



World Health  
Organization

European Region

# Toolkit for tackling misinformation on noncommunicable diseases

Forum for tackling  
misinformation  
on health and NCDs



# Toolkit for tackling misinformation on noncommunicable diseases

Forum for tackling misinformation  
on health and NCDs



---

European Region

## Abstract

Noncommunicable disease related, and health misinformation is a growing concern as more and more individuals obtain their health information from digital venues such as search engines or social media platforms. While increased access to information on health issues can be seen as generally positive, the spread of inaccurate medical information is of course problematic. It can lead to harmful lifestyle or dietary choices, self-medication, the abandonment of medical treatment and incorrect diagnoses.

As such, three meetings were hosted to discuss the topic with representatives from Member States, the media and social media sectors, and civil society. The outcomes of these meetings are reflected in this Toolkit.

This Toolkit was drafted following these meetings, and includes the concerns, challenges and conclusions shared during those conversations by all discussion partners. It is the product of an intense iterative process, of arguments between competing views and interests, and of the constant upgrades in available knowledge.

It reflects, to the extent possible, the developments that occurred after the meetings, but it should be read with the knowledge that it does not presume to contain everything there is to know about this topic.

**Document number:** WHO/EURO:2022-6260-46025-66542

### © World Health Organization 2022

Some rights reserved. This work is available under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO licence (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo>).

Under the terms of this licence, you may copy, redistribute and adapt the work for non-commercial purposes, provided the work is appropriately cited, as indicated below. In any use of this work, there should be no suggestion that WHO endorses any specific organization, products or services. The use of the WHO logo is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons licence. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: "This translation was not created by the World Health Organization (WHO). WHO is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition: Toolkit for tackling misinformation on noncommunicable disease: forum for tackling misinformation on health and NCDs. Copenhagen: WHO Regional Office for Europe; 2022."

Any mediation relating to disputes arising under the licence shall be conducted in accordance with the mediation rules of the World Intellectual Property Organization.

**Suggested citation.** Toolkit for tackling misinformation on noncommunicable disease: forum for tackling misinformation on health and NCDs. Copenhagen: WHO Regional Office for Europe; 2022. Licence: CC BY-NC-SA 3.0 IGO.

**Cataloguing-in-Publication (CIP) data.** CIP data are available at <http://apps.who.int/iris>.

**Sales, rights and licensing.** To purchase WHO publications, see <http://apps.who.int/bookorders>. To submit requests for commercial use and queries on rights and licensing, see <http://www.who.int/about/licensing>.

**Third-party materials.** If you wish to reuse material from this work that is attributed to a third party, such as tables, figures or images, it is your responsibility to determine whether permission is needed for that reuse and to obtain permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

**General disclaimers.** The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by WHO in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by WHO to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall WHO be liable for damages arising from its use.

# Contents

<b>Acknowledgements</b> .....	<b>iv</b>
<b>Abbreviations</b> .....	<b>v</b>
<b>Foreword</b> .....	<b>vi</b>
<b>Executive summary</b> .....	<b>vii</b>
<b>Introduction</b> .....	<b>1</b>
The rise of disinformation and misinformation .....	1
Definitions.....	4
<b>Understanding the problem: disinformation and misinformation in health</b> .....	<b>6</b>
Sources of misinformation .....	9
Main challenges .....	13
A multilayered problem .....	18
<b>Consulting stakeholders: outcomes</b> .....	<b>20</b>
Member States.....	20
Industry .....	28
Civil society .....	40
<b>The way forward: a triple entente against health disinformation and misinformation</b> .....	<b>59</b>
Building on the COVID-19 experience: a triple entente for the future .....	62
<b>Conclusions and lessons learned</b> .....	<b>65</b>
The role of Member States in the triple entente.....	65
The role of industry in the triple entente .....	71
The role of civil society in the triple entente.....	74
<b>Afterword</b> .....	<b>77</b>
<b>References</b> .....	<b>78</b>
<b>Annex</b> .....	<b>93</b>

## Acknowledgements

The technical concept of this report was developed by Kremlin Wickramasinghe and Carina Ferreira-Borges (WHO European Office for Prevention and Control of Noncommunicable Diseases (NCD Office)), and João Breda (WHO Office on Quality of Care and Patient Safety).

The WHO NCD Office would like to thank the authors of the report: Francisco Goiana-da-Silva, João Marecos and Francisco de Abreu Duarte (NCD Office) for drafting this report and for the organization of the consultations.

The WHO NCD Office would like to acknowledge all the Member States, industry and civil society representatives who participated in the three consultation meetings of the Forum for Tackling Misinformation on Health and NCDs and shared valuable input on the final toolkit. Their participation was fundamental to understand the problem and map potential solutions. The WHO NCD Office also thanks all participants that provided inputs and reviewed the final product. The NCD Office expresses its gratitude to Nino Berdzuli (Division of Country Health Programmes, WHO Regional Office for Europe) for her input and support in the development of the technical concept.

