International Civic Forum

Countering misinformation and disinformation

Supported by

Sweden
Sverige

Date	01 – 03 December 2020
Location	virtually

Outline

The International Civic Forum 2020 focused on ***"Countering misinformation and disinformation that target civil society organisations and the communities they serve".*** The event took place virtually in shorter blocks over a course of three days. Individual sessions were building on each other for those who could attend all three days but individual days had a specific focus to enable also participation of people who were able to join only one day. Between 30 and 45 people joined panel discussions and breakout sessions each day. We discussed risks that mis- and disinformation pose, exchanged on ways of countering them, looked at useful tools and strategies, and explored new models of addressing mis- and disinformation.

Summary of discussions

Setting the scene: Why do we need to counter mis- and disinformation?

There is a difference between misinformation, disinformation and malinformation based on the intent and what the person spreading the information wants to achieve. If the information is false but the person spreading it does not know that, we talk about *misinformation*. If the person intends to mislead people with false information, we talk about *disinformation*. When it comes to *malinformation*, the information being spread is generally true but it is released at a time that negatively impacts public conversation. Lies and false information are not a new phenomenon but the internet has accelerated the speed of spreading information, including mis-, dis- and malinformation. In the context of the current Covid-19 pandemic, lies about it spread faster than the truth catches up. It is not enough to simply counter a lie with a fact, but it is also important to understand why people would believe the lie in the first place. This brings up bigger questions of social dynamics, the role of traditional media leaving people vulnerable to disinformation, and the reasons behind people turning to disinformation.

Context is key – regardless of the quality of information available, people still have to make life-changing decisions based on information available to them. In an information poor environment, because of a lack of good information, what is a "rumour" to one person may appear as a "fact" to another. When entering a context, it is important to understand *why* it is information poor (e.g. lack of good information, barriers to access) in order to build a strategy. Not only are we facing *information poor* environments, we are also facing information environments where there is *too much information* - leading to information fatigue - and we need to understand their impacts. Rumours show us what information gaps exist - e.g. existing information barriers or not an appropriate language in which information is being provided can lead to rumours. Continuous analysis, talking to as many people as possible and considering the context are key factors for analysing rumours and assessing their risks – whether the rumour in question might be harmful, whether it could act as an early warning sign, whether it might harm an organisation's work or pose reputational risks. Analysis of rumours should include what caused the rumour and what can be learned from it, assessing the rumour as actionable information and defining a way forward. A rumour-tracking methodology published by Internews is a good starting point of assessing rumours and their impacts.

There are three main areas of concern related to the impact of disinformation on society: erosion of public trust in institutions and political actors; degradation of public debate in quality and content; and interference in democratic processes. Mis- and disinformation hold many inherent risks, including potential and actual risks posed to freedom of expression and to protection of civil society. It is necessary to work on media and digital literacy, support fact-checking community, work with people on the ground and develop more inclusive and responsive communication strategies.

Civil society organisations are a part of the solution to counter mis- and disinformation, but at the same time they have also been targets of them. With civil society organisations targeted, the question is not only about how they can best counter the attacks, but how the attacks might impact the communities they serve as resources and capacities that could have been used to support these communities, need to be redirected towards responding to mis-, dis- or malinformation.

Civil society organisations that have experienced disinformation attacks shared the following lessons and good practices:

- Provide facts and be transparent.
- Challenge (selectively) the narrative against you.
- Stand in solidarity with other civil society organisations that are targeted, provide support, share experience and strategies.
- Develop a risk management strategy if necessary.
- Say who you are and what you stand for. Do not get involved in arguing publicly what you are not but instead talk about who you are and what you do. Showcase your work and its positives.
- Have your work rooted in the communities that you serve.
- Involve your allies and ask people who know your work and believe in it to support you by speaking about your work publicly and to different stakeholders.
- There is no silver bullet to respond to disinformation. A fitting strategy needs to be developed on a case-by-case basis.

Sharing experience: What can we learn from different tools and strategies?

