Background Paper - Disinformation Online: Reducing Harm, Protecting Freedoms

Many international partners are seeking to develop new regulatory approaches to tackle online harms. International fora such as the Internet Governance Forum (IGF) provide an opportunity for stakeholders to share experiences as well as learn from partners This proposed panel session on disinformation at the 2019 IGF will engage representatives from government, civil society, academia and industry to discuss the problem, emerging challenges and effective solutions.

Session overview and objectives

Disinformation is a multifaceted problem with no single solution. It is a global issue, with many countries concerned about its potential harmful impact on security, health and societal cohesion. The objective of this panel session is to discuss responses to disinformation, drawing on international examples and views from government, industry, civil society and academia, and encouraging cooperation and collaboration among partners.

This panel will focus on the following areas and policy questions:

1. *Developing regulatory approaches to tackling disinformation while upholding freedom of expression.*

This part of the session will discuss options for a regulatory approach to tackling disinformation that protects users from harm while upholding freedom of expression, a key issue that many countries are grappling with.

Key policy questions:

- What is the role for regulation in tackling disinformation?
- How can regulatory regimes ensure freedom of expression is protected?

2. *Addressing vulnerabilities in the online environment.*

Disinformation is not a new problem but the online environment has provided very simple tools that allow disinformation to be disseminated quickly, at low cost, and often targeted with high precision. Furthermore, the tools that exist offline to help people assess information are often not available online. In response, industry has developed tools to detect and flag instances of disinformation. This includes the social media platforms themselves who have been under intense pressure from governments, the media and the public to do more. This part of the session will discuss technological solutions to disinformation and discuss industry's role in tackling it.

Key policy questions:

- How can technology be used to tackle disinformation?
- What role should service providers play in tackling disinformation on their platforms?
- How can Internet platforms and media outlets work together to fight disinformation?

3. *Audiences: Impact, public perspectives on the problem, and the role of education.*

Internationally there is great concern about the potential harmful impact of disinformation on societies. Evidence points to increasing public concern about disinformation yet suggests that audiences are often not well equipped to recognise it. There is growing recognition internationally of the importance of education in tackling disinformation, and initiatives have been launched across the world to boost digital and media literacy in response to this challenge. This part of the panel session will invite discussion on educational responses to disinformation, the success of various initiatives launched globally, and how to reach vulnerable audiences.

Key policy questions:

- How can audiences' resilience to disinformation be increased?
- Who are the vulnerable audiences?
- Are audiences informed about disinformation?
- How can the impact of disinformation on audiences be measured?
- How do the public perceive the problem?

4. *Emerging challenges*

Disinformation will ultimately continue to evolve with technology. Developments in AI make it possible to generate fake content (text, audio and video) - known as 'deepfakes' - and it could become increasingly difficult for humans and algorithms to detect this content. Immersive and augmented reality technologies could also present opportunities for hostile actors to try to manipulate people. The final part of the session will discuss emerging challenges and seek to progress the international conversation on to finding solutions to tackle wider online manipulation.

Key policy questions:

- What are the key emerging challenges in this area?
- How should governments responding to emerging technological challenges, and those that do not yet exist?
- How can governments respond to wider forms of online manipulation?

<u>Background</u>

The global technology and online landscapes present increasingly complex risks and opportunities. Despite the great benefits the internet and other digital technology has brought for the economy and wider society, it has also brought a host of unintended consequences. The threats facing democracies across the world are increasingly of a digital nature, including disinformation and emerging technologies.

Democratic society is built on confidence in public institutions; trust in electoral processes; a robust, lively and plural media; and hard-won democratic freedoms that allow different voices, views and opinions to freely and peacefully contribute to public discourse. However, disinformation - information which is created or disseminated with the deliberate intent to mislead; this could be to cause harm, or for personal, political or financial gain - threatens these values and principles, and can threaten public safety, undermine national security, fracture community cohesion and reduce trust.

Disinformation, propaganda and attempts to mislead are by no means new issues. However, the online environment has enabled disinformation to spread faster and reach larger audiences than before. Barriers to entry are small and the tools used to spread disinformation vary dramatically. Very simple tools allow high volume, low cost disinformation to be disseminated quickly and easily at low cost. But there are more sophisticated actors too, who use fake accounts and bots to manipulate and distort the online debate, crowding out new voices and distracting from alternative viewpoints. These exploit social media algorithms to amplify false information ad target specific groups.

Evidence continues to emerge which suggests this problem is not going away. For example, a study from the University of Oxford's Computational Propaganda Project found evidence of organised social media manipulation campaigns in 48 countries in 2018[1].

[1] https://comprop.oii.ox.ac.uk/research/cybertroops2018/