Online Harms and a thriving democracy

What the government has proposed

On 12th February, the Government issued their initial response to the consultation on the Online Harms White Paper which examined how to regulate tech companies to reduce 'Online Harms'. The Government is seeking to establish a legal Duty of Care for reducing the risk of harms to their users and "is minded" to appoint Ofcom as the regulator with responsibility to oversee tech companies' compliance. The Government's stated goals were to:

- Promote fair and efficient markets where the benefits of technology are shared widely across communities;
- Ensure the safety and security of those online; and
- Maintain a thriving democracy and society, where pluralism and freedom of expression are protected.

Welcome progress

We welcome:

- The ambition to establish a duty of care that respects freedom of expression.
- The clarifications given by the Government that this regulatory regime will operate at the level of the systems that the platform companies design, not at the level of individual pieces of content placed on platforms by users.
- The prominence given to some threats to democracy including online abuse and hate speech.

Where the latest proposals fall short

However, in key areas the Government has not backed up the ambition set out in the White Paper. In doing so they risk falling short of their goal of world-leading regulation that helps maintain a thriving democracy, safe for all UK users.

The Online Harms White Paper - released in April 2019 - highlighted a number of harms that are exacerbated by decisions the tech companies make in the design of their products. Yet the initial consultation response is largely silent on how regulation will address these:

- **Disinformation** is used to disseminate hate speech, confuse citizens' understanding of vaccines, and to suppress voter turnout among already-marginalised groups. As the White Paper pointed out, the tech platforms themselves make decisions that in large part determine the prevalence and impact of disinformation.
- **Manipulation of the information environment**: as the White Paper rightly set out, "[a] combination of personal data collection, AI based algorithms and false or misleading information could be used to manipulate the public with unprecedented effectiveness." For example, the platforms' design systems encourage users to go from a video about

vaccinations to a series of anti-vax videos. To hold the attention of users, the platforms design their algorithms with the effect that they disproportionately boost hateful, divisive or misleading content because it is attention-grabbing.

- **Abuse and intimidation, especially of public figures who are women or from a minority background.** This constitutes a threat to the healthy public debate that is essential for our democracy. Black, Asian and Minority Ethnic (BAME) women MPs receive 41% of abusive tweets. Abuse often takes the form of threats of sexual violence, with one UK MP receiving 1000 sexually abusive Tweets per week. It also has what Amnesty UK have identified as a 'silencing effect', particularly on marginalised groups like women and girls, and BAME communities.

How the Online Harms regulation can help

A statutory, systemic duty of care - if properly established - would enable the regulator to look at the processes involved in platform design and how they exacerbate or reduce online harms. For instance:

- **Transparency about the impact of design decisions**: a 'systems approach' was recommended by the independent advisory body, the Centre for Data Ethics and Innovation: the Duty of Care should mean that companies take responsibility for the system design decisions and processes underpinning their services, not individual pieces of content. The regulator should have the authority and capability to scrutinise:
    - How design decisions may incentivise harmful (but legal) content to be posted or shared; and
    - The processes the tech companies use to target and disseminate content. This would require a means for the regulator to access data held by the companies without undermining users' privacy or companies' intellectual property.
- **Redress**: while it is welcome to see such attention given to redress in the Government's response, the Duty of Care should include better protections for users. If platforms' own community standards are to be the primary means to hold platforms accountable, then users and civil society should have a robust, ongoing, and upstream role in shaping community standards. Further, transparency about how complaints are dealt with is only useful if there are harm-specific benchmarks for complaint levels and outcomes, and if data is available to judge independently whether the number of complaints is proportionate to the levels of corresponding harm.
- **Ensure online safety applies for all:** the regulatory framework must take into account the disproportionate impact of abuse on individuals from marginalised and vulnerable groups, including: women, and particularly women of colour; people with disabilities; and religious minority communities. As the Lord's Communication Committee reported, 'the most vulnerable people in society are particularly susceptible to online harms, but they are less likely to develop digital literacy'.
- **Education and training:** this must go beyond 'empowering users to manage their online safety' and towards digital citizenship education, equipping individuals with skills to practise forms of social participation that are respectful of the human rights and dignity of everyone, especially vulnerable and marginalised groups.

For more information please contact Nick Martlew at Digital Action, who coordinated this briefing: Nick@digitalaction.co