To help civil society organisations get better prepared for disinformation attacks, InterAction developed a Disinformation Toolkit. The toolkit offers a baseline of definitions and different practices to guide organisations on how to address disinformation attacks based on 5 WHs: *Who, what, when, where* and *why*. The toolkit further lays out advantages and disadvantages of countering an online attack. Not all disinformation attacks are worth being countered – some do more harm if responded to them. There is *no one-size-fits-all solution*. It is important to create a strong narrative for your work and a public reputation of your organisation so if an attack occurs, it does not impact the public as the public knows what the organisation is about and what it stands for. When countering a disinformation attack, support from a coalition of civil society actors is useful to not feel alone and to come up with different strategies. The toolkit helps civil society actors prepare an action plan and assess risks as well as shows how to be proactive and form partnerships. InterAction is currently updating its Disinformation Toolkit (to be published in 2021) to incorporate received feedback, include more recent examples from the US and other countries and provide more evidence-based resources and best practices.

We further looked at three examples of tools and strategies developed to address mis- and disinformation:

Fact-checking: How can we pre-bunk misinformation and use WhatsApp for fact-checking?

Africa Check applies two strategies to target communities or groups which have been exposed to false information - either deliberately or as collateral damage: a) *reactive fact-checking*: applying methodology, publishing and distributing the fact-check, but it is unlikely to reach everyone who has been misled because of the different ways in which the false information is spread. This is complemented with 2) *pro-active work (pre-bunking)*, with factsheets and guides to get ahead of the curve on topics which are likely to confuse people or where they are likely to be susceptible to being misled. Expert consultations are complemented with community involvement to make sure that the response is rooted in actual problems rather than assumptions, and true information is presented in a simple and easy-to-understand way. Participatory approach and hearing from communities what is important to them are key factors of success.

In 2019 Africa Check started *"What's Crap"* on WhatsApp. They use the voice note feature, record and edit short audio notes, and distribute them via their broadcast lists. This has been a fun and accessible way for people to engage with their resources, and Africa Check published a handbook on how everyone can do this: How to start a WhatsApp fact-checking podcast.

Disinformation detection: How can we track online disinformation on subjects that we care about?

Not every organisation can or should become a disinformation detective. But disinformation can threaten activities, objectives and individuals associated with civil society groups and their work. Organisations witnessing or targeted by disinformation therefore require a baseline understanding of the threats posed by disinformation and how to spot them while conducting their work. The disinformation starter kit The 101 of Disinformation Detection published by the Institute for Strategic Dialogue (ISD) lays out an approach that organisations can undertake to begin to track online disinformation on subjects that they care about. This includes four steps:
1) *Preparation*: Define a strategy; review and revise it; agree and distribute it.
2) *Data collection*: Build a list of relevant keywords and actors; create queries based on these lists.
3) *Spotting false information*: Produce a probability sample and analyse it; apply filters and take steps to improve precision. Take steps to improve recall; begin analysis and reporting; develop an appropriate response.
4) *Spotting false behaviour*: Select some accounts that shared disinformation; test for account automation considering profile, posts and point of view.

A document titled Developing a Civil Society Response to Online Manipulation published by the ISD further provides a vision for a pan-civil societal response to online manipulation: developing the capability to detect it; the coalitions to confront it; the strategies to prevent it, and the structures of cooperation and funding needed.

Empowering the targeted: Resilience to disinformation through critical information engagement with Learn to Discern

How to empower those who are targeted by manipulative information – citizens, youth, activists, voters, policy- and decision-makers – to navigate the information ecosystem, including digital spaces, in safe, responsible, empathy-driven ways? With its *"Learn to Discern"*, IREX has developed an approach to build sustainable healthy information engagement habits in individuals, organisations, communities, and systems.

IREX's *"Learn to Discern"* methodology focuses on: building citizen resilience to disinformation; understanding human info-processing vulnerabilities; recognising hate speech and manipulation; preventing violent extremism; building resilience to polarisation in societies; addressing the public health "infodemic", and preparing new generations for digital futures. *"Learn to Discern"* encourages critical thinking so people can identify mis- and disinformation. It increases awareness of emotional and cognitive biases and vulnerabilities. It focuses on skills (*how* to consume media, *not what* to consume) and long-lasting behaviours. It sharpens critical thinking and cognitive reflection skills in a practical way. It improves abilities to identify and resist information manipulation and encourages peer-to-peer support. To learn more about IREX's *"Learn to Discern",* see: Overview and Resources, Tips for Inoculating Yourself Against Disinformation and Media Literacy Trainer's Manual.

After exploring different tools and strategies, we concluded the session by discussing how we can evaluate and measure the effectiveness of countering mis- and disinformation, bringing perspectives of behavioural science into the picture.