The WHO NCD Office also acknowledges and appreciates the work of João Marecos, Francisco de Abreu Duarte and Francisco Goiana-da-Silva (NCD Office) throughout the production and review; and Kathrin Hetz and Olga Zhiteneva (NCD Office) for finalizing the toolkit.

Thanks for the special contributions to this report go to Belinha De Abreu (International Media Literacy Research Symposium), Al Baker and Joe Ondrak (Logically.ai), Ruth Delbaere (Senior Legal Officer at Avaaz Foundation), Megan Marrelli (Digital Health Lab at Meedan), Sophie Randall (Patient Information Forum), Katya Vogt (Global Lead for Media and Information Literacy Initiatives at the International Research & Exchanges Board), Mathias Vermeulen (AWO Agency) and Miguel Telo de Arriaga (Directorate-General of Health).

## Abbreviations

<b>AI</b>	Artificial intelligence
<b>BBC</b>	British Broadcasting Corporation
<b>CDC</b>	Centers for Disease Control and Prevention
<b>DSA</b>	Digital Service Act
<b>EC</b>	European Commission
<b>EDMO</b>	European Digital Media Observatory
<b>EP</b>	European Parliament
<b>EPW</b>	European program of work, 2020–2025
<b>EPI-WIN</b>	WHO Information Network for Epidemics (EPI-WIN)
<b>EU</b>	European Union
<b>G7</b>	Group of Seven
<b>IFCN</b>	International Fact-Checking Network
<b>IREX</b>	International Research and Exchanges Board
<b>L2D</b>	Learn to Discern
<b>NATO</b>	North Atlantic Treaty Organization
<b>NCDs</b>	Noncommunicable diseases
<b>NGO</b>	Nongovernmental organization
<b>NHS</b>	National Health Service
<b>NYT</b>	New York Times
<b>PIF</b>	Patient Information Forum
<b>PIF TICK</b>	Patient Information Forum Trusted Information Creator
<b>RCCE</b>	Risk Communication and Community Engagement
<b>UNICEF</b>	United Nations International Children’s Emergency Fund
<b>WHO</b>	World Health Organization

## Foreword

Our interconnected, online world offers a wealth of information that is truly useful, including content on health and wellbeing. Indeed, we at WHO have never been able to reach as wide a global audience as we can via our websites and social media platforms which collectively have hundreds of millions of followers. However, the flip side is a plethora of misinformation – a veritable ‘infodemic’ – that is disseminated from multiple sources, unverified and unchecked. This is not new, of course; we have seen this ever since the advent of the Internet, but it took the COVID-19 pandemic to truly make it clear how pervasive and corrosive misinformation can be, damaging and endangering individual and collective health and well-being.

Yet, misinformation is not confined to outbreaks of infectious diseases; it jeopardizes our ability to tackle non-communicable diseases (NCDs) as well, from cardiovascular disease to cancer, obesity to diabetes, making it that much harder for legitimate, evidence-based health information to cut through the clutter, reach intended audiences and, most importantly, be trusted, absorbed and acted upon.

The ‘infodemic’ of our day and age insidiously undermines public trust in a range of key sources, including health authorities and other entities whose communications are intended to strengthen individual and community health and safety via measures grounded in the latest advances in science and medicine.

As the promotion of health and disease prevention is one of the main goals of WHO, we take the challenge of tackling health misinformation seriously. What we’ve learnt about addressing the ‘infodemic’ amid COVID-19 can now be used to tackle the mounting challenges regarding NCD-related misinformation.

The need to do so is urgent. As of this writing, NCDs account for almost 90% of all deaths in the WHO European Region covering 53 countries across Europe and Central Asia. These deaths are largely preventable; we know what needs to be done to promote and adopt healthier lifestyles and reduce morbidity and mortality stemming from NCDs. We need to communicate this widely, including and especially via online platforms. But it is a truly formidable challenge; misinformation about NCDs is rife, but we cannot throw up our hands and surrender. The stakes are too high.

The multi-layered phenomenon of health-related misinformation requires a coalition of stakeholders to strategize on how best to address this crisis – and it is a crisis. WHO/Europe, via our WHO European Office for the Prevention and Control of Noncommunicable Diseases, is pleased to have produced this timely toolkit, in the hope that we can bring together and dialogue with a range of constituencies, including WHO Member States, health sector entities, online and multimedia platforms, and civil society advocates, to raise awareness of our common challenge and offer approaches and solutions to addressing it.

Only together can we learn from each other and jointly work towards an online, interconnected ecosystem that genuinely promotes health and wellbeing for all.

*Dr Hans Henri P. Kluge*

*Regional Director, WHO Regional Office for Europe*

## Executive summary

In 2019 the WHO European Office for the Prevention and Control of Noncommunicable Diseases (the NCD Office) created a Forum for tackling misinformation on health and noncommunicable diseases (NCDs).

Misinformation is a growing concern as people increasingly obtain their information from digital venues such as search engines or social media platforms.

According to a 2018 survey, half of all European Union (EU) residents sought health information online in 2017, a figure that has almost doubled since 2008. This points towards a growing trend whereby the young and increasingly digitally literate seek diagnoses, help and advice through Internet searches and self-help tools, and also rely on Internet articles as reliable sources of health information. While there are many positives from increasing access to information, this can lead to often worrisome results. Despite the constant warnings from different domestic and international public authorities, health misinformation continues to rise.

