Training module

How to counter hate speech and manage an online community

ONLINE MODULE 4

Training module
How to counter hate speech and manage an online community

Introduction

The rise of hate speech in recent years has become a dangerous trend in our society. It poses challenges both to the media professionals and our democratic values in society. Media have a vital role to play in challenging hostility and prejudice, encouraging social solidarity and helping to promote understanding and empathy with others.

Media attention is often focused on the critical aspects of the presence of migrants and refugees and newspapers do not offer a correct image about migration phenomena: violent language and manipulated news to misrepresent migrants and other minorities reinforce stereotypes and fuel hatred against them. Media must show responsibility and actively work against hate speech, promoting a paradigm shift through the building of a counter-narrative.

This training module is based on the workshop realised in Rome on 20-21 November 2017 and organised by **COSPE** in cooperation with **Carta di Roma Association** and National Federation of the Italian Press (**FNSI**), within the **Media against Hate** project. The workshop consisted of a roundtable on hate speech, migration and journalism, with speakers from **Carta di Roma**, **Ethical Journalism Network**, **Article 19**, FNSI, and Facebook, and a three sessions workshop on specific topics. The event gathered 25 participants from various EU countries. It is based on the trainings material provided by Tom Law (**Ethical Journalism Network**), Matteo Moretti

(**Dataninja** – University of Bozen – Bolzano) and Valentina Vellucci (**Magilla Guerrilla**) and elaborated by COSPE.

The training module is addressed to journalists, bloggers, media activists, social media managers and professionals involved in the contrast of online hate speech.

1 The role of the journalists in preventing hate speech

The ethical journalism

Media portrayals influence public perceptions and attitudes and, especially for sensitive topics as migration, it's important to be aware of the effects of the information flow. In the era of social media the need of promoting a critical approach and a correct use of the media is even more urgent. The negative portrayal and stigmatisation of migrants, refugees and asylum seekers in the media have led to some people not considering them as human beings. A more careful and sensitive approach and a renewal attention to the ethics and value of the journalism is urgent.

The core values of ethical journalism are:
ACCURACY and fact-based communication
INDEPENDENCE: no propaganda
IMPARTIALITY: tell all sides of the story
HUMANITY: do no harm
ACCOUNTABILITY and transparency

Five- point test of speech

Online hate speech has been an issue both at the political level and in public discussion for some time now, but it has taken on particular importance starting from 2015. Before 2015, in Europe typical hate speech comments were predominantly directed at Muslims, for instance in the context of the Charlie Hebdo attacks, and at specific or unspecific (ethnic or religious) minorities. By early

summer 2015, hate speech comments were increasingly directed at refugees, especially Muslim refugees, and at people acting in favor of them. This development was clearly related to the European refugee crisis, which led to an increase of hate speech in social media.

How do journalists judge what is acceptable and what is intolerable? How do they embed in their daily work routine a way of assessing what is threatening?

It's a tricky task to judge exactly what constitutes hate-speech. There is no accepted international definition and the tolerance levels of speech vary dramatically from country to country.

To find a way through this minefield journalists must take into consideration the wider context in which people express themselves. They must focus not just on what is said, but what is intended. It's not just a matter of law or socially acceptable behaviour; it's a question of whether speech aims to do others harm, particularly at moments when there is the threat of immediate violence.

The following **five-point test** of speech for journalism in context has been developed by EJN advisers and is based upon international standards. It highlights some questions to be asked in the gathering, preparation and dissemination of news and information that will help journalists and editors place what is said and who is saying it in an ethical context.

Dis/Mis - Information

Fake News is an old problem given new life in the digital age. The EJN definition is "*Fake news is information deliberately fabricated and published with the intention to deceive and mislead others into believing falsehoods or doubting verifiable facts.*"

2 Possibilities and tools of data and visual journalism as a way to counter hate speech

Visual Journalism

Visual journalism can be defined as the combination of data journalism and visual storytelling. Data can't tell personal stories and for this reason visual journalism aims at return complex and multifaceted phenomena to a wider audience in a more engaging way.
Visual journalism is a precious resource to restore a more transparent and balanced information and debate. It empowers readers providing them with data and facts, activating and supporting them during the digital mediated debate and contrast hate speech.

