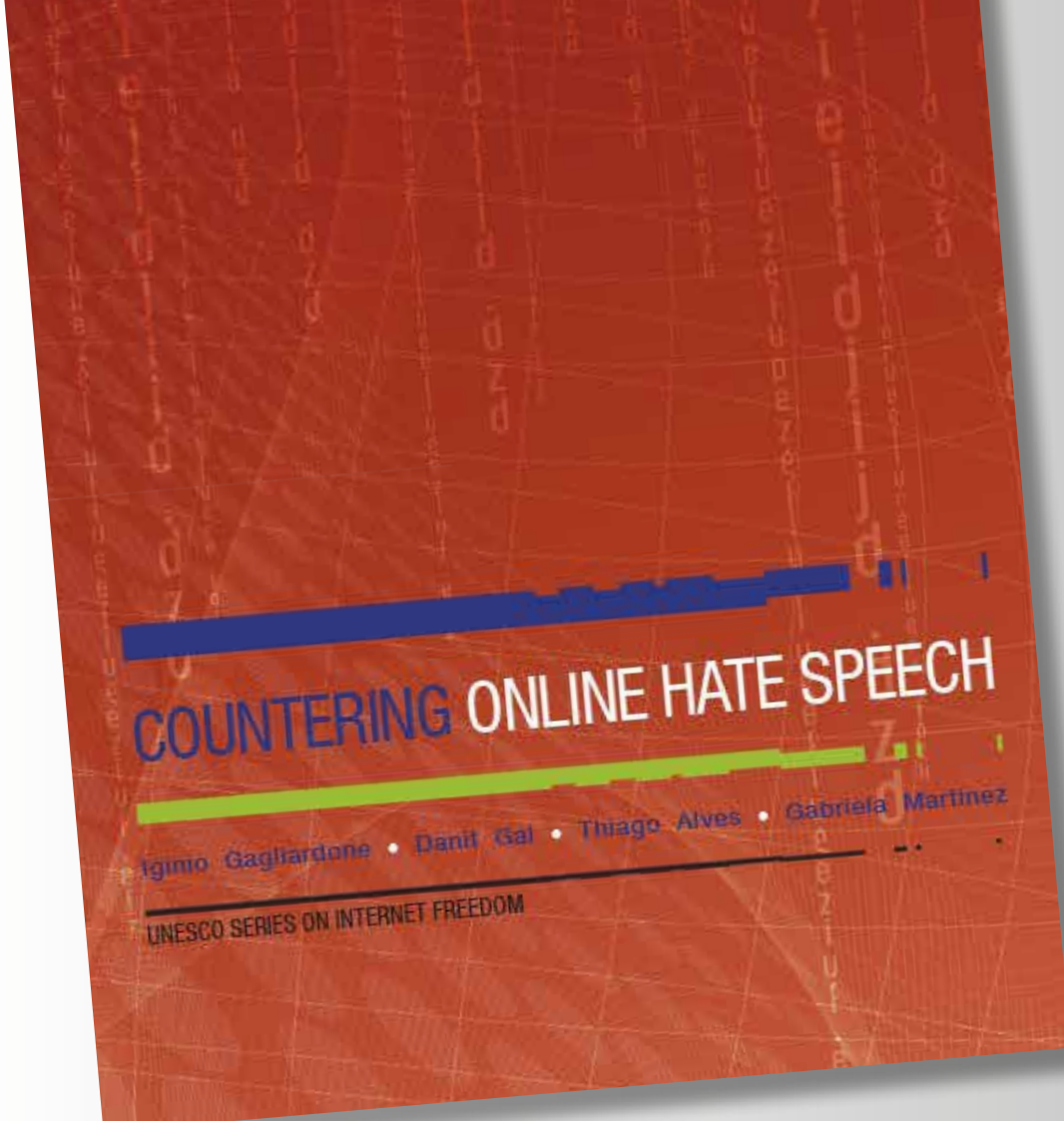




United Nations
Educational, Scientific and
Cultural Organization



Available online



The opportunities offered by the Internet vastly outweigh its challenges. While not forgetting this, we must nevertheless address the problems that come with it. Hate speech online is one such problem - examined in this study with an eye to practical counter-strategies.