The consequences for individual behaviours, especially regarding NCD risk factors, remain underresearched. It is crucial to assess what challenges health misinformation might pose to the well-being of digital users and to suggest new regulatory and policy pathways that can tackle the problem and ensure the well-being of citizens.

A key aspect to tackle the rise of health misinformation is to make sure that stakeholders have an aligned, common strategy, instead of procuring their own solutions.

We believe it is crucial to ensure that every player, ranging from social media platforms to governments and civil society, is committed to adopting a joint regulatory effort to fight the spread of infodemics: a true "triple entente" composed of Member States, industry and civil society, working together and developing novel ways to tackle misinformation and to make sure that health information on key topics such as nutrition, addiction, physical exercise, treatments and prevention of diseases is trustworthy, fact based and scientifically supported.





It was clear, prior to the COVID-19 pandemic, that misinformation can influence people's decisions on health. However, although NCDs are an area where behaviour and its determinants play a crucial role, the impact of misinformation related to NCDs has room for further discussion. The goal of this Toolkit is to map the current landscape of health misinformation – the problems, the actors and the potential solutions – with NCDs at the centre: how can the ongoing initiatives, many born out of the COVID-19 infodemic, be expanded further to tackle misinformation on NCDs and their risk factors. Existing innovative practices can serve as a model to contribute to digital literacy around NCDs and help citizens to critically analyse health information online.

The goal of the Forum was to bring together a broad range of stakeholders in a series of meetings to gather the necessary expertise to tackle health disinformation and help build a toolkit of policy initiatives for the future.

As such, three meetings were hosted to discuss the topic with representatives from Member States, the media and social media sectors, and civil society. The outcomes of these meetings are contained in this Toolkit, organized as follows.

- In the first part, Understanding the problem, the Toolkit provides an overview of the health misinformation landscape, particularly in relation to NCDs, and expands on the roles of gatekeepers and sources before describing the problem as multilayered and requiring comprehensive and coordinated solutions.
- In the second part, Consulting stakeholders, the Toolkit expands on the roles and initiatives currently taken by governments and international organizations, traditional media and social media platforms, nongovernmental organizations (NGOs) and experts, to provide an extensive understanding of the angles at which the problem of health misinformation is being approached.
- In the third and final part, The way forward, the Toolkit describes how different stakeholders have collaborated during COVID-19 to tackle the infodemic, and how that spirit and approach can and should be taken forward to other types of health misinformation.

The conclusion is clear: the COVID-19 pandemic represents a great opportunity to test how different stakeholders can come together to tackle health misinformation. The room to grow these partnerships exists and the role of this Toolkit is to highlight those opportunities in the hope of inspiring others to adopt some of the approaches presented herein.

Academic research on media literacy, artificial intelligence (AI), source credibility schemes and health communication, all in connection with the fight against misinformation, must be supported as they are fundamental to tackle the challenges that lie ahead.

The debate around robust policy and regulatory changes around NCD misinformation can contribute to safer physical and digital spaces. This is in line with the core priorities of the WHO European Programme of Work 2020–2025, promoting health and well-being for all.

This Toolkit was drafted following the three meetings, and includes the concerns, challenges and conclusions shared during those conversations by the civil society, industry and government entities. It is the product of an intense iterative process, of arguments between competing views and interests, and of the constant upgrades in available knowledge.

It reflects, to the extent possible, the developments that occurred after the meetings, but it should be read with the knowledge that it does not presume to contain everything there is to know about this topic.



# Introduction

## The rise of disinformation and misinformation

As more and more individuals obtain their health information from digital venues such as search engines or social media platforms, incorrect or deliberately incorrect information has become a pressing concern. While increased access to information on health issues can be seen as generally positive, the spread of inaccurate medical information and misinformation – or more acutely, of disinformation – is problematic, as inaccurate information can lead to consequences such as harmful lifestyle or dietary choices, self-medication, the abandonment of medical treatment or incorrect diagnoses. It is important to distinguish misinformation from disinformation. While the former might simply be described as the spread of false or inaccurate information, disinformation refers to the spreading of false information deliberately (and often covertly) in order to influence public opinion or to obscure the truth (Box 1).

### Box 1. Definitions

**Disinformation** is information that is created and shared with the explicit purpose to cause harm (1). WHO has also used the definition from Merriam–Webster: "the proliferation of false information deliberately and often covertly in order to influence public opinion or to obscure the truth" (2).

**Misinformation** is information that is inadvertently false and is shared without intent to cause harm. Considering the difficulty in distinguishing between intentional and unintentional purposes, the term misinformation is often used to mean any false information, regardless of intent to cause harm.

**Fake news** comprises false information transmitted in the form of "news", often by sources attempting to pass-off as online newspapers. The term has become highly politicized, most recently and notably being used to refute statements that the recipient does not like or agree with.

