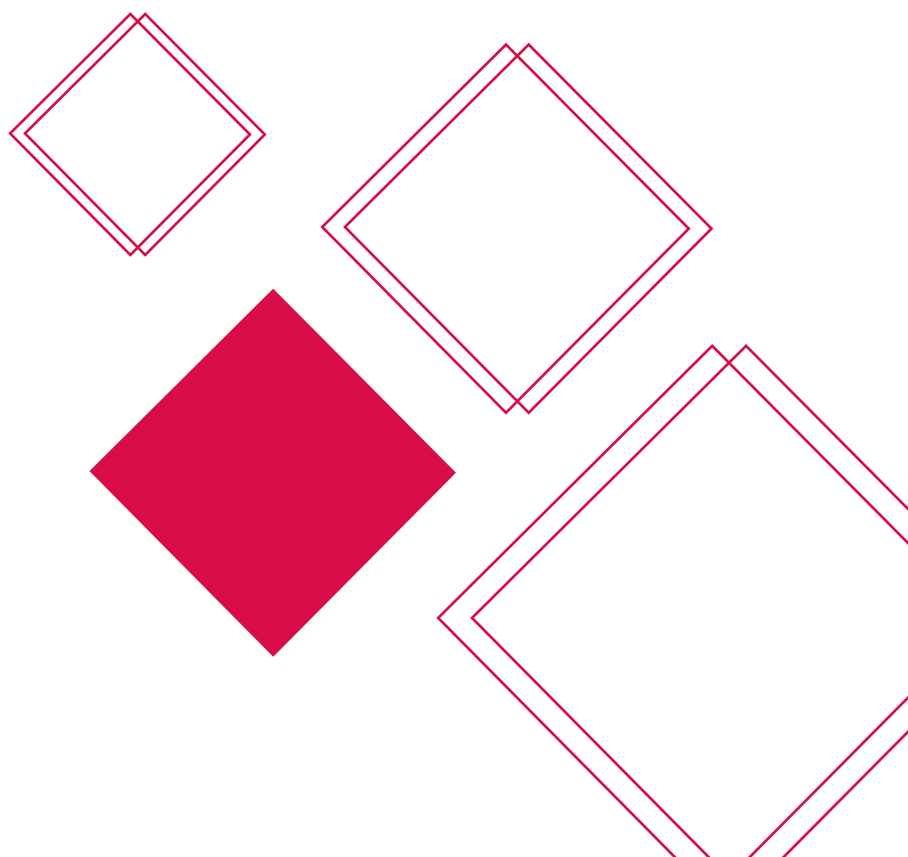


Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EACEA. Neither the European Union nor the granting authority can be held responsible for them.



Table of contents

Introduction	3
Disinformation and hate speech	3
Relation between disinformation and hate speech	4
Causes of hate speech and disinformation	6
Logical fallacies associated with hate speech and disinformation	7
Cognitive biases	8
Importance of reaction to hate speech and disinformation	9
Reaction to hate speech and disinformation	10
Narrative, counter narrative and alternative narrative	11
Style and content guidelines of responding to hate speech and false content	12
Diversity of possible actions	14
Critical thinking and media literacy	14
Fact-checking	15
5Ws and H – questions to analyse information	18
Lateral reading	18
Education activities	20
Activity 1. Why we care	20
Activity 2. Lateral reading check	23
Activity 3. Mechanisms, biases and fallacies	26
Activity 4: Spot the hate speech!	28
Activity 5. How to stay cool	30
Helplines	31



Introduction

This educational toolkit is intended for stakeholders such as teachers, educators working in non-governmental organisations (NGOs) and other civil society organisations (CSOs), particularly those involved in media education, fact-checking, anti-discrimination and human rights education. It is also intended for local and national public authorities. Consisting of an introduction to hate speech and disinformation, as well as how these phenomena interact, the toolkit also contains an educational and methodological proposal, and activities aimed at working with young people. The toolkit is intended to help young people better understand and react to hate speech and disinformation online.

Disinformation and hate speech

Although disinformation and hate speech are closely related, they are quite distinct phenomena. There are no universally agreed-upon definitions of both terms, however, based on existing legal, policy and scholarly documents, it is easy to determine their characteristic features.

Hate speech is defined as the expression of offensive, discriminatory opinions, content or communications towards specific groups of people or individuals on the basis of their identity characteristics, such as gender, age, sexual orientation, nationality, ethnicity, religion, skin colour, disability or other.¹

Disinformation is false content shared with the intention of causing harm. The creation of this type of content is most often motivated by the desire to earn money, gain publicity or increase political influence.²

The notion of disinformation is closely related to the concept of misinformation and malinformation. **Misinformation** is false or erroneous information shared out of a lack of knowledge and without the intent to cause harm. Finally, **malinformation** occurs when true information is shared in a different context, on different dates, or deliberately with the aim of causing harm, such as the publication of private data.³

¹ Inspired by: What is hate speech? <https://www.un.org/en/hate-speech/understanding-hate-speech/what-is-hate-speech> (accessed 14.08.2025)

² Understanding Information disorder, <https://firstdraftnews.org/long-form-article/understanding-information-disorder> (accessed 14.08.2025)

³ ibidem op.cit.

Relation between disinformation and hate speech

The table⁴ below can help you shed the light on the similarities and differences between hate speech and disinformation. It can also be helpful in identifying both phenomena.

	HATE SPEECH	DISINFORMATION
TARGET	Individuals or groups that share certain identity characteristics, such as gender, nationality, ethnicity, religion, sexual orientation or other	The same as in hate speech, plus: <ul style="list-style-type: none"> - Individuals or groups with the common trait, such as occupation (politicians) - State or non-state actors (government, institutions, non-governmental organisations) - A value or an ideal (e.g. democracy)
CONTENT	<ul style="list-style-type: none"> - Incitement to hate and/or violence - Calls to discriminate on the basis of identity - Content is dehumanising - Use of identity-based slurs - Language does not need to be aggressive 	<ul style="list-style-type: none"> - Deceptive content (e.g. sharing false claims) - Content created to mislead (e.g. claims on the basis of non-existent research, cherry-picked statistics, edited texts) - Content designed to undermine trust in institutions and official processes (e.g. conspiracy theories)
	<ul style="list-style-type: none"> - Content created by humans or with the use of information technology (e.g. AI) - Denial and distortion of some historical events or facts (e.g. Holocaust) - Content designed to emphasise in-group/out-group differences - Created with the intention to harm - It can take many forms - words, pictures, videos, audio material, memes on the internet 	

⁴ On the basis of: Wardle C., A Conceptual Analysis of the Overlaps and Differences between Hate Speech, Misinformation and Disinformation, United Nations 2024: https://peacekeeping.un.org/sites/default/files/report_-_a_conceptual_analysis_of_the_overlaps_and_differences_between_hate_speech_misinformation_and_disinformation_june_2024_qrupdate.pdf (accessed 14.08.2025)

HATE SPEECH

DISINFORMATION

<p>AIM</p>	<ul style="list-style-type: none"> - To dehumanise individuals or groups - To scapegoat (e.g. blame certain groups for the societal problems) - To ridicule and humiliate certain people or groups on the basis of their identity 	<ul style="list-style-type: none"> - The same as in hate speech, plus: - To fuel tension and unrest in society - To lower trust in institutions and/in organisations - To manipulate democratic principles and processes (e.g. elections) - To slander institutions/organisations and ruin their reputation
	<ul style="list-style-type: none"> - to spread certain ideologies and gain support - influencing the way other people think and promoting values that are important to the sender for some reason. This reason is not usually honourable. 	
<p>IMPACT</p>	<ul style="list-style-type: none"> - Hate crimes (targeting individuals or/and their property on the basis of people's identity) - Genocide, crimes against humanity - Reinforcing or intensifying existing stereotypes and prejudices in society 	<ul style="list-style-type: none"> - Impact on elections (lower participation, manipulated results, results not legitimate) - Democracy is undermined and institutions are targeted with violence (e.g. destroying the buildings or flags) - Decline of trust towards institutions/organisations - Social unrest (e.g. strikes, violent demonstrations)
	<ul style="list-style-type: none"> - Silencing people or groups that are targeted - Withdrawal from public life - Psychological and/or physical harm (e.g. depression, suicidal attempts, use of physical violence) - Reputation damage (both individual and institutional) - Increased polarisation and fear - Increased hatred and hostility towards individuals and/or groups 	
<p>WHERE</p>	<p>Offline: brochures, leaflets, posters, murals, graffiti, in conversations, demonstrations</p> <p>Online: posts in social media, online adverts, online groups (closed or public), message apps, use of fake accounts,</p> <p>Media: TV and radio broadcast, newspapers</p> <p>Other: conferences and seminars, songs, public debates</p>	

Causes of hate speech and disinformation

It is difficult to determine the causes of hate speech because each example must be considered on its own merits. Internal factors that might drive people to use hate speech include personality traits, belief in violent and discriminatory ideologies, and a lack of knowledge or information. In some cases, people may be motivated by others to strengthen group identity. There are, however, social mechanisms that are associated with using hate speech.

