Workshop ‘’Social Media Monitoring on Hate Speech’’

On Tuesday January 19th, 2021, we organized an online workshop for monitoring hate speech on social media within the Get the Trolls Out project, the workshop was taught by Eline Jeanne, project coordinator at the London-based Media Diversity Institute (MDI).

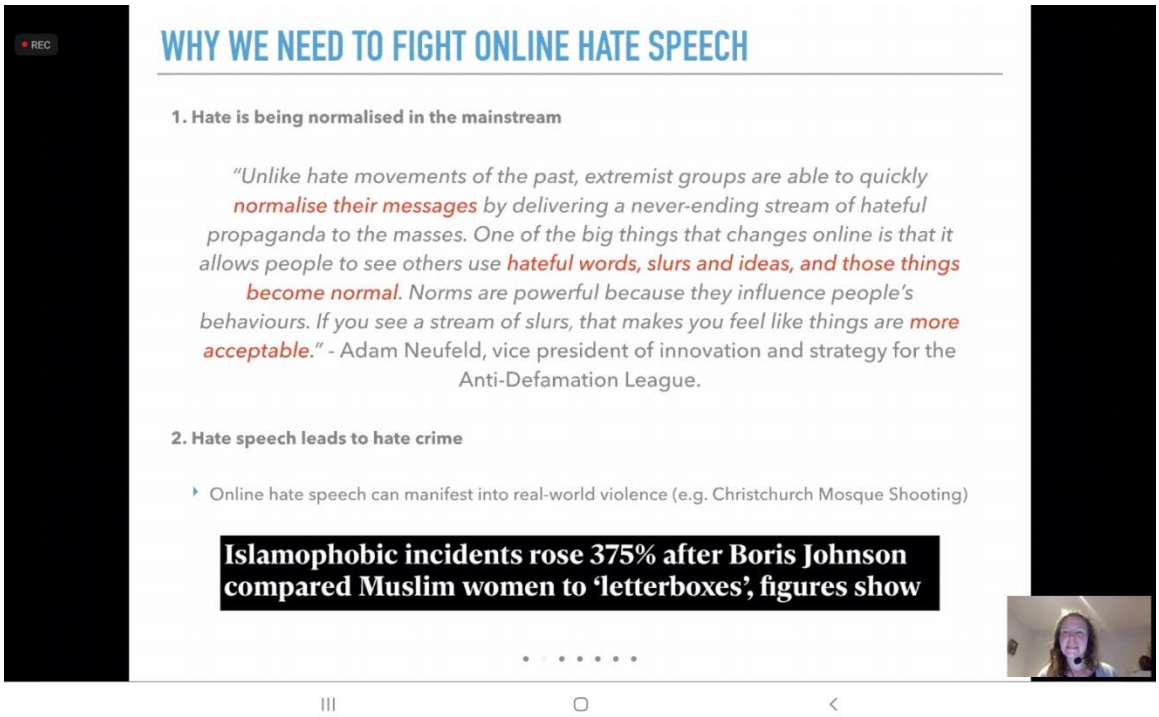
First of all, in preparing the workshop we were very glad of the interest shown by many people from different professional backgrounds to attend it. In Greece, it is not something that is done systematically by many organizations, so we were surprised and incredibly happy by the interest shown for the workshop. On that basis we thought it would be useful first of all to see how widespread or not is the idea of monitoring hate speech on social media in Greece, who engages in that practice and also invite people to discover this practice and include it from now on in their work in a systematic way. Usually when the term media monitoring is used it is applied to mainstream media, when it is done systematically at least.

People who work in related jobs to combatting hate speech, have not necessarily incorporated this technique in their work, so it was a good chance to give them another tool in their fight against hate speech.

The public discussion of whether or not there should be de-platforming of Greek personalities with a large following or/and position of power who spread hate speech on social media, that a lot of times leads to real life violence, is not as present as it is in other countries. Something that was discussed during the Q&A session of the workshop and how the lack of accountability of those figures incites violence and spreads hate speech with many times catastrophic consequences.

As mentioned before the people that wanted to participate and that ended up participating were academics, NGOs employees, researchers, journalists, policy makers etc. After we presented the Get the Trolls Out project and the work, we do within it concerning monitoring anti-Semitism and islamophobia, we asked if the participants if they are aware of what social media monitoring is and if they have used it in their line of work: a lot of them answered yes to one or both questions.