**Conspiracy theories** are explanations of significant events as secret plots concocted by powerful and malevolent institutions, groups, and/or people.

The impact of misinformation on health is apparent today; however, its relation to NCD risk factors remains underexplored.

According to a 2018 survey (3), half of all EU residents sought health information online in 2017, a figure that has almost doubled since 2008. This points towards a growing trend whereby the young and increasingly digitally literate seek diagnoses, help and advice through Internet searches and self-help tools, and also rely on Internet articles as reliable sources of health information. Although this may contribute to increased access to credible information and to tools that empower the patient and build up health literacy, this can also lead to worrisome results. For example, WebMD (a website providing health information) has in the past shown several testimonials that have argued for the

healing effects of apricot seeds in the treatment of cancer. However, the general medical community has primarily advised against this use and instead pointed out its potential poisonous effects (4). Likewise, numerous accounts of a potential connection between autism and vaccination (namely measles, mumps and rubella) can still be found online despite having been proved false by numerous studies (5). This has led the NGO Avaaz to question "Is fake news making us sick?" (6) by analysing how misinformation may be reducing vaccination rates in Brazil.

The experience obtained throughout the COVID-19 crisis can be valuable in addressing other health challenges, namely the spread of false information concerning risk factors connected to NCDs. Initiatives such as those mentioned above can serve as a model to contribute to digital literacy around NCDs and help citizens critically analyse online health information in that field.

In 2019 WHO issued a statement on the Role of Social Media Platforms in Health Information (7), qualifying misinformation about vaccines "as contagious and dangerous as the diseases it helps to spread". Likewise, in February 2020, WHO warned that the COVID-19 pandemic had been followed by an equally dangerous "infodemic" – "an overabundance of information, some accurate and some not – that makes it hard for people to find trustworthy sources and reliable guidance when they need it" (8). This infodemic has had severe consequences for human health and is part of a bigger trend of health disinformation.

Concerning NCDs in particular, the harmful impact of such misinformation and disinformation practices has not been adequately addressed by national and international scientific authorities yet, even though it has been an area where behaviour (and external influences, such as false information) plays such a determinant role.

This calls for more research to be conducted, including debating robust policy and regulatory changes and a substantial investment in e-health literacy among the general population.

With this in mind, the WHO NCD Office created the Forum for Tackling Misinformation on Health and NCDs (9) to bring together a broad range of stakeholders in a series of meetings to gather the necessary expertise to tackle health misinformation and help to build a toolkit of policy initiatives for the future. This multistakeholder approach looks at three different levels of governance: (i) governments of the Member States; (ii) industry, comprising both social media and traditional media outlets; and (iii) civil society at large, encompassing works by nongovernmental organizations, the scientific community and academia. Three reports were produced, one for each meeting, and are annexed to this Toolkit.

Three meetings were organized and held online due to restrictions caused by the pandemic.

- On 9 December 2020, the first meeting took place, with speakers and attendees from civil society, including academia, NGOs and start-ups.
- On 5 February 2021, the second meeting took place, with speakers and attendees from industry, including traditional media and social media platforms.
- On 1 September 2021, the third meeting took place, with representatives from Member States.

The goals of the forum were to get participants to:

- identify barriers, challenges and possible ways to fight health misinformation at the three levels;
- discuss ongoing initiatives in the field of health misinformation, exchanging best practices and relevant experience;
- debate the role of NGOs, academia, social and traditional media and public authorities, including public health bodies, in tackling health misinformation;
- discuss best practices against health disinformation among public administrations;
- assess the role of civil society, industry and local and regional administrations in tackling the spread of NCD-related disinformation;
- discuss digital literacy initiatives as a means to reduce online NCD-related disinformation; and
- showcase success stories and innovative practices among the different participants.

Those meetings were opportunities to exchange best practices, challenges and hopes, and to learn from others.

The majority of the participants accepted the diagnosis: there is a widespread and much-discussed overabundance of information online, coming from everywhere and covering all topics, opinions and sensibilities. Misinformation, disinformation and even a barrage of accurate health information can erect barriers to accessing timely, credible and valuable information, translating into barriers to accessing timely, safe and effective care.

As a result of these meetings, the NCD Office has prepared this Toolkit for policy-makers interested in pushing forward initiatives against health misinformation.

The Toolkit has three main aims:

- to map the health misinformation problem, framing it within the larger misinformation issue, and addressing stakeholders, gatekeepers, sources and main challenges. The goal is to provide policy-makers with an overview of the problem, the landscape and the main actors;
- to showcase multiple approaches and initiatives implemented by Member States, industry and civil society. The goal is to understand the breadth of solutions being pursued, their targets and challenges, and how the problem of health misinformation lacks a silver bullet but has many potential pathways to pursue; and
- to advance proposals for a way forward, based on a triple entente between governments, industry and civil society, anchored in the idea that only a multilayered, synchronized and complementary approach will ultimately benefit everyone's end goal of tackling the spread of health misinformation.