How to design a visual/ data journalism project:
1#DATA: Data alone means nothing, connecting them create new stories (data storytelling)
2#STORY: Visualize them following some basic rules
3#GO PUBLIC: Embed your data in an engaging story framework

Best practice

People's Republic of Bolzano was the **Winner, Data Journalism Award 2015** and **Bronze Award, European Design Award 2016**.

The project People's Republic of Bolzano reshapes the identity of the local Chinese community through digital media, in order to restore more transparent and balanced information, allowing a broader audience to inform itself on such a complex and multifaceted issue. This small project is part of an emergent phenomenon to counterbalance misrepresentation, in this case over the issue of migration.

The project started with an analysis of how the Alto Adige newspaper was depicting the Chinese migration in the area as an "invasion". To counter this narrative, interactive digital artefacts, including data visualizations and videos, have been constructed and disseminated. In this perspective, design becomes "public design", and it aims to expose controversies and to participate in collective

meaning making, through the re-presentation, re-designing, re - informing, re-aligning, re-reading of information.

3 Creating and managing a community in order to prevent and counter hate speech online

Being a journalist nowadays is as much about community building as about news reporting.
It is important to radically rethink journalistic activity within the new digital framework, an activity not circumscribed by the printing or broadcasting of a piece but which continues in the flow of conversations that seek interaction with users and try to follow cues and suggestions for new articles.

Social media policy and active moderation that is prepared and participatory, as well as an inclination to the production of third-party content gathered from the flow of conversations – in short, a comprehensive interaction with the community – have become priorities. This requires social media managers, community managers and content curators, but above all it requires every editorial entity to plan for and establish such professional roles through studying and understanding the new editorial dynamics that the digital revolution demands.

How social networks are trying to face hate speech?

FACEBOOK

The social media giant uses a combination of artificial intelligence (AI) and human content reviewers to find hate speech, which, The Verge says, removes "66,000 hate mail posts per week". But the site says Facebook "relies heavily" on users reporting content as offensive or hateful. In May, Facebook hired an additional 3,000 people to its "team of content reviewers", says TechCrunch, bringing the total up to 7,500. This was triggered by several global "content moderation scandals", including the use of Facebook Live "to broadcast murder and suicide." But Wired says there have been situations where the company's "algorithmic and human reviewers" have labelled comments or posts as offensive without considering the context. For instance, the website says some words

"shouted as slurs" are sometimes "reclaimed" by groups "as a means of self-expression."

TWITTER

Twitter's hateful conduct policy says the service does not "allow accounts whose primary purpose is inciting harm towards others" on the basis of areas such as race or sexual orientation. The Independent reports that over the past months the site has introduced more systems and tools to detect and remove hate speech, as well as improving the process where its users manually report offensive material. But the social media site has been in a "fair bit of hot water in recent months regarding a perceived lack of action in the wake of perceived threats", says TechCrunch.

GOOGLE

While Google isn't exclusively a social network, the tech giant plays a key role in directing internet traffic and the social apps that users can access. Google says it does not support "content that promotes or condones violence against individuals or groups" based on certain criteria, adding that it depends "heavily upon users to let us know about content that may violate our policies".

Best practices

Guidelines for community and social media managers

"BRICkS – Building Respect on the Internet by Combating Hate Speech" is a project promoted by COSPE and other EU organizations, aimed at contrasting the spread of online hate speech against migrants and minorities through media literacy and active involvement of web users and web content producers. The results of the project ended in 2016 include an some national studies and an European research about the management of hate speech in online newspapers and the guidelines addressed to social managers:

La Stampa – visual netiquette

The term netiquette is commonly used in reference to popular forms of online communication, including email, forums and chat. While some netiquette issues are technical in nature, all fundamentally concern how people relate to each other. Once the technical nuances are grasped and abstract concepts are understood, online etiquette is not much different than offline etiquette. *La stampa*, an Italian newspaper that allows comments only on Facebook created a visual netiquette. When they face a discriminatory comment, they don't delete it nor hide it but they post what they call in Italian a "galateo" (etiquette) under the comment, an image reminding the rules of online participation (see below).

ATTACHMENT
- Tom Law's presentation
- Matteo Moretti's presentation
- Valentina Vellucci's presentation