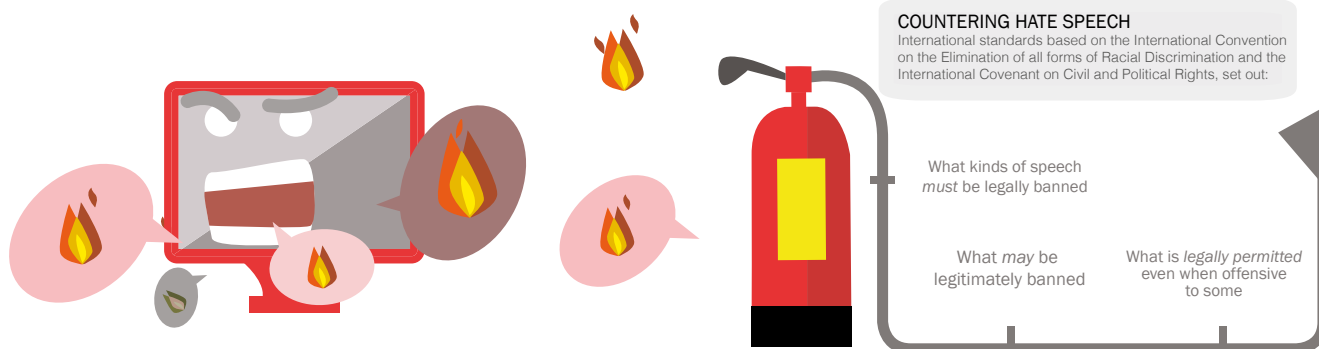
An understanding of the issues around hate speech can draw from international, regional and national normative frameworks, including international standards on freedom of expression. The complexity of defining and identifying hate speech, including online, often calls for tailored solutions.

The study points to specific elements to hate speech on the Internet, which call for multiple and broad-based responses, rather than entrusting only one or a limited number of actors to deal with the issue.

Special attention is given in this research to social and non-regulatory mechanisms to identify and counter the production, dissemination and impact of hateful messages online.

What is hate speech?

International standards and free expression



FOOD FOR THOUGHT – THE NEED TO CONSIDER CONTEXT

Hate speech online covers a very broad set of phenomena occurring on different technological platforms. The architectures on which these platforms are based, however, may vary significantly and have important repercussions on how hate speech spreads and can be countered. As the UMATI project reported in Chapter 4 of the study indicates, Twitter’s architecture makes it easier for users to shun those posting hateful messages and calls to violence. At the same time, the use of hashtags that gather different users around a common issue offer the possibility to share a tweet or a whole account, and allow processes of collective moderation. In the case of Facebook, similar mechanisms are less likely to occur because the platform allows multiple streams of conversations to develop partially hidden from one another. Other platforms, such as Snapchat, reduce the possibility that a message is shared beyond a very close group to a minimum. A more fine-grained understanding how each platform can enable or constrain the production and dissemination of different type of messages is thus an important factor in developing appropriate responses.

As with the right to freedom of expression, on- or offline, UNESCO defends the position that the free flow of information should always be the norm. Counter-speech is generally preferable to suppression of speech. And any response that limits speech needs to be very carefully weighed to ensure that this remains the exception, and that legitimate robust debate is not curtailed.

A distinction can be made between legitimate legal limits on certain hateful speech, and actions to counter those hateful expressions which do not cross the legal limits. International standards for legal limits are:

- For *race-related speech*, the International Convention on the Elimination of all forms of Racial Discrimination calls for a ban on expressing ideas of superiority or inferiority of people categorised by “race”.
- For *hatred on the basis of nationality or religion*, this is criminalised in terms of Article 20 of the International Covenant on Civil and Political Rights (ICCPR) – but with the qualification that the expressions involved should amount to (i) advocacy, which (ii) constitutes incitement to (iii) discrimination, hostility or violence.
- It is possible, but not required, that *hatred such as that based on people’s gender, sexual orientation or other features*, may be limited in terms of the ICCPR (Article 19).