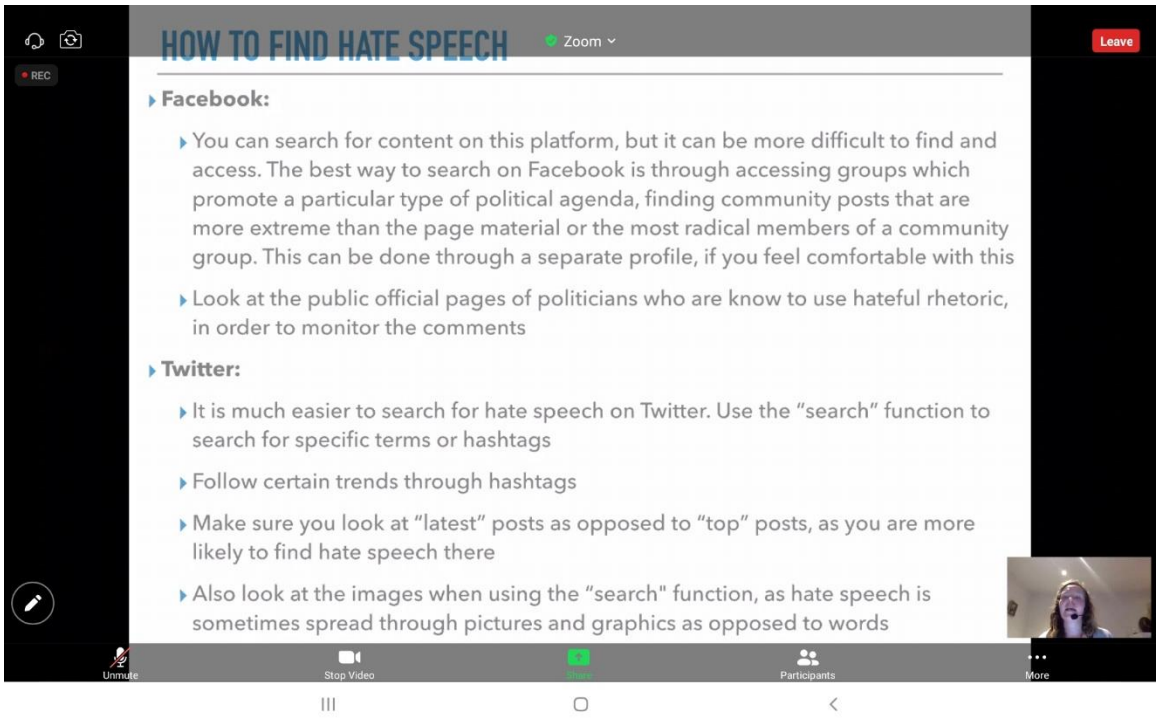
So, we had a good starting point in terms of familiarity with the subject. During the first exercise where they had to identify if the examples of hate speech on social media could be reported, the participants were correct in their assumptions and seemed to understand what can and can’t be reported according to each platform’s guidelines.



Then we split them into groups so they could find and report social media hate speech and practice what they have learned. When we discussed the results of the exercise, we all came to the unfortunate conclusion that it was very easy to find hate speech on social media ranging different topics (homophobia, anti-semitism, anti-migrant propaganda etc). Sometimes they didn't even have to use slurs in order to find the hate speech content it was just enough to use the terms that refers to protected groups for example: One of the participants just used the acronym LGBTQI, and a bunch of homophobic tweets came up without having to search with derogatory terms used by homophobes. This goes to show how pervasive hate speech has become in our daily use of social media.

Another interesting remark came from the group that chose Facebook for their monitoring exercise. According to the participant it was more difficult to find anti-immigrant hate speech through their account since facebook will renew each feed depending on one's interests and will show pages related to those interests. So they tried finding hate groups within the platform where they spotted hateful comments but were not sure if they were ''reportable'' or not according to the platform's guidelines. Lastly, they mentioned that they spotted comments which they called ''hidden hate speech'' meaning it was phrased in a way that it was open to interpretation and could slip through the guidelines and not be reported. Our instructor advised them when a comment is in the ''gray area'' there is no harm in reporting it anyway even if we are unsure or if it will be removed.

Another pertinent issue that came up during the workshop is how often politicians and people with large social media followings propagate hate speech to their followers which can lead to real life violence often, without explicitly describing violent acts or urging their followers to do them but incite them, nonetheless. In these cases, because they know how to use the medium in their favor when those real-life violent acts occur, they can denounce responsibility.



During the Q&A we found out that not everybody could use the monitoring directly in their work for various reasons, however one idea that came of the workshop was for big NGOs that combat injustice, hate speech etc and have large followings to involve the community of their followers and educate them as to how they can do it themselves and help them in their work. However, most of the participants declared that they would and can use it directly in their work.

We as Karpos would like to thank Eline Jeanne one more time for the excellent work she did as instructor of the workshop and we are hoping to do it again due to the demand we had from people to participate.