We need to investigate and understand whether what we are doing (toolkits, campaigns, strategies, working with media etc.) to counteract mis- and disinformation is working, and how and why it is working. Measuring what works is not just a technical issue, it is also political and relational. There is no one-size-fits-all solution when it comes to measuring effectiveness. We need to consider questions that we ask and weights of different voices. Rule of thumb advice:

- Don't leave designing your evaluation till the end, consider it from the outset.
- Don't try to measure everything possible – focus on what is essential.
- Build a strong monitoring framework to collect evidence as the programme is running.
- Apply ongoing context analysis.
- Look at what is done in other areas using similar tactics but don't copy-paste solutions.
- Keep a systems thinking lens on.

It is important to bring humans back in and behavioural science helps us understand the effects of mis- and disinformation on people. The use of emotions is making mis- and disinformation effective as we perceive what is being said based on our emotions. There is a lack of understanding of behavioural mechanisms through which emotions become powerful. Emotions and behavioural mechanisms are context-specific and there is a danger in assuming that all people on social media are the same. Agent, message and interpreter need to be taken into account. Four behavioural aspects help describe why mis- and disinformation spread:

- Risk as feelings (if you are emotional about something, you would perceive the information based on the emotion you are feeling).
- Confirmation bias (if you hear something that already confirms what you believe, you will believe it even if it is not true).
- More exposure effects.
- Social norms (if a message fits the social norms that you live in, you would be more receptive to the message).

Taking actions: How can we move forward?

New forms of mis- and disinformation enabled by artificial intelligence such as *deepfakes* bring the whole discussion to the next level. Deepfakes are new forms of audiovisual manipulation that allow people to create realistic simulations of someone's face, voice or actions. They enable people to make it seem like someone said or did something they did not or an event happened that never occurred. They are getting easier to make, requiring fewer source images to build them, and becoming increasingly commercialised. Deepfakes make it easier to manipulate or fake real people's voices, faces and actions as well as enhance the ability to claim any video or audio is fake. They have become a critical concern for celebrities and politicians, and for many ordinary people worldwide, and raise mis- and disinformation threats as they become easier to deploy. To address deepfakes, don't panic but prepare! To learn more what the key technologies are, what threats deepfakes pose to civil society and what global civil society should prioritise and engage on, see the resources from WITNESS such as Backgrounder: Deepfakes in 2020, Deepfakes: A short typology as well as Prepare, Don't Panic: Synthetic Media and Deepfakes.

Crises are a good lens to look at tackling misinformation better as some events can bring a surge of bad information that impact citizens and the communities civil society organisations work with and for. Can we learn the lessons of previous waves of bad information, from elections to the pandemic, to make sure we are all more ready to contain the next crisis before it unfolds? What should our response be - and that of others - ahead of, during and following a misinformation crisis based on shared thinking? Full Fact is currently developing a framework to help decision-makers respond to and mitigate information crises. A framework for managing misinformation crises is needed to:

- Increase shared understanding of misinformation-related risks.
- Enable faster, proportionate, scalable and well-evidenced responses.
- Develop responses that can be scaled based on prior understanding of what is needed.
- Encourage effective collaboration with clarity on roles and responsibilities.
- Introduce methods of measuring success as a situation de-escalates.

We need to have a common view of the most frequent challenges, shared assessment of risks and agreed aims that can inform and direct responses. Find further information about the misinformation crises framework at FullFact's blog. The first version of the framework will be launched for consultation in early February 2021 and civil society organisations are invited to feed into this work.

Moving forward

In the final session of the International Civic Forum 2020, we discussed what we need to keep from existing practices on countering mis- and disinformation and what we need to change to address them better. These include:

- Increase awareness of full spectrum of mis-, dis- and malinformation. Raise media and digital literacy.
- Continue knowledge and information sharing, work together with organisations that have expertise in this field and stay close to the communities.
- Get faster to respond and prepare better for the next generation of threats. Enable better access across geographies.
- Apply an interdisciplinary approach when navigating through mis- and disinformation (including psychological approach as mis- and disinformation target emotions).
- More panic and pressure on tech companies to deal with the issue. Push for regulations and policy change in the countries.
- Support collective efforts, form inclusive coalitions and collaborate more internationally.

The topic of countering mis- and disinformation will be further dealt with within the scope of the Solidarity Action Network (SANE) which brings together international civil society organisations and their local partners to support each other when faced with undue threats and challenges to their operations or civic space restrictions more broadly. If you want to follow up on the International Civic Forum 2020 or get involved in further steps within the Solidarity Action Network, please contact Eva Gondorová, Project Manager: egondorova@icscentre.org.