» **Stereotypes**⁵: in social sciences, stereotypes are specific cognitive representations of different groups of people. Stereotypes are usually very persistent and do not change easily. These cognitive representations are extremely generalised and simplify reality. All of this is associated with strong emotions that are evaluative in nature (this is good, this is bad). We learn most stereotypes from our social environment – family, school, work, friends, etc. They are also reproduced in culture (films, school books, advertisements). Language plays an important role in the transmission and perpetuation of stereotypes (using negative terms to describe the ‘other’ and positive terms to describe the group with which we identify). Examples: ‘Italian people cook well and are always very funny’ or ‘Africans are poor and hungry.’

» **Prejudices**⁶: stereotypes can evolve over time and become prejudices – attitudes towards certain groups of people based on unproven, generalised and incomplete information. On an emotional level, they are associated with feelings such as fear, anxiety, disgust or a sense of danger. On a behavioural level, they usually lead to the avoidance of certain people or groups of people. Example: feeling disgust towards gay people.

» **Discrimination**⁷: prejudices may lead to discrimination – treating people unfairly because of who they are, e.g. excluding them from public events, firing from work, etc. Example: not hiring someone because of their ethnic background.

» **Scapegoating**⁸: unjustified blaming somebody for something. Scapegoating can happen on a microscale (e.g. at school, where a particular person is spotted by a group of other students and blamed for things he or she did not do) or macro scale level (e.g. in a time of crisis, people look for understanding, explanation and healing, and find a ‘culprit’ who, if eliminated from the group or murdered, would heal the situation). Example: A teacher blames one student for the whole class being noisy, even though many students were talking.

⁵ CONSTELLATIONS, A manual for working with young people on the topic of racism and invisible racism: <https://www.salto-youth.net/tools/toolbox/tool/constellations.2998/> (accessed 14.08.2025)

⁶ Ibid.

⁷ Ibid.

⁸ Rothschild, Z. K., Landau, M. J., Sullivan, D., & Keefer, L. A. (2012). A dual-motive model of scapegoating: Displacing blame to reduce guilt or increase control. *Journal of Personality and Social Psychology*, 102(6), 1148–1163: <https://psycnet.apa.org/record/2012-10977-001> (accessed 14.08.2025)

» **Ethnocentrism**⁹: the belief that one's own ethnic group or nation is superior to others. It often manifests itself as a dislike of other people who are culturally or ethnically different. This can lead to discrimination or violence. Example: A teacher only celebrates holidays from their own culture and ignores or dismisses others, like only recognizing Christmas but not Diwali, Eid, or Lunar New Year, even though students from those cultures are in the classroom.

Logical fallacies associated with hate speech and disinformation

The content of hate speech and disinformation may sometimes look very convincing. This is because people who perpetuate hate speech often rely on logical fallacies – constructs hate groups use to justify their causes and their belief that people who come across this content would consider it as reasonable and truthful. Logical fallacies can use false premises, true premises but false conclusions, or true premises and conclusions but a false argument. Some fallacies are produced with the intention of deceiving or making misinformation viral.

There are many fallacies¹⁰ that show up in hate speech. Here are some of them:

- » **Argumentum ad populum** – a non-substantive way of arguing in which the speaker appeals to the tastes, preferences and prejudices of the crowd. Based on national egoism, instincts, stereotypes or prejudices, the speaker tries to win over the crowd to their thesis (e.g. The majority of people in our country thinks like this, so this is the way we should act.).
- » **Ipse dixit** – the way of arguing without providing a valid proof (e.g. The political party said it so it must be true.).
- » **A slippery slope fallacy** – occurs when someone claims that a position or decision will lead to a series of unintended negative consequences (e.g. We cannot let the migrants stay in our country as they will bring their families here and have children, which will lead us to losing our jobs and the collapse of the social security system, and as a result we will be poor and lose control over our country.).
- » **Either-or fallacy** – where someone claims there are only two ways of approaching the issue, while there are many other options (e.g. You can either vote for this party or live your life in constant danger.).

⁹ <https://www.merriam-webster.com/dictionary/ethnocentrism> (accessed 14.08.2025)

¹⁰ If you are interested to read more about them, see here, here, here, here, here and here (accessed 14.08.2025)

- » **Post hoc, ergo propter hoc** – logical error in wrongly concluding that there is a causal connection between two events merely because one occurred before the other (e.g. Ever since we got a new maths teacher, I've been getting bad grades. She must be the reason I'm doing poorly.).
- » **Argumentum Ad Hominem** – when someone attacks the person making an argument rather than the argument itself (e.g. Don't listen to Peter's opinion on the group project — he's lazy and never does anything right.).
- » **Cherry picking** – when someone selectively presents evidence that supports their argument while ignoring or dismissing evidence that contradicts it (e.g. Hate speech isn't harmful. Look at this one study that says it's just offensive, not dangerous – while ignoring dozens of other studies showing the psychological and social harm of hate speech.).
- » **Appeal to tradition** – when an argument assumes something is good or true simply because it is being traditionally done or believed, without considering whether it is still valid in the present context (e.g. In our school we have always had more men than women in the school council, and it always worked well.).
- » **Argument from anecdote** – someone uses a single, often personal, experience or story to support a general conclusion or claim (e.g. I don't think hate speech leads to real harm. I've heard people make offensive jokes all the time, and no one I know has ever been seriously affected.).
- » **Ad ignorantiam fallacy** – someone argues that a claim is true or false simply because it has not been proven the opposite (e.g. No one has ever proven that students cheat on online exams, so it must not happen.).

Cognitive biases¹¹

Our brain uses shortcuts to help us make quick decisions in everyday life. These shortcuts help us focus only on the most important information, so we don't get overwhelmed by everything we see and hear. This makes the brain work more efficiently. However, the downside is that these shortcuts can sometimes lead us to make mistakes or misunderstand things because of our thinking (cognitive) biases.

We are all affected by cognitive biases in our decision-making process, and those who spread disinformation understand these biases very well. Here are some examples of cognitive biases:

¹¹ If you are interested to read more about them, see here and here (accessed 14.08.2025)

- » **Membership bias** – tendency to favour own group over other groups (e.g. The committee chose to fund the project led by their colleague, even though another proposal was stronger, saying We know and trust our own people.).
- » **Availability bias** – tendency to rely on information that comes quickly to mind (e.g. After seeing a news report about plane crashes, someone insists flying is more dangerous than driving, despite statistics showing otherwise.).
- » **Causality bias** – occurs when people believe that there is a relationship between events that are actually independent of each other (e.g. Every time I wear my lucky socks, my team wins – so they must be the reason.).
- » **Blind spot bias** is the tendency to see oneself as less biased than other people (e.g. I'm fair and objective, unlike others who let their feelings get in the way of facts.).
- » **Confirmation bias** – tendency to search for, interpret, favour and recall information in a way that confirms prior beliefs (e.g. Belief vaccines are dangerous only reads articles that support this view and dismisses scientific research that disproves it.).
- » **Authority bias** – tendency to be more influenced by the opinions of authority people (e.g. It must be true – a famous doctor said so on TV – even if there's no supporting evidence.).

