CyberPeace
Institute

Cyber Peace Institute Contribution to *Report on Disinformation* Issued by Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression

We welcome the open call for contributions issued ahead of the annual thematic report on disinformation, due to be presented in June 2021. The CyberPeace Institute is an independent NGO based in Geneva, working to encourage accountability in cyberspace and towards the protection of human security, dignity and equality in digital ecosystems. Since the outbreak of COVID19, health-related disinformation and attacks against the healthcare sector have become a priority area for us. The contribution below draws on this work, identifying the key challenges raised by disinformation and information-enabled cyber operations and recommending ways to address them.

Key challenges

The proliferation of false, misleading, and inaccurate information amid facts fundamentally undermines human security, dignity and equity in cyberspace. Exploiting confusion and an urgency to search for accurate information, cyber criminals spread messages to plant malware, phish for credentials and ask for donations, but they also undermine democracy on a global scale, reducing trust in authoritative institutions and democratic processes. In this contribution, we identify four key challenges related to disinformation.

1. High stakes: distinguishing between accurate and false information costs lives.

The spread of disinformation in the context of a pandemic has proven to be particularly dangerous in relation to ineffective or untested remedies. This so-called COVID-19 infodemic has been facilitated by the spread of both deliberate and non-deliberate wrongful information (i.e. disinformation & misinformation) by private as well as state actors. These “informational viruses" threaten to undermine the local and global responses to the actual biological viruses by interrupting or disabling medical services, diverting staff away from their regular activities, redirecting valuable resources and weakening public trust in authoritative institutions. Such a virulent information environment has been linked to a greater death toll during the recent Ebola outbreak in West Africa[1] and may have played a role in the death of over 700 Iranians, who ingested toxic methanol thinking that it could cure COVID-19[2]. Through the spread of misinformation, state actors may be both **infringing on the rights of people to seek and receive information** and **violating their obligation to respect and protect the rights to life and health.**

Background information:

While the origins of healthcare-related disinformation campaigns are not always evident, some have been exposed as a deliberate tactic employed by state actors. In September 2020, documents of the Richard Lugar Research Center were stolen and leaked on a “foreign website” together with falsified documents[3]. The US-funded center has not only been at the forefront of Georgia’s pandemic response but has also featured prominently in a Russian disinformation campaign that paints it as a research lab for biological weapons[4]. Similarly, in January 2020, hackers compromised the content management system of a Lithuanian news site and planted a falsified story of US soldiers spreading COVID-19 in the Baltics[5]. Such cyber enabled information operations against healthcare

[1] Allgaier and Svalastog, 2020. “The Communication Aspects of the Ebola Virus Disease Outbreak in Western Africa – Do We Need to Counter One, Two, or Many Epidemics?”. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4655935/

[2] Al Jazeera, 2020. “Iran: over 700 dead after drinking alcohol to cure coronavirus”. Available at: https://www.aljazeera.com/news/2020/4/27/iran-over-700-dead-after-drinking-alcohol-to-cure-coronavirus

[3] IDFI, 2020. “Cyberattack on the Ministry of Health and Russian Trace“. Available at: https://idfi.ge:443/en/strategy_of_russian_cyber_operations

[4] Civil.ge, 2020.“Tbilisi Decries Russian Disinformation over Lugar Research Center“. Available at: https://civil.ge/archives/353924

[5] Baltic Times, 2020. “Fake report about coronavirus-infected US soldier posted on Lithuanian news website“. Available at: Fake report about coronavirus-infected US soldier posted on Lithuanian news website.

targets directly contribute to the COVID-19 Infodemic and threaten to complicate the pandemic response.

2. Disinformation creates a climate of confusion and uncertainty, leaving everyone vulnerable to cybercriminal exploits.

The global community as a whole stands to lose from actions intended to deliberately deceive the public and lower the level of trust in online information, which impacts our **rights to work and education, freedom of expression and opinion, as well as the rights to personal security and privacy.**

Background information:

A deluge of malicious cyber activity, accompanied the false and misleading information, has led to the destabilization of the information space during the pandemic. The Georgian government claims that Russian hackers stole and leaked data from the Tbilisi-based Richard Lugar Research Center as part of a broader disinformation campaign that aimed to discredit it as a research lab for biological weapons[6]. Threat actors have capitalized on the pandemic-induced fear and uncertainty in their social engineering attacks by using COVID-19 themed lures in their phishing attacks and malicious domains[7]. The geographic targeting of these campaigns has largely correlated with the spread of the virus[8] as illustrated by Emotet's COVID-19 malspam campaigns, which targeted users in Japan and Italy as the virus was spreading there in January and March 2020[9].