But in all three cases, any limitations need to be specified in law, rather than arbitrary. They should also meet the criterion of being “necessary” – which requires the limitation to be proportionate and targeted in order to avoid any collateral restriction of legitimate expression which would not be justifiable.

This points to the importance of assessing apparent hate speech in concrete context – which affects how dangerous and influential this speech is likely to be, and how necessary in practice it is to restrict it. The UN’s Rabat Plan of Action provides a six-part threshold test for forms of speech that merit being prohibited under the application of criminal law. The test takes into consideration: the context of incitement to hatred, the speaker, intent, content, extent of the speech, and likelihood of causing harm.

International standards also require that any limitation of expression also has to conform to a legitimate purpose, and cannot just be an exercise of power. Besides for the objective of upholding the rights of others noted above, these purposes can also be national security, public morality or public health.

This complexity calls for an informed and nuanced approach to identifying and legally limiting “hate speech”. Nothing less can ensure an appropriate balance with the free flow of ideas and information. This is especially the case with regard to the world of online expression.

Responses

While the Internet is not separate from the realm of laws, there are risks and complications in a strategy that relies exclusively on developing and applying legal responses to perceived online hate speech. It is for this reason that attention should also be given to social responses, which may be considered as complementary to any legal limitations enforced by a state.

A typology of responses further elaborated in the study includes:

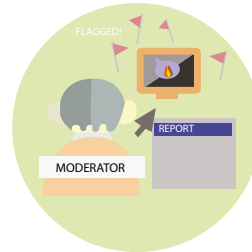
- **Monitoring and analysis by civil society** e.g. of how online hate speech emerges and spreads, set up of early warning systems, development of methods to distinguish among different typologies of speech.
- **Individuals promoting peer-to-peer counter-speech**
- **Organised action by NGOs** e.g. to report cases to authorities or create national and international coalitions to address emergent threats connecting online hatred and violence offline.
- **Initiatives encouraging social networking platforms and Internet Service Providers to increase active response mechanisms to online hate speech**
- **Structural responses through Media and Information Literacy**, empowering users through education and training about the knowledge, ethics and skills to use the right to freedom of expression on the Internet to interpret and react to hateful messages.



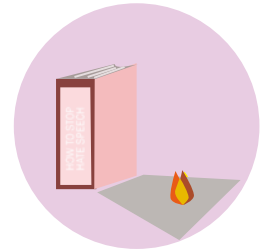
Monitoring and analysis by civil society



Peer-to-peer counter speech



Internet intermediary companies' mechanisms



Media and Information Literacy

Creative societal responses like these can produce results that reduce expressions of group hatred online. They can help ensure that the Internet remains a place of positive potential, and therefore that this network of networks will help us to build Knowledge Societies on the basis of peace, human rights and sustainable development.

Recommendations

The study assesses a set of case studies that analyse the definition of hate speech, taking into account Internet issues such as the jurisdiction of national governments and the role of transnational companies. Also examined in these cases is the character of hate speech online and its relation to offline speech and actions, as well as tried and tested methods to address occurrences of hate speech. From these assessments, several **key recommendations** follow:

- **States** should work collaboratively with organizations and projects that conduct campaigns to combat hate speech, including on the internet.

- **Internet intermediaries** – including internet service providers, search engines and social media platforms – should develop collective principles for proactively addressing online hate speech, in alignment with international human rights law.
- **Civil society organizations** should monitor hate speech and engage in counter speech.
- **Educational institutions and media organizations** should lead media and information literacy campaigns to teach the importance of user empowerment, critical thinking and peaceful dialogue on all media platforms.

