



## **UNESCO Call for Proposals for a Research Report on Freedom of Expression and Addressing Disinformation on the Internet**

UNESCO launches a call for research proposals on behalf of the Broadband Commission for Sustainable Development, particularly the Working Group on Freedom of Expression and Addressing Disinformation.

The report should be the outcome of a survey of the various steps being taken around the world in response to what different actors perceive as disinformation. The framing of this project is founded on the universal right to freedom of expression, and the importance of ensuring that measures concerning disinformation – including through its creation, dissemination and reception – do not impact negatively on the essence of this right.

The research should analyse the different modalities of responses based on empirical data from all parts of the world, and be presented mainly at a level of generality which extrapolates key trends. There should also be a broad assessment concerning each modality of costs, benefits and risks concerning freedom of expression and sustainable development.

### **Terms of reference:**

#### **1. Background:**

The Broadband Commission for Sustainable Development, established jointly by ITU and UNESCO in 2010, is engaged in advancing the 2030 Agenda for Sustainable Development through Broadband connectivity and inclusion. A new Working Group has been set up by the Commission in the light of threats posed by disinformation to the contribution of broadband connectivity to the SDGs. The research described below is an activity of the Working Group. For this research, disinformation understood broadly, and as meaning the production, circulation and consumption of made-up facts that are presented as factual information, or erroneously believed to be such (see section on Definitions below).

Among relevant SDGs impacted by the use of broadband for disinformation are:

SDG 16 which calls for peaceful and inclusive societies.

Target 16.1 calls for reducing violence and indicator 16.10.2 refers to reducing conflict-related deaths. SDG 16A and 16B point to the need to combat terrorism and discrimination. Broadband should not be used for disinformation that is harnessed into advocacy to incitement to violence, hatred and discrimination.

SDG 16.10 which calls for “public access to information and fundamental freedoms”.

Freedom of expression is an incentive to uptake Broadband. Disinformation works against the spirit and purpose of access to information. Moreover, disinformation can trigger overreaching legal measures that entail Internet shutdowns and censorship and such steps risk being disproportional as well as intruding upon fundamental freedoms.

SDG 4 which calls for “quality education and lifelong learning”.

Disinformation works against the purpose of education and learning. On the other hand, digital knowledge and skills, as well as media and information literacy, are essential in order to provide Broadband users with understanding of ethics, human rights and how to defend against the automation, fabrication and dissemination of disinformation.

SDG 9.c which calls for universal and affordable Internet access.

Disinformation – like spam and viruses - can have negative impact on access by deterring the public from seeing the constructive potential of the Internet.

SDG 3 which calls for good health and wellbeing.

Abuse of internet connectivity to spread falsehoods, such as about vaccination for example, undermines the potential of Broadband to contribute to achieving this SDG.

SDG 5 which calls for gender equality.

Online disinformation about gender issues work against efforts to achieve equality.

This background of the SDGs contextualizes the interest of the Broadband Commission in disinformation and what is being done in response to it.

## **2. Definitions**

### Freedom of expression:

Freedom of expression is a fundamental human right, recognized as such in the Universal Declaration of Human Rights proclaimed by the United Nations General Assembly in Paris on 10 December 1948 (General Assembly resolution 217 A).

Article 19 of the Declaration states: Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.

Thus, this fundamental freedom is well defined in international human rights law. It is the starting point for a Broadband Commission assessment of that form of expression that is identified as disinformation.

It is worth noting that like all rights, the right is the norm; and any restrictions should be exceptions. Furthermore, for any exceptions to the free flow of expression to be legitimate, it is required under international law that these must be in law, necessary and for legitimate purpose such as the rights of others.

How disinformation relates to freedom of expression and its legal exceptions is complex. It is not evident that the full range of problems of disinformation can be addressed through laws that meet the international criteria for legitimately restricting content. What is clear, however, is that besides for legislative and regulatory strategies, there are also many other approaches being undertaken around the world to deal with disinformation. Considering the totality of responses is important for framing the matter of disinformation within an holistic approach that is grounded in the international standards for freedom of expression.

### Disinformation:

This term is not defined in international law. Various groups such as the European Union and the Council of Europe have interpreted it to mean fabricated factual content deliberately disseminated with an intention to obscure the truth, covertly influence public opinion or result in specific behaviors. Complexities relate to the challenges of establishing intention, and to different perspectives on the relationship between context and truth, and on how facts are verified and presented in context.