Importance of reaction to hate speech and disinformation

Hate speech has negative consequences on the lives of those affected, but it also has a negative impact on society as a whole. Hate speech coupled with disinformation can lead to stigmatisation, discrimination and large-scale violence.¹² Protecting oneself and others from hate speech is not only the moral duty of every individual, but sometimes even a legal obligation.¹³

Reacting to hate speech is not always easy, it may require some competences, ability to assume the consequences of our actions, lack of fear or lots of moral strength. However, hate speech when left unanswered may escalate violence and generate more hate, which can result in severe consequences. Therefore, we believe that reaction is always a necessity. This is our motivation and spirit of Hatedemics project.

¹² Bradshaw S., Disinformation and Identity-Based Violence, Stanley Center for Peace and Security, 2024 and Case Study: Hate Speech, UNU CPR: <https://unu.edu/cpr/report/case-study-hate-speech> (accessed 14.08.2025)

¹³ Criminalisation of hate speech and hate crime in selected EU countries, Briefing 04-11-2024, European Parliament: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/766226/EPRS_BRI\(2024\)766226_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/766226/EPRS_BRI(2024)766226_EN.pdf) (accessed 14.08.2025)

Reaction to hate speech and disinformation

There are many ways you can take a stand, even if you are not personally targeted with hate speech or disinformation:

- **EDUCATE** yourself and others: increase your knowledge about stereotypes, prejudice, discrimination and violence and educate others about hate speech and disinformation;
- **CRITICALLY ANALYSE** media messages and practise your critical thinking skills;
- **RESPOND** to hateful and disinformation posts on social networks and in media through evidence-based information;
- **SUPPORT** people who experience or have experienced hate speech or have become victims of hate crimes;
- **REPORT** hate speech and disinformation either directly to the website administrators or to law enforcement agencies, if a criminal offence is suspected;
- **ACTIVELY ENGAGE** in the activities of informal groups and non-governmental organisations that counteract hate speech, hate crimes and disinformation.

The work practices of the organisations behind the Hatedemics project in counteracting hate speech and disinformation show that these are the benefits of reacting to hate speech and disinformation:

- It sends a clear message that the individuals or groups targeted by hate speech and disinformation are not alone — that someone cares.
- It helps create safer spaces for everyone, especially in online environments.
- It provides a sense of self-fulfillment and personal growth.
- It encourages others to speak out and take action.
- It restores faith in human solidarity and action.
- It contributes to building a more just and equitable society.
- It helps reduce the impact of hate speech and disinformation on those affected.

Narrative, counter narrative and alternative narrative

Hate speech can be a narrative that uses words, videos and images to dehumanise and humiliate people because of who they are. It also shows fears, which often arise when ideas and fantasies that have nothing to do with reality are given free rein.

One of the ways to react to hate speech and disinformation are counter narratives, which can be described as concise responses to content, aimed at deconstructing and discrediting messages.¹⁴ They act as fact-based, non-aggressive reactions to messages, aiming to deconstruct, discredit, and demystify the harmful ideas underlying hate speech.¹⁵

Counter narratives are used to block or challenge hate in two main situations: first, by responding quickly to hate targeting people after specific incidents, and second, by confronting deeply rooted hateful narratives tied to societal prejudice or political power. These approaches empower activists, educators, and youth workers to counter hate more effectively.

Alternative narratives, in contrast to counter narratives, seek to replace or overshadow hateful messages by flooding the discourse with positive, constructive content, promoting positive values and what society stands for, rather than what it stands against. They provide new perspectives or stories that promote inclusion, diversity, respect, and human rights.

While both strategies aim to combat hate speech, **counter narrative are reactive** and challenge harmful ideas head-on, whereas **alternative narratives are proactive**, providing new, empowering stories that nurture a culture of inclusion and respect.

Example:

- Counter narrative: Responding to xenophobic content by correcting falsehoods, explaining facts about migrants, and debunking myths.
- Alternative narrative: Sharing stories of successful integration, celebrating diversity, and highlighting the benefits diverse communities bring to society.

Alternative narratives present a world in which the human rights of everyone are respected, every human life has value, there are no easy answers to problems and the world is not black and white.

¹⁴We CAN! Taking Action against Hate Speech through Counter and Alternative Narratives, Council of Europe 2017, p. 80: <https://www.coe.int/en/web/no-hate-campaign/we-can-alternatives> (accessed 14.08.2025)

¹⁵Toolkit for Human Rights Speech: <https://bjp-eu.coe.int/en/web/human-rights-speech/step-3-deciding-and-defining-what-you-want-to-transfer-human-rights-based-counter-or-alternative-narratives> (accessed 14.08.2025)

The primary goal of alternative narratives based on human rights values is to build alternative ways of thinking and living in society by challenging the discourse that is popular in society, which is exclusionary and oppressive to certain social groups. While hate speech is a tool of oppression, alternative narratives are a tool of emancipation. Alternative narratives therefore offer a different picture of the world than the one often portrayed in the media or in conversations with friends. These narratives give us the opportunity to show our position and express the values that are important to us. Alternative narratives do not polarise society – they do not divide it into ‘us’ and ‘them’ – rather, they focus on taking a critical look at reality and proposing an alternative way of thinking, taking into account facts and emotions, and that there is space for all social groups to coexist in peace and harmony.¹⁶

Example:

Amnesty International's campaign, **Look Refugees in the Eye**, features a video that humanises the refugee experience by focusing on personal stories rather than statistics. The video showcases refugees as individuals with families, dreams, and aspirations, challenging stereotypes and promoting empathy. The campaign has garnered millions of views across various platforms.¹⁷

Style and content guidelines of responding to hate speech and false content

Regardless of the situation, it is important to remain calm and to maintain a polite and respectful attitude. Here are some key insights into this process:

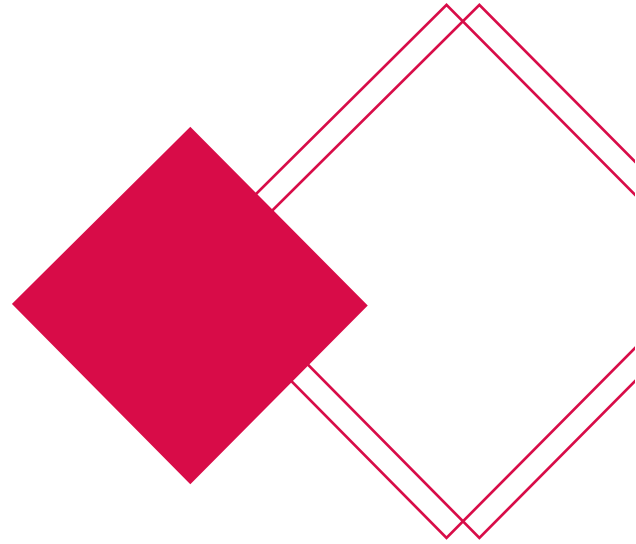
- Avoid abusive language;
- Address the message, not the author of the message;
- Refrain from using divisive labels (e.g. racist, fascist);
- Use sarcasm cautiously — it may be perceived as hostile or simply misunderstood;
- Empathise with the underlying fears or anxieties that caused the expression of hate;
- Respond with kindness and positivity.

