

Global Alliance Against Digital Hate and Extremism

We are an international movement mobilising collective action for a safe, fair and just online experience for humans everywhere. The evidence is clear: harmful behaviour permitted and exacerbated by social media companies - the spread of disinformation, hate speech, incitement to violence, and extremist propaganda online - leads to offline violence, including genocide, and threatens human rights and democracies.

Our organisations have witnessed these harms first hand. Social media companies are inadequate in their so-called efforts to mitigate harm in the world, and to live up to commitments they have set forth in their own Community Standards and Terms of Service, and in the public commitments they have made to the UN's Guiding Principles on Business & Human Rights.

Indeed, their actions are fueling harms globally.

An unfair and discriminatory approach to providing resources, particularly outside the U.S., to protect the integrity of the platforms and the safety of the users has led to a gross breach of faith in the social media companies. Real-world, systemic issues - from far-right extremism to genocide to gender bias and racism - are made worse by the business practices of major social media companies. We cannot strengthen democracies and protect human rights while companies amplify and reward hate and extremism with no accountability.

Our movement will focus on making progress in the following four areas:

1. **Fairness:** for companies to afford appropriate attention and resources to the experience of platform users in the Global South as those in the U.S.
2. **Justice:** for companies to adapt their policies to mitigate regional and local harms, including genocidal campaigns and denial, when identified by civil society.
3. **Safety:** for companies to live up their commitments to keep users safe online and prevent the organization of offline violence.
4. **Transparency:** for access to real time information on the allocation of resources, content and account moderation decisions, and the application of policies across countries.

Priorities

- Define “dangerous individual and organization” designation policies and their enforcement, with the input of civil society, not only in the U.S. and Europe, but in the Global South.
- Expand and ensure proportionate language and cultural competency for all regions.

- End exemptions to content moderation for the politically powerful and influencers globally, implement fact-checking for political ads.
- Fix algorithms to end their amplification of disinformation, hate and extremism.
- Full transparency from tech companies on their policies, systems, practices and internal research relevant to disinformation, hate, and extremism globally.