In summary, the research being commissioned is not aiming to produce a definition of disinformation. Instead, it operates with the assumption that the term exists as a pole on a spectrum with its opposite being “information”. It assumes that “disinformation” functions in counter-position to “information”, through its purveyance of “facts” that are inherently false and cannot stand up to credible verification. The research also will examine trends that are moving away from deliberate falsehoods into the amplification and manipulation of information in a way that distorts the information ecosystem by sidelining scientific information as well as journalistic information that is produced professionally and focused on public interest criteria. In this perspective, disinformation is wider than formats that exploit the style and conventions of news. It is also distinct from opinion, as well as being different from hate-speech if the latter is broadly understood as advocacy for incitement to violence, hostility or discrimination. At the same time, it is evident that although each of these kinds of expression can exist without the other, disinformation is often combined with opinion and incitement to constitute a particularly toxic package.

Accordingly, the approach is to have a broad working conceptualization of disinformation which can apply across countries. In this regard, a straightforward semantic analysis can help towards distinguishing meanings of the term and its differences to other words. Such an analysis may note that in popular usage, disinformation is sometimes used as a synonym for “fake news” as well as “hate speech” and “misinformation”. Considering these one by one, the following points can be made:

- Some disinformation may fraudulently be presented as if it were verified and professional news information, and this may also be reinforced when actors masquerade as established media entities. However, disinformation can also exist in other informational formats (eg. it can be inserted into Wikipedia entries). This means that disinformation therefore covers more than “news” formats. Meanwhile, the term “fake news” itself is losing a common meaning because it has also been used in very many ways and to refer to completely different issues.

- There are many instances where unverified information is harnessed into advocacy for incitement to violence, hostility and discrimination. However, it can be noted that “hate speech” also often involves non-informational formats (eg. strong opinion, “memes”, entertainment formats). In this sense, disinformation and “hate speech” are therefore not synonyms. Disinformation can exist independently of hate speech (for example, in false messages about vaccination risks). Hate speech does not necessarily rely on disinformation. There may be different responses by concerned actors in terms of combating each of these qualities of expression, even when the two are combined.
- The term “misinformation” is sometimes used as a synonym for “disinformation”. In other cases, however, the term is used as a verb to describe the effects of “disinformation” – i.e. that people become misinformed. In yet a further use, “misinformation” refers to the dissemination of false information without knowing that it is false, and therefore without malicious intent. Responses may vary depending on the distinction made concerning intention, which is significant for research. At the same time, the overarching commonality is that whether done deliberately or inadvertently, is that made-up information is circulated in the communications ecosystem, with potential or real harm to human rights and sustainable development.

Considering the semantic analysis above, it is recommended that the term disinformation should not be conflated with “fake news” or “hate speech”, or “misinformation”, although it can appear in combinations with these. This scoping of the term provides guidance for this research project to be able to identify the key activities being deployed to counter the phenomenon. In many cases, the interventions may be self-described as steps against “fake news”, “hate speech” or “misinformation”, but should be assessed on the basis of intersection with the perspective on disinformation formulated in these terms of reference. Where actors explicitly deem their activities to be tackling disinformation, the meaning they attribute to the term should also be unpacked and assessed in terms of the semantic distinctions made above.

This research recognizes that there may generally be a relationship between how the diverse actors describe what they are doing, and how they seek to intervene at moments of the production, transmission, and consumption of disinformation. It is therefore of value to assess how those actors who intervene against disinformation define their rationales and objectives, and how their approaches shape their types of responses as well as how they evaluate the results of their interventions.

Thus, this research should pay attention to the differences in how actors approach the definitional question. However, the primary emphasis of the research is to go further and to put most focus upon the responses which the respective actors either propose or are already implementing as their perceived solutions to the issues at stake. These responses are categorized in a typology of modalities presented below.

### **3. Objectives:**

The objective of the proposed research is to produce knowledge and understanding about efforts to counter the abuse of broadband for the dissemination of what diverse actors may define as disinformation.

In particular, this research foresees the collection of information, and its compilation into a report containing useful information for governments, courts, regulators, educators, companies, academia, media, civil society organizations and others.