¹⁶ We CAN! Taking action against hate speech through counter and alternative narratives, Council of Europe 2017: <https://www.coe.int/en/web/no-hate-campaign/we-can-alternatives> (accessed 14.08.2025)

¹⁷ <https://www.amnesty.org/en/latest/news/2016/05/look-refugees-in-the-eye/> (accessed 14.08.2025)

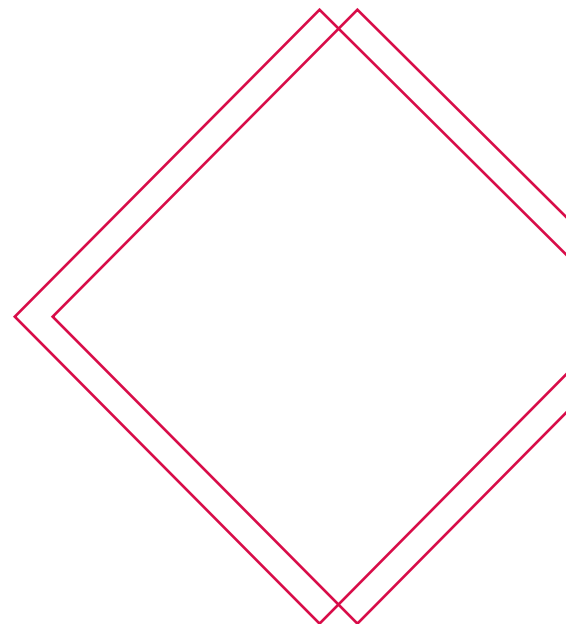
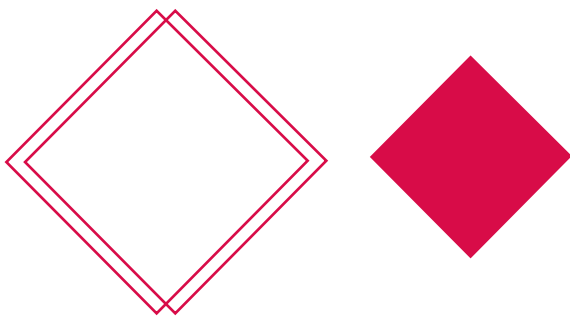
Responding to hate speech:

- Encourage the interlocutor to avoid generalisations;
- Show support for the targeted people/group;
- Challenge negative stereotypes using:
 - Facts,
 - Your personal stories and experiences;
 - Provide context on the origins of stereotypes.



Responding to false content:

- Counter it with accurate information using facts and statistics and citing the sources;
- Offer context and caution against overgeneralisation;
- Highlight the consequences of spreading misinformation;
- Suggest consulting reliable sources;
- If out-of-context, inaccurate, or outdated news and altered images: cite original sources, and compare with accurate facts;
- If satirical news: clarify the joke's origin and context, compare with accurate information.
- Describe used manipulation method (eg. when used scapegoating, define how it works);
- When debunking the untruth, communicate facts first and last, as repeating information helps remembering it.¹⁸



¹⁸ When correcting the record, make a 'truth sandwich': <https://newslit.org/tips-tools/news-lit-tip-rinse-repeat/> (accessed 14.08.2025)

Diversity of possible actions

Reacting to hate speech and false content is important and can take many different forms. Sometimes, you may want to report it to the service provider, sometimes you may be willing to engage in the discussion and bring facts that do not support the statement made, and on other occasions you may want to explore why the person says things that can be considered hate speech or false. However, there are situations when you may decide to ignore hate speech. It may be relevant when it comes to people spreading hate and false content as a form of seeking attention, as engaging can amplify their influence. In some cases, responding may escalate conflicts or spread harmful ideas further. You may wish to ignore it or even block the user when you do not feel safe. In such situations, you may want to talk to someone about it and see how they can support you in choosing the relevant way to respond. You should always decide for yourself what the best response is in a given situation, and if you have any problems or dilemmas, contact someone close to you, someone you trust, and think about how to respond together.

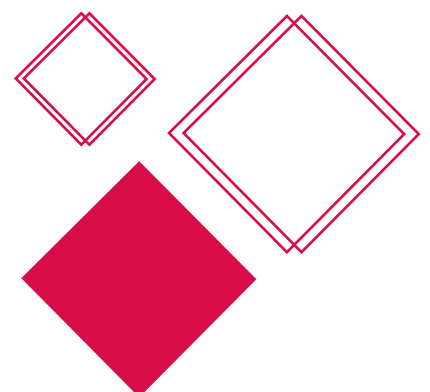
Critical thinking and media literacy

Critical thinking is a key competence to prevent disinformation and hate speech – an action of reflecting and becoming aware of all possible data, context and circumstances related to the observed situation or gathered information in order to understand it in the best way.

One way to develop critical thinking skills is to learn about media and become more media literate, as this helps people become more informed and responsible consumers and creators of media in today's information-rich world. Media Literacy can be defined as ability to:

- **Decode media messages (including the systems in which they exist);**
- **Assess the influence of those messages on thoughts, feelings, and behaviours; and**
- **Create media thoughtfully and conscientiously.**¹⁹

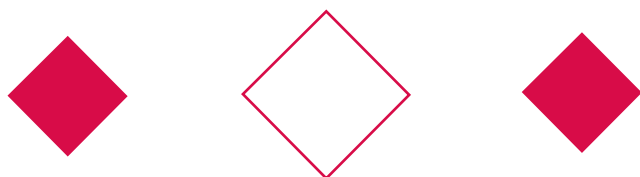
¹⁹ <https://medialiteracynow.org/challenge/what-is-media-literacy/> (accessed 14.08.2025)



Fact-checking

The role of fact-checkers is not new: **accuracy, rigour, and information based on data and facts are at the heart of quality journalism**, and in this sense, verification is part of good journalistic practice. In the United States, since the 1930s, prestigious media outlets such as the **New Yorker magazine** have had professionals dedicated to verification. Their task was to review journalists' work and confirm that they met ethical standards (accuracy, representation of diverse points of view, contrasting sources and data, etc.).

At the beginning of the 21st century, however, fact-checking began to grow rapidly in the United States. Various organisations hired journalists to verify the statements made by politicians on TV programs or in parliamentary sessions, such as **FactCheck.org**, **Politifact.com** and Washington Post Fact Checker²⁰. Major changes in the media ecosystem - such as widespread internet access (now available to more than half of the world's population)²¹, increased mobile phone use, and the proliferation of social media - have made it easier to both produce and consume a variety of content formats. We have never had so much access to information, but at the same time, we live under constant attack from hoaxes and lies.



With the rise of fake news, the need for fact-checkers became even more evident. In addition to verifying political and journalistic discourse, public discourse also needed to be scrutinised. Hoaxes appeared everywhere and in multiple formats — from manipulated data cited by politicians on X (former Twitter) to memes, audio clips, and decontextualised photos — all containing fake information designed to mislead. Fact-checking, which was already part of journalistic methodology, began to be recognised as a distinct genre and is now considered one of the most important tasks in the digital journalism ecosystem.²²

²⁰ You can read more about it in Lucas Grave book 'The Rise of Political Fact-Checking in American Journalism', <https://cup.columbia.edu/book/deciding-whats-true/9780231542227/> (accessed 14.08.2025)

²¹ Number of internet and social media users worldwide as of February 2025: https://www.statista.com/statistics/617136/digital-population-worldwide/?__sso_cookie_checker=failed (accessed 14.08.2025)

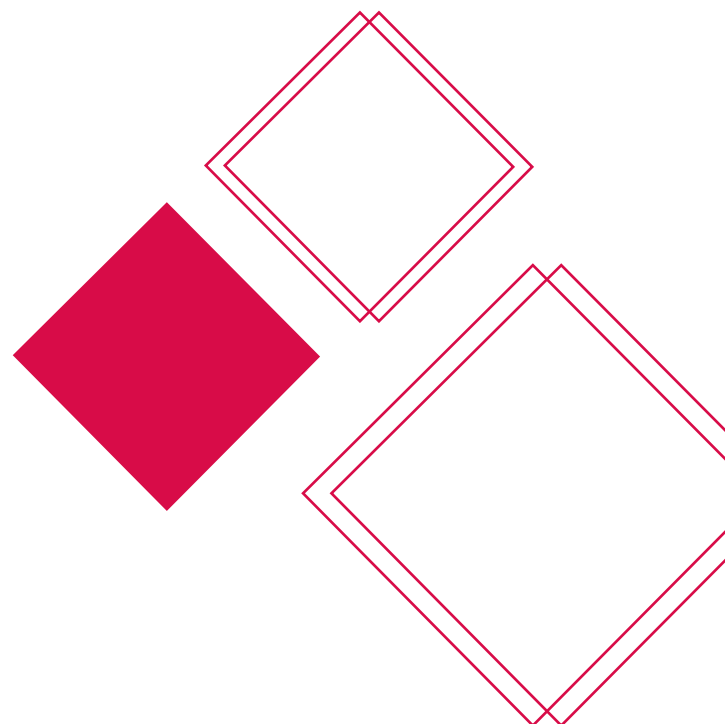
²² Graves L., Amazon M.A., Fact-Checking as Idea and Practice in Journalism, Oxford Research Encyclopedia of Communication: <https://ora.ox.ac.uk/objects/uuid:a7450b2f-f5a7-4207-90e2-254ec5de14e2/files/m3da900b71ba84b74d67f8e8c65880f1a> (accessed 14.08.2025)