## Definitions

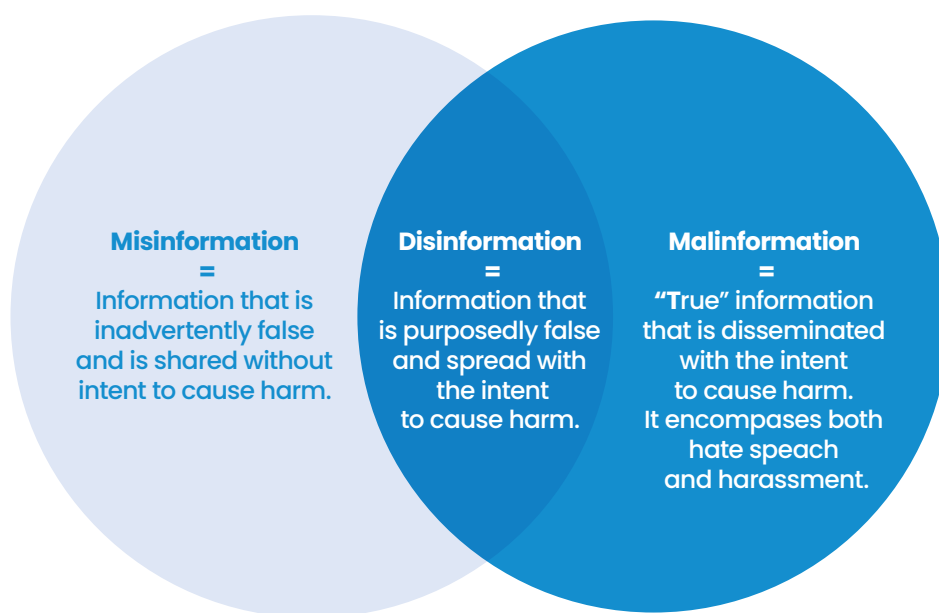
The phenomenon of false information spreading is one with many ambiguities and these start with the words used to identify the many moving parts that are encompassed by it.

It is worth noting the distinction made by the Council of Europe regarding the different modalities of **information disorders**. In their report *Information disorder: towards an interdisciplinary framework for research and policy-making*, the organization distinguished three types of information disorders: **misinformation**, **disinformation** and **malinformation** (10).

The difference between the three terms lies mainly in the intent to cause harm.

While both misinformation and disinformation refer to false or misleading information, the former is not created with the intention to cause harm, whereas the latter is. Malinformation, by comparison, concerns true information that is disseminated with the intent to cause harm. It encompasses both hate speech and harassment (Fig. 1).

**Fig. 1.** The distinction between misinformation, disinformation and malinformation



The European Union's Democracy Action Plan distinguishes between four different types (11):

- misinformation, largely with the same scope;
- disinformation idem;
- information influence operation, which pertains to "coordinated efforts by either domestic or foreign actors to influence a target audience using a range of deceptive means", that may include the suppression of independent information sources in combination with disinformation; and
- foreign interference in the information space, defined as being frequently carried out "as part of a broader hybrid operation", that can be understood as "coercive and deceptive efforts to disrupt the free formation and expression of individuals' political will by a foreign state actor or its agents".



These definitions align with the WHO distinction between misinformation – as false information that is not created with the intention to harm others – and disinformation; false information created with the intention of profiting from it or causing harm. This distinction is obviously challenging as it is often impossible to properly assess intent when spreading information. This has led many organizations to use the term misinformation in a broader sense to relate to both phenomena, regardless of intent.

This report, for ease of reference, will also refer to misinformation as the larger phenomenon of creation and spread of false information, regardless of intent. References to disinformation will carry the meaning of that added intention to cause harm or deceit.

It is essential to distinguish these information disorders from other widespread phenomena.

The first is the popular term "fake news". This generic term relates to the transmission of false information in the form of news, capitalizing on the credibility of traditional media. Fake news often encompasses traits of both disinformation and misinformation, depending on whether it is part of a general campaign to spread falsehoods or the result of involuntary sharing by media outlets. Due to its popular use by politicians, the concept has lost any meaningful scientific meaning, especially in light of the ambiguity with which it has been used in recent years to also encompass disagreeable or inconvenient information, despite having circulated as a working concept for almost 100 years (12).

The European Union's High-Level Panel on Disinformation strongly argues that "fake news" should be altogether abandoned as a working term in the field for two main reasons: one is the above mentioned weaponization of the term by politicians; and the other is the acknowledgement that misinformation involves much more than the spread of "fake content or the mimetization of "news", and includes techniques such as astroturfing, fake accounts, deepfakes, targeted advertising and organized trolling, among others.

For these reasons, this report will also abstain from using the term unless the context demands otherwise.

The second information disorder is conspiracy theories, which encompasses a much larger reality in which one can also find disinformation and misinformation practices. Conspiracy theories have been described as "morality tales based on archetypal narratives about right versus wrong, good versus evil" (13).

Conspiracy theories attempt to explain certain realities by referencing secret plots attributed to ill-intentioned individuals or organizations that are alleged to have been revealed against the plotters' will (14). They are often appealing stories that frame factual events and justify actual conduct by connecting them in a made-up narrative.

