Promoting Fundamental Freedoms and Tackling Hate Speech and Violence Online

Examining the problem of gender-based hate speech and violence

Anita Gurumurthy
IT for Change

Gender-based hate speech: Research findings

Profitable Provocations

Pervasiveness Of Misogynistic Speech

All women in our sample regardless of their political or ideological standpoints received some amount of abuse on the platform. No one was entirely spared.

Herd Aggression

Trolls tended to strategically target certain women and certain posts to exploit the affordances of virality and the algorithmic amplification of content.

"Light-hearted" Trolling

A majority of abusive messages directed at the women were of a supposedly milder variety of tongue-in-cheek jokes and remarks.
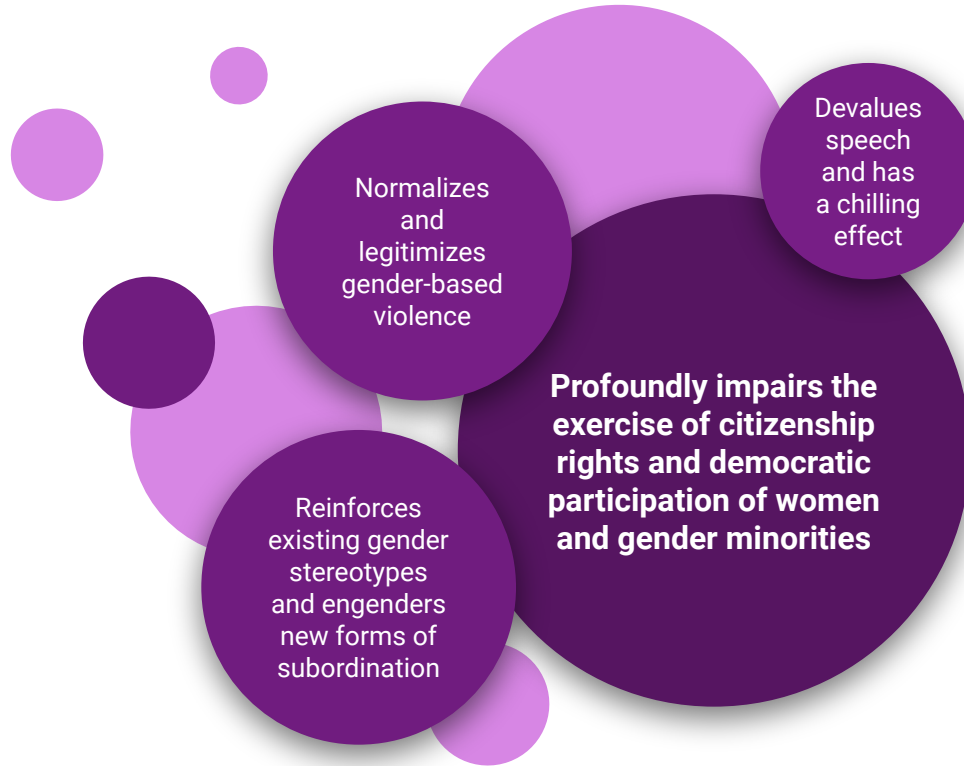
Intersectional Violence

Muslim women, political dissenters, and political commentators received an overwhelming majority of abusive messages.

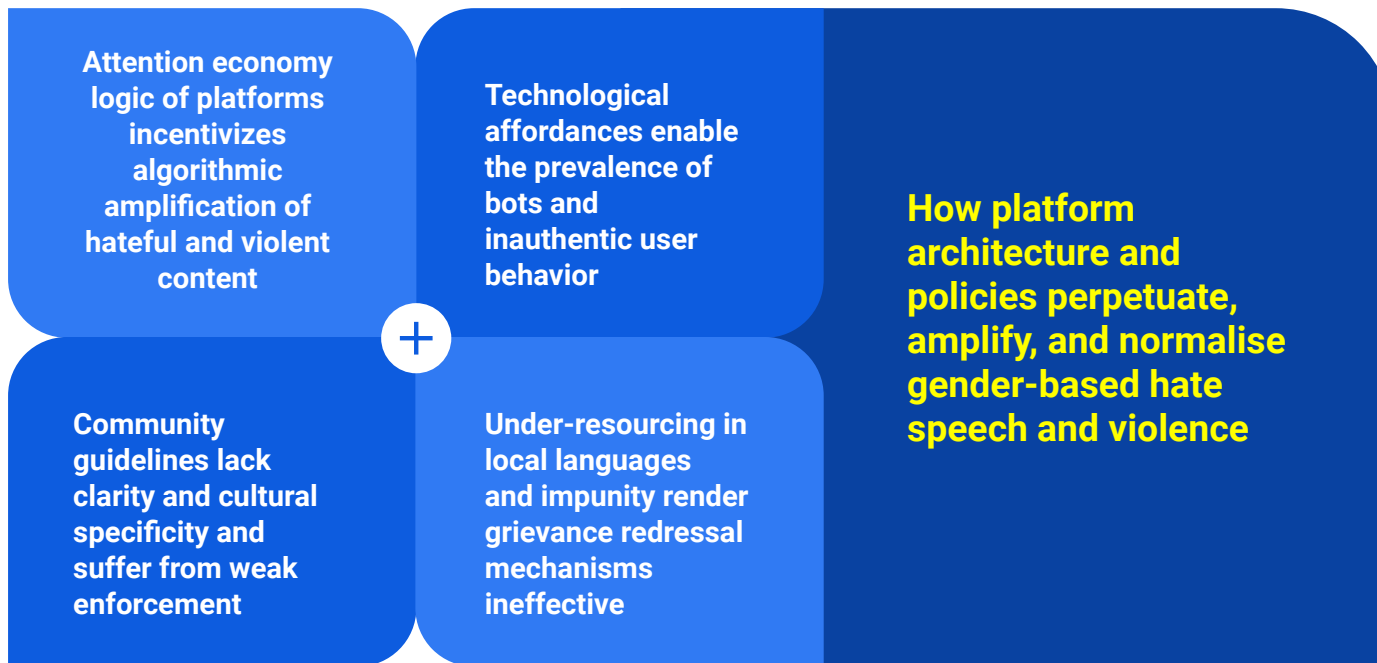- Gender-based hate speech is a veritable threat to the democratic participation of women and gender minorities, particularly those in public-political life and/or belonging to marginalized groups.