The selling of branded COVID-19 vaccines online has also emerged as a lucrative business, with the price of individual doses of alleged branded vaccines between 20 USD and 1,575 USD[10]. The personal information required for an online order for a vaccine has also been used in phishing and malware exploits. In December 2020, the US authorities shut down fake websites of alleged biotech companies which stole users'

[6] IDFI, 2020. "Cyberattack on the Ministry of Health and Russian Trace". Available at: https://idfi.ge:443/en/strategy_of_russian_cyber_operations

[7] Recorded Future, 2020. "Capitalizing on Coronavirus Panic, Threat Actors Target Victims Worldwide". Available at: https://www.recordedfuture.com/coronavirus-panic-exploit/

[8] Pilkey, Adam, 2020. "Coronavirus email attacks evolving as outbreak spreads". Available at: https://blog.f-secure.com/coronavirus-email-attacks-evolving-as-outbreak-spreads/

[9] Id.

[10] CodaStory, 2021. "Newsletter - infodemic, February". Available at: https://www.codastory.com/newsletters/infodemic-february-5/

information for cyberattacks[11] , whereas the EUROPOL has warned about the high-quality faking and counterfeiting of COVID-19 test documentation for travel purposes[12].

To appear legitimate, attackers have also been observed imitating organizations involved in the pandemic response on a national and international level[13]. The imitation of the digital presence of organizations can have several repercussions: first, depending on the content of the email, it can threaten the victim's health. Second, it can undermine the organization's credibility and therewith its ability to perform its role.

3. Online disinformation undermines democracy

Disinformation can single handedly affect and disrupt democratic systems, whether during elections or in times of pandemic. Key elements of destabilisation include the erosion of trust in democratic principles and processes, easily undermined by malicious cyber activity, but also greater surveillance of individual users and efforts to render certain groups more vulnerable and dismiss alternative voices.

Background information:

Foreign and domestic players eroding democratic principles: The information space has been abused not only by malicious foreign actors but also by democratic governments who cover up data, send out conflicting reports, and spread misinformation. This has made civilians more vulnerable and created additional options for external actors to attack.

Increasing surveillance: Covid-19 disinformation is used as an excuse to launch more surveillance activities against civilian populations. In many cases, such software has also been used by states to track people's movements during the pandemic[14]. Some tech firms have also tried to sell their surveillance technology for civilian purposes.

[11] Becker's Hospital Review, 2020. "Officials shut down Modern Regeneron websites that allegedly stole users' information for cyberattacks. Available at: "https://www.beckershospitalreview.com/cybersecurity/officials-shut-down-fake-moderna-regeneron-websites-that-allegedly-stole-users-info-for-cyberattacks.html

[12] EUROPOL, 2021."EUROPOL warns on the illicit sale of false negative COVID-19 test certificates". Available at: https://www.europol.europa.eu/newsroom/news/europol-warning-illicit-sale-of-false-negative-covid-19-test-certificates

[13] Kaspersky, 2020. "The year of social distancing or social engineering? Phishing goes targeted and diversifies during COVID-19 outbreak". Available at: https://www.kaspersky.com/about/press-releases/2020_the-year-of-social-distancing-or-social-engineering

[14] BBC News, 2020. "Coronavirus: Israel halts police phone tracking over privacy concerns". Available at: https://www.bbc.com/news/technology-52395886

Reducing the plurality of voices: Disinformation spreads as a decentralized threat via social media, reinforcing echo chambers and increasing engagement levels around ‘fake news’. While the moderation of content has become a standard practice, this response shows systemic limitations, such as the imprecision of automated systems for reviewing content and the heavy reliance upon human workforce. This reality accentuates the asymmetry of capacities between major players and smaller platforms, the latter being unable to hire sufficient workers to cope with the challenge at hand.

