

## Open Net Association's input to UN SR FOE on disinformation

February 15, 2021

1. What do you believe are the key challenges raised by disinformation? What measures would you recommend to address them?

Disinformation has misled the public into dangerous political discourses, inciting among others racial violence, refusal to cooperate with COVID-19 pandemic, blind loyalty to authoritarianism, etc. The harms of disinformation can be mitigated by greater transparency and sharing of accurate information.

The new trend of disinformation coincides with the rise of political leaders vested in dividing people into the good and the bad, and disinformation campaigns by people supporting those leaders usually involve construction of "enemies" and spreading false information discrediting them. Therefore, political leadership demonstrating a commitment to representation of all people and racial justice is crucial to combatting the new disinformation phenomena.

As to specific government measures, see below 2a-2c.

2. a. What legislative, administrative, policy, regulatory or other measures have Governments taken to counter disinformation online and offline?

Governments around the world have resorted to punitive measures against disinformation such as "false news" crimes or criminal defamation. However, such punitive measures not only hamper counter-speech aiming to abate the harms of disinformation but also produce distrust of the society in general, generating receptivity to conspiracy theories and similar disinformation trends (See examples below).

Most recently, Myanmar's military government that came into power through coup d'etat on February 1 proposed a draft cybersecurity law that punishes both speakers and intermediaries up to 3 years of imprisonment for disinformation (Section 65 for individuals; Sections 29 and 61 for intermediaries). It is likely that this law will be used to suppress voices and people calling for democracy, and what is worse, make intermediaries exercise comprehensive censorship to that effect. It is urgent that the Special Rapporteur speak on this instance of anti-disinformation measure.

2.b. What has been the impact of such measures on i) disinformation; ii) freedom of opinion and expression; and iii) other human rights?

"False news" crimes and criminal defamation have been abused by authoritarian governments to suppress and persecute dissident/rival political forces and opinions which are often truthful. For instance, in Korea, an internet pundit whose by and large accurate critique of the government's exchange rate policies was criminally prosecuted and jailed for minor inaccuracies in his blog posts in 2009. Also, TV news producers were criminally prosecuted for making a show questioning the government's beef importation policies again with minor inaccuracies in 2009. The theory of guilt was that such news defamed the government officials responsible for those policies. The

international human rights literature is full of similar abuses spanning several decades from the 1960s around the world, and several international human rights bodies have denounced “false news” crimes and criminal defamation laws.

Especially, an election-related “false news” crime was used back in 2011 in Korea to imprison an opposition politician for allegedly spreading false information about a presidential candidate’s involvement in a stock price manipulation scheme but in 2019 the allegation was confirmed to be true in a trial that found the now former president guilty of the scheme. This is a case in point showing that disinformation prosecution actually begets and propagates disinformation.

Additionally, such overzealous prosecution creates a siege mentality and a sense of victimhood on the political groups thus persecuted, who then take on their own versions of conspiracy theories and disinformation tactics to defend their political gains and attack the opponents, causing failure of deliberative democracy where opposing political camps stop discussion and engage in shouting matches calling each other “fake news”.

### 2.c. What measures have been taken to address any negative impact on human rights?

In both cases, the role of the judiciary was crucial in checking the pro-incumbent prosecutions of “false news” crimes and criminal defamation. In the above internet pundit case, the criminal court acquitted the pundit on the ground of his reasonable belief in the truth of his speech and the constitutional court found the “false news” law unconstitutional for penalizing speech for its truth value or lack thereof it only on account of the vague notion of “interfering with public interest” (2010). In the TV news producers case, the courts also acquitted the TV show producers on the ground of their reasonable beliefs in the truth of the content broadcast (2011).

### 3.a. What policies, procedures or other measure have digital tech companies introduced to address the problem of disinformation?

Facebook and Twitter de-platformed the accounts and/or took down postings inciting violence with false information (e.g., Trump’s account for spreading false information that the election was rigged). Twitter’s Trust and Safety Council and Facebook’s Oversight Board are also the attempts to address disinformation conducive to incitation of violence.

Likewise in Korea, Naver and Kakao created a cross-industry self-regulatory mechanism called KISO (Korea Internet Self-Governance Organization) who deliberates on contents flagged and escalated by the member companies.

### 3.b. To what extent do you find these measures to be fair, transparent and effective in protecting human rights, particularly freedom of opinion and expression?

When the platforms have an opportunity to invest time and resources into deliberating upon and justifying takedown/deplatform decisions, they seem to make fair, transparent and effective decisions whether to take down or leave in tact postings.

KISO's decisions in Korea and Facebook and Twitter's deliberated decisions globally seem to stay within the reasonable margin of appreciation under international human rights law, balancing the users' freedom of speech against the rights of the possible victims of harmful speech and the freedom of platform operators.

What is problematic, there are too many non-deliberated takedown or leave-in decisions that are made, that do not benefit from calm balancing, in the primary stage of content moderation. What is worse, there are so many decisions that, even when objected to, do not escalate properly to a higher level of deliberation or a full disclosure of the reasoning behind the decisions.

3.c. What procedures exist to address grievances and provide remedies for users, monitor the action of the companies, and how effective are they?

Scalability seems to be the most important variable. Many primary take-down measures taken by digital tech companies are often over-inclusive and under-inclusive for the simple reason that there are too many postings wrongly flagged or too many illegal/dangerous postings not flagged, compared to the human and machine resources for reviewing the former or flagging the latter. We believe that transparency and due process is the solution to such scalability, meaning by sharing information with and engaging with the users and flaggers to properly escalate to a higher level of deliberation.

4. Please share information on measures that you believe have been especially effective to protect the right to freedom of opinion and expression while addressing disinformation on social media platforms.

Intermediary liability safe harbor allows platform operators to take down content or deplatform users without having to assume the role, risk and liability for aiding the publication of the problematic content (US's CDA 230). Also, the safe harbor in the EU E-Commerce Directive 2000 protects platforms from liability for user-created contents that platform operators are not aware of, and therefore, taking down known illegal contents does not give rise to a greater risk of liability for platforms. These principles incentivize platform operators into innovating with different processes and scopes of abating disinformation online, without being attributed to all the legal burdens of publishers.

5. Please share information on measures to address disinformation that you believe have aggravated or led to human rights violations, in particular the right to freedom of opinion and expression.

Criminalization of speech through "false news" crimes and criminal defamation has aggravated freedom of opinion and expression (See 1, 2b above).

6. Please share any suggestions or recommendation you may have for the Special Rapporteur on how to protect and promote the right to freedom of opinion and expression while addressing disinformation.

In relation to the benefit of intermediary liability safe harbors for encouraging self-regulatory anti-disinformation measures, there must be a clearer announcement on

the definition. For instance, the relationship between NetzDG, Germany's national law aiming to address online disinformation, and intermediary liability safe harbor enshrined in the 2000 E-Commerce Directive and the new Digital Services Act has not been clear. NetzDG seems to impose liability on platform operators when they are only aware of the existence of the content but not its illegality. Such legislative practice is not consistent with the principle of intermediary liability safe harbor, which is an enhanced corollary of the general axiom that accomplice liability attaches only when there is knowledge of the illegal nature of the main perpetrator's conduct. Platforms pushed toward the brinkmanship of liability by such law will avoid innovating with anti-disinformation measures. Taking down too many contents will result in erosion of trust and therefore creates a fertile soil for conspiracy theories and other disinformation.