The objective of fact-checking is to provide citizens with tools and information that allow them to form opinions with all the data and verified facts at hand and, consequently, make informed decisions.²³ False information can influence how society thinks and how it deals with everyday problems, and that's why fact-checkers' central mission is to shed light and data on these issues. Broadly speaking, **fact-checking involves detecting hoaxes, verifying their content, and spreading the denial.**

Working as a fact-checker means to follow a strict code of principles, transparency, and methodology. In 2015, the International Fact-Checking Network (**IFCN**) was created, to support the various fact-checking initiatives by promoting best practices and exchanges in this field; as well as to certify that fact-checkers worldwide conduct their work according to the highest standards of quality and ethics. In 2022, nearly 50 fact-checking organizations in Europe joined forces to create the European Fact-Checking Standards Network (**EFCSN**). The central objectives were to define higher verification standards appropriate to European reality and to build a professional integrity code for independent European fact-checkers and OSINT (Open source intelligence) organizations.

When some people think of a fact-checker, they only see a journalist who looks into a claim, checks it and confirms whether it is true or not. Nevertheless, fighting disinformation effectively nowadays takes much more than that:

- quickly detecting possible **hoaxes** before they go viral and affect more people;
- explaining that not everything is black or white, helping people to be well-prepared against falsehoods through Media Literacy or regulation;
- creating tools to make it harder for others to deceive us, or fighting together against manipulation are some of the fact-checkers tasks nowadays.



²³ Cooperating with fact-checkers, civil society, media and academia:
https://commission.europa.eu/topics/countering-information-manipulation/cooperating-fact-checkers-civil-society-media-and-academia_en
(accessed 14.08.2025)

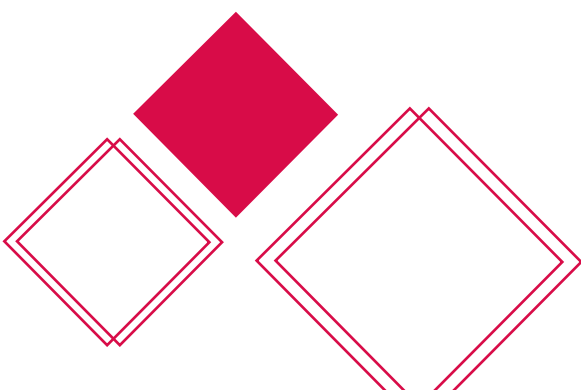
Disinformation narratives

It is also important to learn how to identify the disinformation narratives: stories that spread hoaxes and misinformation, endangering society by generating perceptions not based on data and facts. Disinformation narratives often appeal to emotions and can damage integration, equality, and democratic systems. The disinformation narratives are **what the disinformation creators want you to believe**. When an idea is repeated many times, people will believe it more. Also, one hoax can contain several disinformation narratives. Debunking these stories with data and facts stops the chain of misinformation and prevents citizens from believing similar hoaxes.

How to verify like a fact-checker

Sometimes, verifying a content is very complex and demands a great amount of time and the help of tools that allow to geolocalize a photo or prove that a video is a deepfake. In other situations we just need to use the verifier's main tool: observation. Most of the misinformation content we see everyday could be debunked by observing carefully and not sharing questionable content.

In order to learn how to debunk like a fact-checker, you can consult a **toolbox** designed to teach the basic tools of verification from **First Draftnews**. In this toolkit you can learn, for example, how to use a **Reverse Image Search**: an easy way of getting information about a picture and verifying its context. You can also practice verifying videos using **Invid**, a great tool to save time and be more efficient especially when verifying videos.



5Ws and H – questions to analyse information

The five Ws of journalism is a checklist used to ensure that the lead (the first part of the journalistic work) contains all the essential points of a story. To write a good lead, the reporter should answer:

Who is the story about?
What is the story about?
Where did the story happen?
When did the story take place?
Why did it happen?²⁴

These questions after adding an 'H' for '**How**' can be developed to a checklist which can help to analyse the context of any information and practice critical thinking skills. You can ask about details, benefits and harms, reasons, results and solutions, shown perspectives and its lack, (un)importance, influence and interdependence, potential, sources, similarities and differences, connected values and lifestyles. Check the **Ultimate Cheatsheet for Critical Thinking**²⁵ prepared by the Global Digital Citizen Foundation for inspiration.

Lateral reading

Lateral reading is a technique that allows us to quickly assess whether a web page's content comes from reliable sources and is based on quality data, references, and information. In a **research conducted by the Stanford History Education Group**²⁶ educators discovered that it is an efficient way to assess the quality of an unfamiliar website.

This powerful strategy²⁷ used by fact-checkers to evaluate information online involves not reading a website 'vertically', as if it were a printed document to evaluate its quality, but instead using a method called 'lateral reading'. This means analysing the webpage by opening several tabs to answer key questions about the content.

- Who is the author of the content? *Is it an expert source?*
- Can we find evidence? *Is it based on facts and data?*
- Is it published by other reliable sources? *What do they say about it?*

²⁴ You can read more about its origins here: https://en.wikipedia.org/wiki/Five_Ws (accessed 14.08.2025)

²⁵ Ultimate Critical Thinking Cheat Sheet: <https://blog.education.nationalgeographic.org/2017/01/20/12-things-we-learned-this-week-8/ultimate-critical-thinking-worksheet/> (accessed 14.08.2025)