THE NEED FOR MEDIA AND INFORMATION LITERACY

Initiatives promoting greater media and information literacy have begun to emerge as a more structural response to hate speech online. Given young people's increasing exposure to social media, information about how to identify and react to hate speech may become increasingly important. While some schools have expressed interest in progressively incorporating media and information literacy in their curriculum, these initiatives, however, are still patchy and have often not reached the most vulnerable who need the most to be alerted about the risk of hate speech online and offline. It is particularly important that anti-hate speech modules

are incorporated in those countries where the actual risk of widespread violence is highest. There is also a need to include in such programmes, modules that reflect on identity, so that young people can recognise attempts to manipulate their emotions in favour of hatred, and be empowered to advance their individual right to be their own masters of who they are and wish to become. Pre-emptive and preventative initiatives like these should also be accompanied by measures to evaluate the impact upon students' actual behaviour online and offline, and on their ability to identify and respond to hate speech messages.



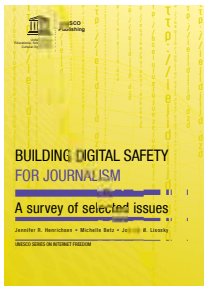
Upcoming:
Protecting Journalism Sources in the Digital Age

The legal frameworks that protect confidential sources of journalism are essential to reporting information in the public interest. However, these frameworks are under significant strain in the digital age, and there's a need to strengthen them. These are the findings of global research, undertaken for UNESCO. Read more about the study, its assessment tool, and gender dimensions of protecting journalism sources on UNESCO's website. This research also formed the foundation for a dedicated chapter in World Trends in Freedom of Expression and Media Development – Special Digital Focus 2015.



World trends in freedom of expression and media development: Special digital focus 2015

The publication explores emerging opportunities and challenges for press freedom in the digital age. With a focus on online hate speech, protection of journalism sources, the role of internet intermediaries in fostering freedom online, and the safety of journalists, the report highlights the importance of new actors in promoting and protecting freedom of expression online and off-line. In a media environment transformed by digital technologies, this special volume in the World Trends series is a key reference for Governments, journalists, media workers, civil society, the private sector, academics and students.



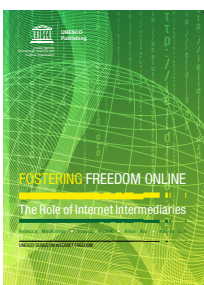
Building digital safety for journalism: a survey of selected issues

As technologies develop, so do opportunities as well as threats to journalism. This research explains some of the emerging threats to journalism safety in the digital era, and proposes a framework to help build digital safety for journalists. Examining 12 key digital threats to journalism, ranging from hacking of journalistic communications, through to denial-of-service attacks on media websites, it assesses preventive, protective and pre-emptive measures to avoid them. It shows too that digital security for journalism encompasses, but also goes beyond, the technical dimension.



Keystones to foster inclusive knowledge societies: access to information and knowledge, freedom of expression, privacy and ethics on a global internet

UNESCO's vision of universal Knowledge Societies builds on a free, open and trusted Internet that enables people to not only have the ability to access information resources from around the world, but to also contribute information and knowledge to local and global communities. What can UNESCO do to move towards the realization of this vision of Internet-enabled Knowledge Societies that can foster inclusive sustainable human development worldwide?



Fostering freedom online: the role of Internet intermediaries

With the rise of Internet intermediaries that play a mediating role between authors of content and audiences on the internet, this UNESCO publication provides in-depth case studies and analysis on how internet intermediaries impact on freedom of expression and associated fundamental rights such as privacy. It also offers policy recommendations on how intermediaries and states can improve respect for internet users' right to freedom of expression.



FOR MORE INFORMATION

<http://www.unesco.org/new/en/communication-and-information/>

While the research was conducted for UNESCO and funded by Sweden, ideas, opinions and findings are those of the authors, Iginio Gagliardone, Danit Gal, Thiago Alves, and Gabriela Martinez, they do not necessarily reflect those of UNESCO or Sweden and do not commit these parties.