## Understanding the problem: disinformation and misinformation in health

According to Google (15), the total daily health-related queries at the end of the 2010s amounted to more than 70 000 per minute (more than one billion every day), ranging from conditions, symptoms and medication, through to health insurance questions. These numbers are likely to have increased, even if they may have been somehow rendered incomparable due to the onset of a pandemic. Studies have confirmed this: increasingly, we are looking for health information online (16).

The Internet has lowered the costs of access to health information dramatically, having multiplied sources and cheapened or made accessible access to expert advice, facilitated by the widespread adoption of smartphones: cheap, small, individual and portable points of access that can presently be found in more than three billion pockets around the world (17), even without considering laptops, desktops, tablets and other Internet-access devices.

Simultaneously, and quite inevitably, the threshold for broadcasting health information has also lowered dramatically: the same devices that allow us to instantly find information about anything also enable us to create, publish and share information online, submit it to information catalogues searched by others and design it to look however we want.

People online arrange themselves around semi-public social communities with which they share and from where they obtain content and information, with a considerable degree of trust.

Misinformation – which has been around for centuries (18) – has become more abundant and its effects more impactful due to the powerful combination of instantaneous disintermediated communication and a newfound ability to amplify content at a global scale.<sup>1</sup>

This was, in hindsight, a logical development. It was simply made unequivocal by the emerging and highly competitive market for people's attention (19) and the now-apparent commercial benefits of sensationalizing and oversimplifying complex topics in exchange for viral content, views, likes, shares and the accompanying advertising revenue in a surveillance capitalist model (20).

---

1 Disintermediation is a term used to describe the reduction or elimination of the use of intermediaries between producers and consumers.



## Misinformation on NCD risk factors

This Toolkit interprets health misinformation as encompassing the spread of false information from any source concerning health topics whether causes, symptoms, effects, diseases, treatments, prevention or risk factors, through to the accessibility and status of health-care services and other related topics such as medical research information or public health policy developments.

The existing literature on this topic has mainly focused on vaccination-related misinformation, a topic that gained national prominence with the attachment of many celebrities to the antivaccination movement. Such prominence led WHO to identify so-called "vaccine hesitancy" as one of the top 10 threats to global health in 2019. Research around the topic is often focused on communication: how the Internet influences vaccination decision (21); how the Internet is used to spread anti-vaccine sentiment and misinformation (22); how audiences perceive risk and make vaccination decisions (23); how we can measure sentiment around it (24); how negative sentiment leads to activism (25); and how one can participate in online debates about the topic in order to convert antivaccination proponents (26). Research has also developed around the media through which health misinformation flows: from YouTube videos (27,28), to tweets (29), to Facebook comments (30).

Epidemics have also been fertile ground for health misinformation research, from misinformation analysis about Zika on Twitter (31), Facebook (32), Instagram (33) and YouTube (34), to Twitter and Sina Weibo manifestations (35) and YouTube misinformation about the Ebola virus (36). More recently, the coronavirus pandemic has highlighted the need for a comprehensive approach to health misinformation, as rumours and false information spread quickly and provide an added layer of complication to the implementation of public health safety measures (37-39). Over 25% of the most-viewed YouTube videos on COVID-19 contain misleading information, reaching millions of viewers worldwide (40). Twitter has gone as far as to temporarily suspend influential users from its platform for retweeting misleading information related to COVID-19, as it is now clear that the platform serves as an essential hub for the spread of health misinformation (41).

A recent systematic review of literature on the prevalence of health misinformation in social media revealed that health misinformation was most prevalent in studies related to smoking products and drugs such as opioids and marijuana. The proportion of posts containing misinformation reached 87% in some studies. Health misinformation about vaccines was also very common (43%), particularly regarding the human papilloma virus vaccine. Health misinformation related to diets or pro-eating disorder arguments were less prevalent (36%). Studies focused on diseases such as NCDs also reported moderate misinformation rates (40%), especially related to cancer. For a systematic review of existing literature see the paper Infodemics and health misinformation: a systematic review of reviews (42).

The impact of misinformation on chronic NCDs has benefited considerably less from researchers' attention, covering to date some cancer-related topics (43) and anticancer screening; a Twitter analysis on gynaecological cancer misinformation (44); and some research on the quality of health information on YouTube for particular diseases: heart disease (45); hypertension (46); psoriasis (47) and diabetes (48).

Diet and nutrition topics are seldomly researched from a misinformation perspective, despite being some of the most-accessed content online (49). A study on anorexia-related misinformation disseminated through YouTube videos reveals that pro-anorexia content, despite being less common than informative videos, is more highly favoured and rated by its viewers. Other relevant studies on the topic of anorexia have been published since, focused on other platforms such as Instagram (50,51). Likewise, although there is plenty of research on the intersection of tobacco and social media, the specific impact of misinformation contained therein remains underexplored (52), although misleading tobacco content is on the rise on YouTube (53).

