IRISH NETWORK AGAINST RACISM:

NETWORK POSITION ON ONLINE HATE SPEECH

5th March 2020

INAR and its membership support the re-iteration by the UN Committee on the Elimination of Racial Discrimination (CERD) of Article 4 of the International Convention on the Elimination of Racial Discrimination (ICERD), requiring member states and organisations to curb incitement to racial discrimination or violence, the propagation of ideas of racial superiority, and to prohibit organisations that promote or incite discrimination. At network meetings, INAR members frequently report being targets of online hate speech. 46% of surveyed members of Uplift, an INAR member organisation, have reported experiencing hate speech online. INAR and others[i] have identified gaps in self-regulation by social media platforms, considering the frequent use of their platforms by individuals and organisations who stir hatred, potentially inciting racially motivated violence. The concerns of INAR[ii], shared by CERD[iii], stem from the noted increase in incidences of online racist hate speech and online hate organising, and the correlation between these specific instances of racist sentiment and organising online, and recorded incidences of racially motivated violence.

CERD Obligations and Social Media Platforms

CERD's 2019 Concluding Observations on the combined fifth to ninth reports of Ireland, regarding racist hate speech, include the following recommendations for the State. INAR members believe these imply an obligation on social media platforms to cooperate in their implementation:

➔ Intensify efforts to tackle the prevalence of racist hate speech on the Internet and social media, in close cooperation with the Internet service providers and social media platforms;

➔ Ensure that the online safety and media regulation bill, which is under development, is in line with international human rights standards and expeditiously enact the bill;

➔ Effectively investigate and, as appropriate, prosecute and punish acts of hate speech, including those committed by politicians during election campaigns;

➔ Ensure that the mandate of the Electoral Commission, once established, includes provisions on the prohibition of racist hate speech in the context of elections;

Social network platforms, such as Facebook, are also signatories of the European Commission's Code of Conduct on Countering Illegal Hate Speech Online[iv], strengthening their obligation to adhere to CERD Observations.

Facebook Community Standards

Facebook's own Community Standards[v] state that the company takes pride in policies which are inclusive and ensure users that;

> *"expression that threatens people has the potential to intimidate, exclude or silence others and isn't allowed on Facebook"*.

> *"We do not allow hate speech on Facebook because it creates an environment of intimidation and exclusion, and in some cases, may promote real-world violence. We define hate speech as a direct attack on people based on what we call protected characteristics – race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity and serious disease or disability. We also provide some protections for immigration status. We define "attack" as violent or dehumanising speech, statements of inferiority, or calls for exclusion or segregation."*

Speaking to the Oireachtas, Facebook Ireland's public policy lead Dualta Ó Broin said; *"We take our role in keeping harmful content off our services very seriously"*[vi], and highlighted the role of platform users in reporting content which violates community guidelines.

Gaps in regulation

But there remains a gap between Facebook and other social media platforms' stated aspirations to maintaining a safe environment, and the everyday reality for minority

and other users who do report. In this regard, INAR and its members are concerned at an apparent lack of transparency, regarding the decisions by Facebook that offending content, such as calls for the sterilisation of members of the Traveller Community, or for the deportation of people of African descent from Ireland, should not warrant removal. This is despite being in clear violation of Facebook's own Community Standards.

Investigations[vii] by INAR have found consistent failings by social media companies in addressing reported hate content, even in those instances escalated by INAR using its Trusted Flagger status. INAR members also cite gaps in the regulation of the content of political advertising on Facebook, as well gaps in policy relating to pages and individuals who violate Facebook Community Standards on multiple occasions. INAR and its members are concerned that social media platforms' frequent non-enforcement of their own community standards constitutes a situation of non-adherence to the EC Code of Conduct, to UN CERD and Council of Europe recommendations and to EU Directives.

The harm of failing to regulate hate content

Recent increases in recorded online hate speech have been linked by INAR to the increased coordination by hate groups, who utilise social media platform features such as closed groups and community pages to recruit, propagandise and incite racial violence. INAR and its members further connect these increases in reported hate content to the orchestration of hate targeting and racially motivated public demonstrations. As highlighted by both INAR and CERD, these developments are linked instances of serious violence, including the recent arson attacks on proposed locations for Direct Provision centres.

INAR Network Members recommendations:

INAR's position is that Facebook should have a duty of care on its platform. For the safety of its users, as the largest and most influential social media platform in the world, Facebook should;

➔ Take a strong, proactive position in the education on the effects of hate speech, and the promotion of Facebook Community Standards.

- Encourage and enable users to report instances of online hate speech more easily.

- Have the best trained content review teams to quickly respond to instances of Online Hate Speech and make the decisions on reviews transparently.

- Engage with local specialist NGOs to ensure staff training and AI learning incorporates local and historical context (including vocabulary).

- Improve response times for reports of racist hate speech on Facebook.

- Engage and cooperate with specialist NGOs, Trusted Flaggers, and An Garda Síochána to protect those threatened with real-life violence.

- Monitor pages which are set up with the specific intention of inciting racist hate speech.

- Improve Facebook policy to address serial offenders who utilise Facebook to stir racial hatred.

- Address the use of Facebook to distribute unverified political advertisements, and to ensure all advertisement on Facebook is in line with the Irish equality legislation on non-discriminatory advertising standards.

[i] Rogan, A. (2019) Social media giants fail to act on hate speech. Available at: https://www.businesspost.ie/news-focus/social-media-giants-fail-to-act-on-hate-speech-3ae2b233
[ii] INAR (2019) Alternative Report on Racial Discrimination in Ireland. Available at: https://inar.ie/wp-content/uploads/2019/11/INAR-ALTERNATIVE-REPORT-to-CERD-WEB.pdf
[iii] CERD (2019) Concluding observations on the combined fifth to ninth reports of Ireland. Available at: https://tbinternet.ohchr.org/Treaties/CERD/Shared%20Documents/IRL/INT_CERD_COC_IRL_40806_E.pdf
[iv] European Commission (2016) Code of Conduct on Countering Illegal Hate Speech Online. Available at: https://inar.ie/wp-content/uploads/2019/11/INAR-ALTERNATIVE-REPORT-to-CERD-WEB.pdf
[v] Facebook (2020) Community Standards. Available at: https://www.facebook.com/communitystandards/
[vi] Oireachtas (2019) Joint Oireachtas Committee on Justice and Equality Opening Statement. Available at: https://data.oireachtas.ie/ie/oireachtas/committee/dail/32/joint_committee_on_justice_and_equality/submissions/2019/2019-10-09_opening-statement-dualta-o-broin-head-of-public-policy-facebook-ireland-ltd_en.pdf
[vii] Rogan, A. (2019) Social media giants fail to act on hate speech. Available at: https://www.businesspost.ie/news-focus/social-media-giants-fail-to-act-on-hate-speech-3ae2b233