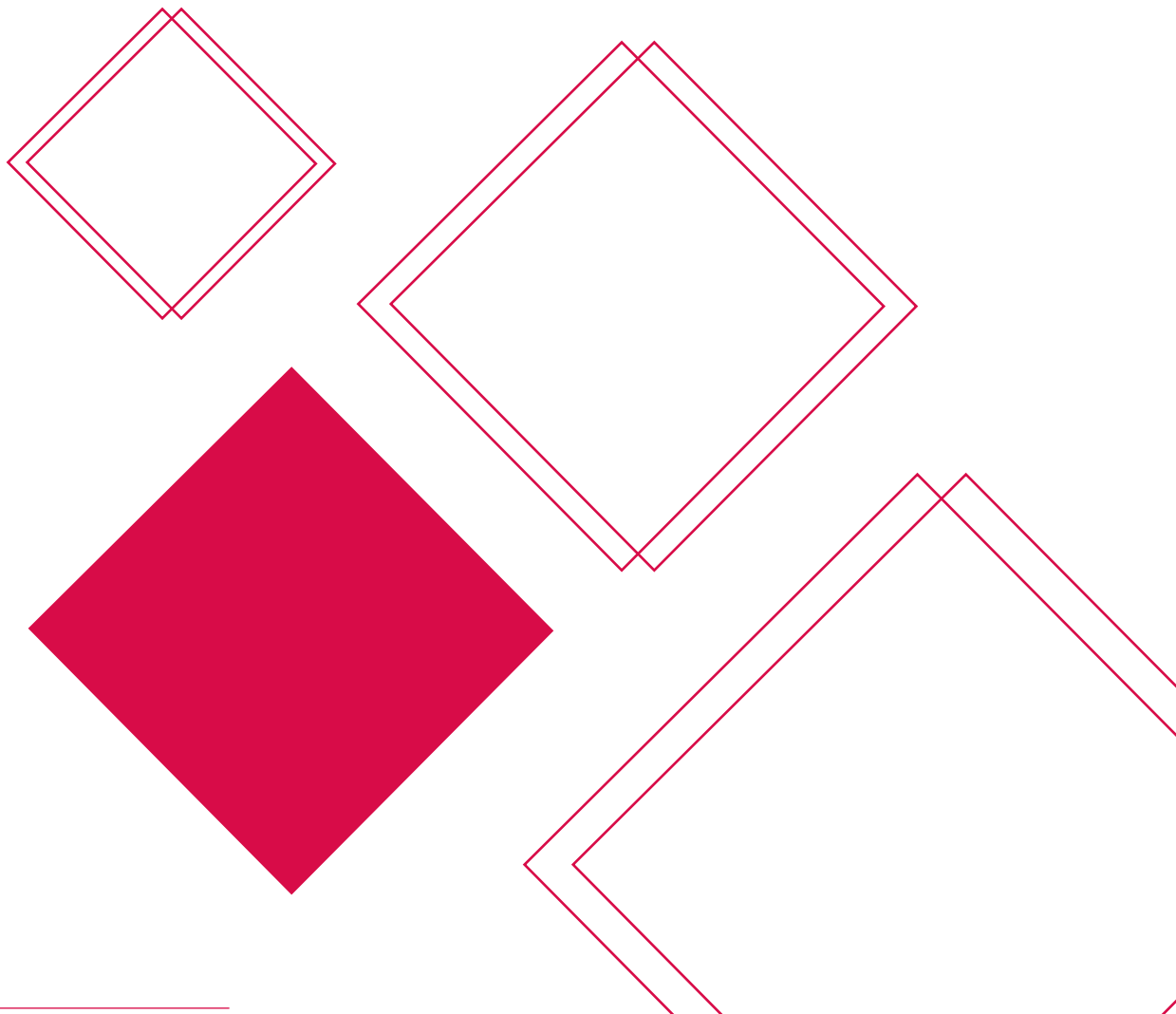
²⁶ Lateral Reading on the Open Internet: <https://cor.inquirygroup.org/research/lateral-reading-on-the-open-internet> (accessed 14.08.2025)

²⁷ Teaching Lateral Reading: <https://cor.inquirygroup.org/curriculum/collections/teaching-lateral-reading> (accessed 14.08.2025)

In addition to the above key questions, the lateral reading technique invite us to promote critical thinking and go deeper analysing others aspects of the content, such as:²⁸

- Who funds or sponsors the site where the original piece was published?
- What do other authoritative sources say about that site?
- When you search for the topic of the original piece, are the top results from fact-checking organisations?
- Have any questions been raised about other articles the author has written?
- Does the information you find elsewhere contradict the original piece?
- Are credible news outlets reporting on — or perhaps more importantly, not reporting on — what you are reading?

Lateral reading helps us better evaluate our sources of information — that is, to distinguish between truth and lies - and to consciously exercise our right to choose how we access information.



²⁸ Expand your view with lateral reading: <https://newslit.org/tips-tools/expand-your-view-with-lateral-reading/> (accessed 14.08.2025)

Education activities

DISCLAIMER

The activities in this section may trigger different emotional responses in participants, particularly when dealing with challenging topics such as hate speech. To ensure everyone's comfort and safety, it is a good idea to clearly define the rules of discussion at the beginning of the session, including mutual respect, the right to express one's emotions, and the option to withdraw from the exercise at any time. Facilitators should emphasise that the aim is to learn how to recognise and counteract hate speech, not to reproduce or accept it. By creating an atmosphere of trust and openness, encourage participants to share their feelings and provide space for discussion after the exercises. This will allow participants to safely explore difficult topics and develop empathy towards others.

ACTIVITY 1 **Why we care**

The fundamental human rights outlined in the **Universal Declaration of Human Rights**²⁹ protect and uphold the dignity, equality, and freedom of all people. Some of them are listed below in a simplified version.



Objectives:

- Developing knowledge about human rights;
- Rising awareness about connection between human rights and hate speech and disinformation;
- Learning possible arguments for reaction to hate speech and disinformation



Learners' age: 10+



Time: 60 minutes

You can show a short video about human rights before starting the activity: **[What are the universal human rights? - Benedetta Berti](#)**



Before the activity, prepare the puzzle pieces.

²⁹ Universal Declaration of Human Rights: <https://www.un.org/en/about-us/universal-declaration-of-human-rights> (accessed 14.08.2025)



Instructions:

1. Ask participants to form 12 groups. Each group picks up one jigsaw puzzle piece.
2. Each group discusses the meaning of the human right written on their piece. After 5 minutes, the first group places their puzzle piece on a wall or a board and explains what the right means to them and how it connects to hate speech. You may add to their explanation if needed.
3. The next group adds their piece to the previous one, explaining the meaning of their piece and its connection to hate speech. This process is repeated until all groups have shared their pieces and the full jigsaw puzzle has been created.



Debriefing

- Did you already know all the human rights mentioned? Which ones were new to you?
- Was it easy or difficult to explain the meaning of these rights and find a connection between them and hate speech?
- How can knowing about human rights help when it comes to reacting to hate speech and disinformation?



<p>Art. 1. All human beings are born free and equal in dignity and rights.</p>	<p>Art. 2. Everyone has the right to be treated in the same way, irrespective of race, colour, sex, language, religion, political opinion, property, birth, or other status.</p>	<p>Art. 3. Everyone has the right to life and to live in freedom and safety.</p>	<p>Art. 5. Everyone has the right to be free from torture and from inhuman and degrading treatment.</p>
<p>Art. 7. The law is the same for everyone.</p>	<p>Art. 13. Everyone has the right to live and travel freely within state borders.</p>	<p>Art. 14. Everyone has the right to go to another country and ask for protection if they are being persecuted or are in danger of being persecuted.</p>	<p>Art. 16. Everyone has the right to marry and have a family.</p>
<p>Art. 18. Everyone has the right to believe whatever they wish (including, but not confined to, religion).</p>	<p>Art. 19. Everyone has the right to say what they think and to give and receive information freely.</p>	<p>Art. 29. Everyone must respect the rights of others.</p>	<p>Art. 30. No-one has the right to take away any of these rights.</p>

ACTIVITY 2 Lateral reading check

Lateral reading is a technique that allows to quickly assess whether a web page's content comes from reliable sources and is based on quality data, references, and information.



Objectives:

- Learning about transparency and source credibility factors;
- Developing competence in identifying trustworthy sources of information.



Learners' age: 13+



Time: 45 minutes



Instructions:

Choose any website and evaluate its credibility using the lateral reading method. Learn about the steps you can use from the examples below.

Examples:

The lateral reading technique can be applied to any webpage. The following examples show how students can practise using this technique on a computer with multiple tabs open to quickly verify whether a website can be considered a quality source of information. Select the website you would like to check and follow the instructions below.

Example 1: <https://medlineplus.gov>

1. Who is the author of the content? First, we look for information about the authors.

We find that the page is created by the National Library of Medicine of the United States and that it provides information about its authors in the 'About' section:

<https://medlineplus.gov/about>

The website states that it belongs to the NIH (National Institutes of Health) and indeed redirects to the National Library of Medicine. In addition, the page has a .gov domain (used by official government websites), which provides another level of credibility regarding its content.

2. Can we find evidence? The page publishes its methodology, which includes consulting multiple government sources such as the Food and Drug Administration (FDA) and other recognized scientific sources. The information from these sources is reviewed and updated at different times. They also indicate a list of expert reviewers.

3. Is it published by other reliable sources? We open other tabs and search for 'MedlinePlus'.

The National Library of Medicine states that MedlinePlus is 'reliable, up-to-date health information to you'. Also **USA.gov** states that you can 'Get reliable health information from MedlinePlus'.