Data show that experts have been studying health misinformation more seriously since the late 2000s, with the number of papers devoted to the topic increasing. A systematic literature review revealed that more than half of these were about communicable (infectious) diseases and included vaccination; a tenth of the literature reviewed covered chronic NCDs such as cancer and cardiovascular diseases, and another tenth covered nutrition and smoking (54). Currently, most existing online research on health misinformation is focused on the spread and characteristics of the information shared. Although adverse consequences are widely recognized, the influence of health misinformation on individual behaviours remains underexplored (55).

One of the main challenges facing research, according to researchers and other non-industry stakeholders, is a difficulty in accessing digital platforms' data, particularly by outside researchers. This means that researchers often work with partial or insufficient data that hinders their ability to achieve meaningful results (56). This difficulty was also addressed by the European Digital Media Observatory, one of the participants of the high-level meetings hosted by the WHO, in a report on platform-to-researcher data access (57).

This imbalance in research has left important gaps which should be acknowledged. These gaps are more than merely theoretical and have practical implications: they affect social media platform structures and behavioural patterns of health information recipients, and condition the strategic communication of health authorities.

It is essential, therefore, to identify these gaps to both inform high-quality research and policy-making on misinformation in the field of NCDs and overall nutrition information, and to propose viable regulatory pathways.



## Sources of misinformation

The sources of misinformation are often complex and multilayered. They stem from multiple actors who use different platforms to willingly and unwillingly spread false health claims. However, misinformation displays some patterns which help to trace it back to its primary sources.

One thing to keep in mind is that not all misinformation starts with the intention of tricking another person. This became apparent during the COVID-19 infodemic, where many well-intentioned people uncritically shared unconfirmed health information based on the good-faith belief that they would be helping others to fight the virus. In comparison, actors such as WHO worked actively to manage the infodemic positively (58). Also pertinent is that it may be tough to pinpoint the source of a particular misinformation event, especially if it starts on private messaging apps and spreads in private chats. Such sources are often unwilling and unknowing sources of misinformation.

Others critically spread verifiably false information justified in their rejection of scientific evidence from a sincerely held belief of the accuracy of their knowledge, ignorance and a deep mistrust of institutions. It is debatable whether these were willing or unwilling spreaders of misinformation.

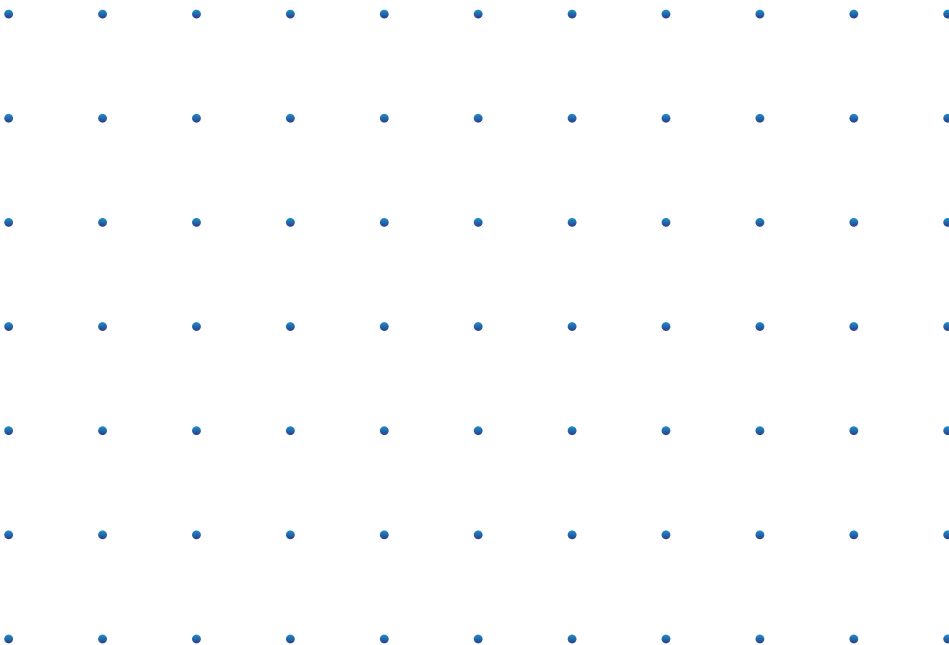
There are plenty of other clear and identifiable sources of disinformation, knowingly creating and spreading false information, motivated by political or commercial interests and taking advantage of the lack of media literacy and human nature. These sources take advantage of a lack of moderation to generate attention and traffic, generating profit or political gain.

**Fake news farms** are businesses that employ people to create false information concerning viral and topical issues and push them onto social media to generate traction, clicks and money (59). Some of these have been exposed in the past by journalists. In an age where the attention economy generates massive profits, it is not surprising that such business ventures have found a way to proliferate. A study by the Center for Countering Digital Hate (60) found that 65% of false information about COVID-19 vaccines shared on the leading social media platforms could be traced back to only 12 sources. Selling "likes" is also a profitable business: one that researchers have pointed out could be better addressed by platforms in their efforts to combat inauthentic behaviour (61).