Rendering certain groups more vulnerable: Beyond its obvious cost, this reliance on manual review of content creates a new risk beyond the infodemic itself: exposing ever more staff to traumatizing and manipulative content potentially brings vicarious trauma at scale. Moreover, due to the design of social media platforms and business models supporting the ‘eyeball economy’, structural incentives exist to produce ‘fake news’ that attract more attention and have rendered users more vulnerable to disinformation.

4. Disinformation thrives on the absence of accountability.

In the absence of adequate accountability frameworks, cyber criminals continue to have a *carte blanche* for their illicit operations against vulnerable persons. The burden of proof lies primarily in the strength of the causal link between disinformation and the damage or harm that it causes to infrastructure, public institutions and individuals. **Not being able to hold those responsible accountable makes it impossible to exercise the right to an effective remedy.**

Measures to address the challenges identified

1. Condemning the targeting of the healthcare sector with disinformation tactics

As working, studying and socializing have moved online during the pandemic, the weaponization of the information landscape directly affects our well-being and our lives. The spread of information and disinformation about healthcare professionals and international health organizations, rumors and fear mongering about vaccine development and risks and information-enabled cyber attacks make it difficult to find essential and trustworthy information, leave communities vulnerable and can lead to loss of lives. This proliferation undermines patient safety, and more broadly, human security at a global scale.

Moreover, the infodemic threatens all concepts of peace, not only in cyberspace but also in the public sphere. It locks populations within echo chambers of misinformation, accelerates sectarianism and impacts some of the fundamentals of modern society — the sense of community, the confidence in science, the sharing of common values. It is dangerous at all times, but even more so in times of pandemic. It is primordial that the targeting of the healthcare sector is publicly condemned in clear terms and becomes unacceptable as a destabilization tactic against states and against individual institutions.

2. Calling on governments to refrain from orchestrating disinformation campaigns and information operations and to set in place measures to protect vulnerable groups

It is crucial to call on states to refrain from engaging in any and all disinformation campaigns or information operations. Beyond affecting fundamental rights, such activities are also long-term damaging for our democratic system, increasing violence and dividing communities. State actors also have positive obligations to protect their own citizens from disinformation campaigns that impact or disrupt their life, work or education. They must also act in order to prevent human rights violations. In the case of information attacks on a healthcare facility or misinformation campaigns in which the rights to life and health of people are in jeopardy, the state has an obligation to act to prevent harm. Setting in place adequate safeguards for vulnerable populations and key sectors likely to be destabilized by disinformation is much needed.

3. Calling on the private sector to adopt human-centric content moderation practices and bring platform architecture in line with responsible design

Any response to the challenge of content moderation of large and small social media platforms must be human-centric. For example, human-based moderation must be supported with technology and psychological support for those who are exposed on a daily basis to the worst of the Internet. In addition, any technical procedures which support moderation must be widely accessible and shared with the entire industry and civil society, which should have the capacity to audit such technology in terms of its potential impact on fundamental human rights.

Actors at the top of the information chain, such as the media and the influencers, have a key role to play in terms of fact-checking. Any action taken will be worthless if these actors serve as a disinformation haven.

4. Calling for greater accountability

The devastating impact of the infodemic on society is an avid reminder that this issue cannot be relegated to the sidelines of international debate. Acknowledging that there are new vulnerabilities and threats to online security, influencing the level of trust in the digital ecosystem, the accountability framework to be established needs to respond to this volatility and to the continuous evolution of risks. Effectively measuring the human and societal impact of disinformation and related cyber attacks is instrumental not only in helping to better understand the problem, but also in building sufficient evidence to trigger the application of domestic, international, and human rights law, and therefore leading to the stronger protection of vulnerable communities.

The CyberPeace Institute has developed a proposal for an accountability framework with the goal of mapping accountability and pushing increased responsible behaviour onto the numerous actors of the digital world. This framework has four key elements:

1. Identify **clear and common expectations** of what constitutes responsible behaviour in cyberspace
2. Establish **stakeholder commitments** to uphold expectations
3. Track stakeholder **adherence to commitments**
4. **Implement the consequences** for failure to uphold commitments and reward or incentive for upholding commitments

By using this framework to guide its actions and serving as a neutral and independent source of information regarding the activities of stakeholders active in cyberspace, the aim is to effect change and increase responsible behavior by all stakeholders in cyberspace.

Overall, we call for more action to tackle disinformation while protecting individual rights and to hold accountable those responsible, building on a human-centric approach.