To this end, the research will include:

- A “gap analysis” through reviewing studies, surveys, and other research about this phenomenon and aspects of the communication system in which misinformation and hate speech are disseminated. This literature review will help refine the niche and value-add of this study.
- Research based on experiences around the world into the typology of modalities described in section 6 below, and on this basis providing descriptions as well as an assessment in each case, at a general level.
- Proposals for the way ahead, such as:
  - Identifying options for creative responses that can effectively reduce the obstacles raised by disinformation in relation to Broadband’s SDG potential;
  - Signaling a research agenda for academia and the research community;
  - Ideas for developing toolkits to support decision-making on these issues that can empower policy makers in government, civil society and corporate sector.

Overall, the research aims to provide insight that can stimulate creativity and innovation in tackling problems in broadband use while reinforcing freedom of expression and sustainable development, as well as help set agendas for future research, and provide useful resources for decision-makers.

### **4. Research fields – the modalities of responses to disinformation:**

As indicated, the research will prioritise information gathering about the responses to disinformation. The typology of categories for organizing this research, as set out below, are not through the prism of particular actors, nor the problems being responded to. Instead, the research fields that follow are based on modalities, i.e. the kinds of activities that are being undertaken.

It is recognized that many actors are engaged in more than one modality, and that a number of interventions are undertaken by partnerships of diverse actors. Where there are partnerships these should be signalled in the research. It can also be noted that some modalities constitute a response to a range of problems of “harmful content” in which disinformation is identified as just one of the components, or is not directly labelled as “disinformation”. Lastly, there may be mixtures of modalities, although a predominant

character is likely to be found in specific interventions to be identified empirically. By putting focus on the modalities, the research should be able to surmount these complexities.

The ambition of the research is not to map each and every instance in any given modality category. Rather it is to consider trends that can be extrapolated from many empirical instances, and which reflect a geographically-representative set of cases. Similarly, the assessment of each modality – which should cover costs, benefits and risks concerning freedom of expression and sustainable development – should be generic as opposed to singling out specific actors or interventions.

The proposed modalities to be considered in the research (with purely indicative examples) are:

1. Monitoring responses (*Eg. Italy reporting portal; Poynter Institute*)
2. Legislative / regulatory responses, including constraints on connectivity as well as electoral regulation (*eg. Germany, Malaysia, Singapore*)
3. Countering disinformation: (*eg. “EU vs disinfo” website, Pakistan government’s “fake news buster” on Twitter; how anti-hate groups are dealing with disinformation*)
4. Economic responses: (*eg. Companies disincentivising the production of disinformation for financial gain; transparency of advertising*).
5. Curatorial responses (*eg. company terms of service, codes of conduct, appeals mechanisms, limits on sharing, enlistment of users, crowd curation, recognition of certain actors for monitoring and priority complaints*)
6. Technical / algorithmic responses to disinformation (*eg. hash databases, automated ranking, upload filters*)
7. Ethical responses (*e.g. UNESCO campaign to advise people to ‘think before sharing’*)
8. Educational responses (*eg. Finnish public education system, Media and Information Literacy initiatives, the BBC’s ireporter educational game, events like misinfocon, cautionary signals featured by Facebook on certain posts*).
9. Fact-checking responses (*incl partnerships, Mexico - verificado, France - crosscheck, Brazil – comprova, South Africa - Real411*)
10. Journalistic responses (*eg. Strengthening professionalism to avoid being hoaxed; conducting investigations into orchestrated campaigns; creation of dedicated beats; accreditation initiatives such as Newsguard and the Journalism Trust Initiative; UNESCO’s handbook for training journalists – “Journalism, ~~fake news~~ and disinformation” )*

The research should be produced over a maximum of six months, with a five to 10 page draft outline and draft Table of Contents due in mid-September and an interim report due mid-way. The outcome should be a publishable report in mother-tongue level English, of approximately 100 pages, plus references, and including a five-page Executive Summary for translation. Oversight and final editing will be by a working group of Broadband Commissioners as well as external experts in the field. The final document will be published under a Creative Commons license (CC BY-NC-SA 3.0).

Proposals should include comments on the Terms of Reference, detailed description of the research methodology, description of the proposed team, including updated CV, deliverables, timeline requested funding, and what the researchers can offer in terms of promotion of the final study. To ensure that experience is covered internationally, consortia of researchers from different countries are encouraged.

UNESCO, therefore, invites interested researchers and organizations to submit their proposals, according to these Terms of Reference, by email to before **noon (CET), 5 August 2019** to [J.Hironaka@unesco.org](mailto:J.Hironaka@unesco.org).