- Gender-based hate speech on digital platforms is characterised by the **speed, volume** and **frequency** of hateful messages.

- What we see online - overt threats of violence to "light-hearted" trolling.

Gender-based hate speech: a threat to democratic integrity

Normalizes and legitimizes gender-based violence

Devalues speech and has a chilling effect

Profoundly impairs the exercise of citizenship rights and democratic participation of women and gender minorities

Reinforces existing gender stereotypes and engenders new forms of subordination

Gendered hate and violence is a feature - not a bug - in platform design

Attention economy logic of platforms incentivizes algorithmic amplification of hateful and violent content

Technological affordances enable the prevalence of bots and inauthentic user behavior

+

Community guidelines lack clarity and cultural specificity and suffer from weak enforcement

Under-resourcing in local languages and impunity render grievance redressal mechanisms ineffective

How platform architecture and policies perpetuate, amplify, and normalise gender-based hate speech and violence

What needs to be done?

1. Recognition of the problem - and the right to publicness

- Gender based hate speech and associated role of platforms in perpetuating such violence a critical issue for digital governance.

- 'Gender' as a ground for hate speech needs to be recognized.

- UN Special Rapporteur Irene Khan's recommendation (2021)
  EU Directive - 'Cyber incitement to violence or hatred on grounds of sex or gender - Combating Violence against Women (Article 10).
  Canadian law has recognized online hate speech on grounds of gender.

- A gendered perspective on the right to freedom of expression must include a framework for promoting positive freedoms (freedom for), such as the right to public participation, and the right to publicness of digital space, as well as negative freedoms (freedom from).

Right to publicness :

- Publicness may be seen to refer to:

  a specific mode of relationship among people based on visibility and access, which is essential for the processes of collective self-understanding constitutive of democratic societies.

- In the context of the silencing effect of hate speech and misogynist trolling, the right to publicness, as opposed to elaborations of the legal right to freedom of expression or public participation, would emphasise that the **right to be heard be placed on an equal footing as the right to speak.**

2. A gender-responsive platform regulatory framework

The platform regulatory model must move towards tackling an attention economy-focussed approach - a model that seeks accountability for the hostile and abusive online environments that platforms foster and profit from.

2. A gender-responsive platform regulatory framework (contd.)

Strong platform liability framework	Independent National Regulatory Body	Human rights impact assessment with focus on gender	Transparency and data access
Hold platforms accountable for enabling or facilitating harms including online gender-based violence, disinformation, hate speech, incitement to violence and for any systematic or deliberate failure to take steps to prevent or mitigate the harm.	Establish an independent regulatory body, with investigative, inspectorial, supervisory & other functions over platforms to ensure their compliance with prescribed standards. (Recommended by UNESCO Guidelines on Regulating Digital Platforms, Canada's Online Harms Act)	Institute human rights impact assessments to identify systemic risks to the rights of women and gender minorities, especially from marginal social locations, arising from platform operations, including their algorithmic systems. (EU DSA, Art 34)	Mandate proactive transparency about the design, development, and deployment of content moderation systems and content curation systems, and redressal of grievances. Make data access (gender-disaggregated) for public interest research on illegal and harmful content a requirement (EU DSA, Art 40)

3. Gender-sensitive techno design changes

Platforms must be bound by law to arrest algorithmic amplification of misogynistic content as a commitment to their statutory duty of care owed to users.

01	Capabilities to arrest the viral spread of misogynist content	• Invest in systems to predict the reshare cascades of content • Implement steps like internal circuit breakers to arrest algorithmic amplification
02	Friction in design	• Introduce friction, through contextual information about content and/or warning labels. • Helpful to combat gender disinformation
03	Multidisciplinarity in engineering and safety teams	• These teams should be equipped equipped to develop algorithmic solutions for various types of gender-based violence
04	Technical capability in regional languages	• Develop ML tools to detect problematic speech in regional languages.

Conclusion - Key messages

- The crisis of information integrity on digital platforms is reflective of de-democratisation

- It cannot be addressed without effectively tackling the role of platforms in delegitimizing gender equality.

- States have a duty to acknowledge and address the systemic undermining and silencing of women and girls by the ruthless logic of digital marketplaces. Ignoring the issue sets us back in the collective struggle and movement for gender equality everywhere.

- Efforts to tackle gender-based hate speech and violence should be rooted in a right to publicness of digital space; the right to be heard as equal to the right to speak

- Gender considerations in platform regulatory approaches are an urgent imperative - a duty of care that must be mandated by law

- Regulatory oversight over moderation practices of platforms must not be limited only to decisions to takedown or reinstate content, but also extend to design choices and algorithmic processes of amplification of content.