In Wikipedia (in the article with proper references) we find that 'MedlinePlus' is an online information service provided by the U.S. National Library of Medicine. It provides health information free of charge, in English and Spanish. Material must meet a strict set of selection criteria for inclusion.'

In principle, we found no information questioning or criticising the content of MedlinePlus. We can conclude that MedlinePlus is a website that offers high-quality, easy-to-understand health information. Its methodology is clear and transparent, based on multiple sources. Its authors are correctly identified and traceable. Therefore, it is a reliable source for information on health-related topics.

Example 2: <https://thepeoplesvoice.tv>

1. Who is the author of the content? First, we look for information about the authors. We cannot find this information, just a vague mention about 'our team of writers' in the 'about us section'. We can read that 'The ThePeoplesVoice.tv website (the 'Site') is operated by 'Fact Checked Limited.'

A search for 'Fact Checked Limited' does not provide information about the authors of the website. Instead, we can find debunks from fact-checking websites such as **Africa check**, **Maldita** and **TheCanadianPress**, which highlight false content published by Fact Checked Limited.

If we check some of the 'news articles' published we can find that **Baxter Dmitry** is the author of several articles. The search for Baxter Dmitry does not bring information about the alleged journalist (such as his biography, a LinkedIn profile or other contributions in different media). Instead, we can find some articles that show that Baxter Dmitry is a **regular contributor** of this 'alternative reality web' or that had previously been posing **as an unrelated Latvian man using a stolen profile photo**.

2. Can we find evidence? The page does not publish its methodology. We can find that 'The People's Voice' is a for-profit, independent news and entertainment website. "We generate revenue via advertising" but it is not clear who finances the website.

A disclaimer under the **terms of use** of 'The People's Voice' states that its content may not be accurate or reliable and may 'include inaccuracies'. Obviously, any media site can make mistakes, but it is a little bit suspicious that a media site assumes that inaccuracies are a frequent occurrence. We cannot find information about sources, 'alleged investigations' or sponsors.

3. Is it published by other reliable sources? In **Wikipedia** we find that 'The People's Voice (formerly known as 'NewsPunch' and 'Your News Wire') is an American fake news website based in Los Angeles. Wikipedia's article shows a list of references, including fact-checkers' articles and academic investigations that report 'The People's Voice' has published several false stories, including **The Pizzagate Conspiracy Theory** and **Anti-vaccination hoaxes**.

After applying the Lateral reading method, we found that 'The People's Voice' is not a reliable source. It has the appearance of a media website and publishes content with a 'Fact Checked stamped', but in fact, the website publishes misinformation that is shared as real.



Debriefing

- Was it easy or difficult to do a lateral reading check?
- Do you practice it sometimes?
- Do you plan to practice? Why?
- When is it especially important to practice it?

ACTIVITY 3 **Mechanisms, biases and fallacies**

Social mechanisms, cognitive biases and logical fallacies are factors which help hate speech and disinformation spread. By getting to know some of them, we can be more immune to them and be more critical to other forms of manipulation.



Objectives:

- Developing awareness about social mechanisms, cognitive biases and logical fallacies;
- Developing critical thinking competence.



Learners' age: 13+



Time: 45 minutes



Instructions:

1. Ask participants to form 7 groups. Provide each group with Situations' card and a List of logical fallacies/biases/social mechanisms.
2. Each group needs to read through situations on their cards and match them with the logical fallacies, biases and social mechanisms (more than one can fit). After that, ask the participants to come together and share their responses with the rest of the group.



Debriefing

- Was it easy or difficult to match the situations and logical fallacies/bias/social mechanisms?
- Did you know some of them? Which? Which were totally new for you?
- Can you provide concrete examples from your reality that represent some of those fallacies/bias/social mechanisms?
- Do you know any other examples of logical fallacies/bias/social mechanisms?
- In which way knowing logical fallacies/biases/social mechanisms can help in acting against hate speech and disinformation?

SITUATIONS' CARDS

- 1 During a heated online discussion about black people, a user posts: Everyone knows that these people are dangerous and bring crime. Just look at how many people in our town agree with me. We need to get them out to protect our community.
- 2 During a community meeting about updating local school policies to be more inclusive of LGBTQ+ students, a long-time resident stands up and says: For generations, we have taught that marriage is only between a man and a woman. This is how it has always been and how it should stay. Promoting these new ideas in our schools is an attack on our values.
- 3 During a public debate on refugee resettlement, a speaker points to a recent rise in unemployment that occurred shortly after a group of refugees arrived in the town: Ever since they arrived, our unemployment rates have shot up. It's obvious they are taking all the jobs from hardworking local people.
- 4 When Sarah joins the school computer club, she's excited to try out for the competition team. During planning meetings, her suggestions are ignored, and she's asked to help 'make the PowerPoint slides' instead. Some classmates assume coding is 'just not for girls.' When it's time to select the team leader, the advisor quickly chooses a boy, believing he'll 'speak up better on stage.'
- 5 During a heated debate about religious groups, a participant states: It may be an uncomfortable fact, but unlike those people, whose beliefs are full of prejudice and intolerance, my conclusion is that their religion is inherently backward and dangerous.
- 6 During a debate about allowing a new community center for a minority group, a speaker argues: If we allow them to build this center, next they'll demand their own laws. Soon, they'll be taking over our neighbourhoods and imposing their culture on our children. We must stop this now, or we will lose our entire way of life.
- 7 Jake, who uses a wheelchair, is excited to serve on the student council and help plan the annual school festival. At the first meeting, classmates unanimously decide he cannot join the logistics team, stating it's 'not suitable' for someone with his limitations. The advisor supports their choice and tells Jake to only assist remotely, refusing to let him participate onsite. When a complaint is raised, he's told, 'it's just easier this way for everyone.' Even though Jake speaks up and has the right skills, he is left out just because he has a disability.

List of logical fallacies/biases/social mechanisms

- A **Blind spot bias** – tendency to see oneself as less biased than other people.
- B **Stereotype/Prejudice** – stereotypes are oversimplified beliefs about a group of people who share the same identity, which shapes a way of treating this group, while prejudice is a preconceived opinion or judgment about individuals or groups which is connected with negative emotions and avoidance towards them.
- C **Appeal to tradition** – when an argument assumes something is good or true simply because it is being traditionally done or believed, without considering whether it is still valid in the present context.
- D **Post hoc, ergo propter hoc** – logical error in wrongly concluding that there is a causal connection between two events merely because one occurred before the other.
- E **Argumentum ad populum** – a non-substantive way of arguing in which the speaker appeals to the tastes, preferences and prejudices of the crowd. Based on national egoism, instincts, stereotypes or prejudices, the speaker tries to win over the crowd to their thesis.
- F **A slippery slope fallacy** – occurs when someone claims that a position or decision will lead to a series of unintended negative consequences.
- G **Skapegoating** – unfairly blaming an individual or group for problems or negative events, often to divert attention from the real causes.
- H **Discrimination** – treating people unfairly or denying them opportunities because of characteristics like their gender, sexual orientation, disability, nationality, ethnicity, religion, or other group identity.

ACTIVITY 4 **Spot the hate speech!**

Recognising hate speech is crucial to effectively combat it and protect vulnerable groups from harm. Identifying hate speech helps prevent its escalation into violence or discrimination, thus supporting societal peace and respect for human rights.



Objectives:

- Learning to distinguish between actual hate speech and expressions that may be offensive or controversial but do not meet the definition of hate speech;
- Practicing appropriate reactions to hate speech.



Learners' age: 13+



Time: 45 minutes



Instructions:

1. Hand out the 'Situations' card' and 'Task cards' to each participant. Give them some time to answer the questions. Participants are asked to identify if the given situations are hate speech or not and choose the reaction they would undertake. They can also provide their own proposal for the reaction. Before the activity you may want to present the definition of hate speech and disinformation.
2. Join participants into teams of 3-5 people and let them discuss their answers – ask them to focus on similarities and differences and ask them to explain their choices. The task is not to find common answers but rather to see the diversity of reactions.
3. Ask the groups to provide feedback for the group work and correct (if needed) if you spot participants who wrongly classified the examples (e.g. they did not identify hate speech correctly).



Debriefing

- Did you have problems in identifying the examples of hate speech? What made a statement cross the line into hate speech?
- Is it difficult to react? What skills are needed to do it?
- Why is it recommended to react to hate speech?

SITUATIONS' CARDS

1. Post on social media about migrants: Go back where you came from, you don't belong here!

2. Post on social media about Muslims: Their religion teaches them to hate everyone else!

3. I'm tired of political correctness. Everyone's too sensitive these days.

4. Let's all avoid Peter, he always slows the group down!

5. A meme ridiculing someone's accent, linking it to their ethnicity.

6. Women are too emotional to be leaders!

7. We cannot allow migrants to settle here because they only want to take advantage of the social welfare system and do not want to work. We do not need people like that.

TASK CARD

1. Is the statement hate speech or not? Why?

2. How would you react (you can choose more than one):

Engage in a respectful discussion (if safe)

Challenge the misinformation (if safe)

Report to the platform/moderator

Ignore it

Block the user

Educate the person (if appropriate and safe)

Seek more information about the user's point of view.

Other reaction:.....

.....

SITUATIONS

1	2	3	4	5	6	7

ACTIVITY 5 **How to stay cool**

When reacting to hate speech, it is important to manage your own emotions in provocative situations. This allows for a rational and constructive response while maintaining respect for the other person. Controlling emotions can help defuse tension, prevent the escalation of aggression, and make the conversation safer and more productive.



Objectives:

- Learning strategies to manage difficult emotions and stay regulated when dealing with hate speech;
- Being better prepared to handle and respond to hate speech effectively.



Learners' age: 10+



Time: 45 minutes



Instructions:

To help your students develop better control over their emotions – such as sadness, anger, or stress – choose appropriate exercises and either show short instructional videos or demonstrate the techniques yourself.

The box breathing technique can help in case of feeling anxious, overwhelmed, or having a stressful moment:



Method 5-4-3-2-1 can be used to calm down by grounding in the present moment:



The body scan exercise can help improve focus and overall well-being:



Debriefing

- Was it easy or difficult to learn the calming exercises?
- Which exercises do you like the most?
- In what kinds of situations can it be helpful?
- What other ways of calming down do you practice? You can suggest also:
 - Counting to 10 or 100;
 - Taking a break, getting fresh air;
 - Listening to music, thinking about something else;
 - Walking, stretching, exercising, jogging, biking;
 - Meeting, talking or playing games with friends or family;
 - Engaging in a hobby;
 - Asking for help.

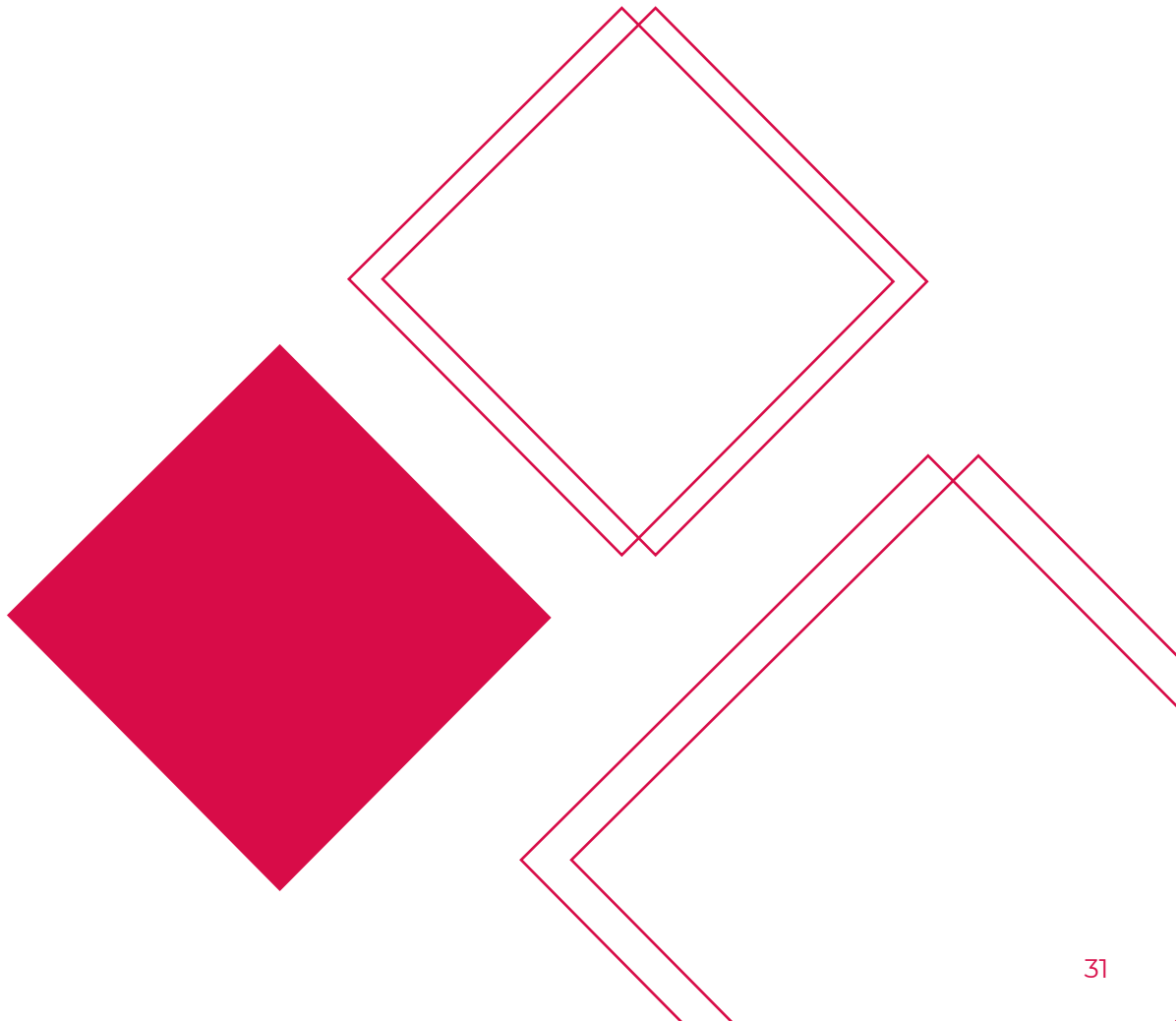
Helplines

When working with young people on issues such as hate speech and disinformation, it is important to support them in dealing with any difficult thoughts or feelings that may arise from participating in such educational activities. Ideally, they will have the support of you, their family or friends. However, we always recommend providing young people with contact information for helplines that can support them with this and other life challenges.

Child helplines are civil society organisations, government bodies or social enterprises that provide help, support and counselling to children (under 18s) and young people (up to 25s) via online and offline channels.

Child helplines often serve as the first point of contact for young people with child protection services and are frequently the most trusted and accessible way for young people to receive further support.

ChildHelpLineInternational.org/helplines



This Education Toolkit was created as part of the HATEDEMICS project — ‘Hampering hate speech and disinformation through AI-based technologies to prevent and combat polarisation and the spread of racist, xenophobic, and intolerant speech and conspiracy theories.’



Project aims to strengthen both preventive and reactive measures against online hate speech and disinformation. It empowers NGOs, CSOs, fact-checkers, public authorities, and young activists to effectively prevent and combat polarisation, the spread of racist, xenophobic, and intolerant speech, as well as conspiracy theories.

Copyright © 2025 HATEDEMICS Consortium Partners. All rights reserved.

HATEDEMICS is a Citizens, Equality, Rights and Values Programme (CERV) project supported by the European Union under Grant Agreement (GA) no. 101143249.



Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EACEA. Neither the European Union nor the granting authority can be held responsible for them.