Certain polarizing topics – such as climate change, politics, migration or even health policy, especially in the context of a pandemic – are particularly prone to misinformation. These topics serve as fertile places for misinformation to develop but are not, per se, responsible for its spread. Here, the role of online platforms and the use of algorithmic content recommendations play a significant role in the dissemination of misinformation. Online platforms are, despite their public efforts to control it, avenues for misinformation to quickly spread by amplifying existing speech while simultaneously clustering information around commercial interests. In many cases, these same systems have been leveraged to spread disinformation campaigns that are then unwillingly shared by Internet users.

E-commerce platforms have also been found to contribute to the amplification of certain types of health misinformation (62). This occurs due to algorithmic curation of content based on past-searches, which increases the risk of locking-up users in misinformation bubbles.

The fact that we are dealing with multiple and widely different sources of misinformation, with different incentives, makes this problem particularly complex. Like a virus, the spread of health misinformation is hard to contain and requires more than good intentions and a passing knowledge of the risks involved.



## Gatekeepers

Disinformation and misinformation travel fast through different channels. Given the amount of information generated every minute, both online and in the physical world, decisions must be constantly made as to what people watch, read and listen to.

An important historical role of players such as traditional media has been to gatekeep information. The term is attributed initially to social scientist and psychologist Kurt Lewin, who in 1947 aimed at understanding how group decisions worked in exchanges of information. The author spoke of the existence of multiple channels through which information flowed and was mediated by "in-or-out" decision "gates", points of decision whereby information could be transmitted or retained (63,64).

In its most simple form, gatekeeping means simply selecting what ought to be transmitted and what ought not to be transmitted between two or more actors. It is "the process by which the billions of messages that are available in the world get cut down and transformed into the hundreds of messages that reach a given person on a given day" (65,66).

Through the works of David Manning White<sup>2</sup> (67), the concept of gatekeepers became popularized within media and the term now refers often to professions such as journalists or other media actors. According to this upgraded view, the editor, publisher or news organization is a benign controller of information, making choices on what information should be conveyed and what should not.

However, traditional media, historically in the frontlines of the fight against false information, has been subject to certain challenges that have made that role harder to uphold.

Financial difficulties, political attacks, a degradation of work conditions and autonomy for journalists and other limitations to press freedom, among other things, have caused journalism to be considered "completely or partly blocked in 73% of the 180 countries" ranked by Reporters Without Borders in 2021 (68). Newspapers, television and radio have been forced to adopt new methods to attract audiences amid increasing competition by new and free social media venues. Click-based advertising became, for many, the main source of income, which promoted the use of sensationalism and click-bait content – headlines designed to trick people into clicking on them just to find out the content has little to do with the headline. This has contributed to erosion of trust in the media.

More specifically to health information, public health authorities, doctors and other health workers have historically acted as information gatekeepers, digesting scientific research and new evidence and producing information that is communicated to audiences. Recent surveys suggest that the pandemic may have contributed to the erosion of trust in public health authorities (69) and that nurses and doctors garner the highest trust scores. Health professionals have been the ultimate source of health information for patients, and their front-line position and expertise are still highly regarded among audiences. This makes them important actors in the effort to promote health literacy and credible sources, as well as to debunk and prebunk misinformation.

---

2 White's work is often cited alongside a different author, Warren Breed – although his contributions have no direct impact on gatekeeping theory.



Additionally, disinformation and misinformation now flow through different channels, mediated by new gatekeepers.

The rise of the platform economy and horizontal exchanges of information justify speaking of the new concept of so-called "digital gatekeepers": digital platforms that mediate exchanges of information between sources and audiences. Importantly, this concept is narrower than the one recently adopted by the EU in the Digital Markets Act, which describes a digital gatekeeper as an undertaking which has a significant impact on the internal market, the control of an important gateway for business users towards final users and an entrenched and durable position (70). This Report uses the term digital gatekeeper to mean a larger reality, namely an online platform with significant impact on the exchanges of information between sources and audiences.

The rise of the platform economy and horizontal exchanges of information has given rise to the new concept of "digital gatekeepers": digital platforms that mediate exchanges of information between sources and audiences. Although online platforms such as Google or Facebook are not considered publishers and hence are not subject to the same gatekeeping obligations (including deontologic obligations) they nonetheless perform exercises of gatekeeping. In practice, through content moderation tools such as terms of service or community guidelines, these companies control and restrict speech, allowing certain views to prevail over others. Just like journalists, social media platforms are responsible for making difficult choices in gatekeeping information.

In addition, as AI mechanisms are used more and more by such online platforms, the original idea of subjective gatekeeping loses more and more meaning; it becomes more important to establish how the algorithm arrived at a decision (explainability), rather than the subjective reasons why it favoured some information over another.



## Main challenges

During the Forum meetings participants, including experts, journalists, social media representatives and government officials, described the main challenges associated with health misinformation, providing exciting and fertile ground for policy-makers to work towards addressing them.

In this report, we will provide an overview of these challenges.

### *Low levels of health literacy*

The first challenge identified was that of low levels of health literacy. Although there have been examples of high-impact awareness campaigns on certain health-related topics over the years (sexually transmitted diseases, tobacco and drugs are topics often covered in schools), health literacy levels are still drastically low (71-73).

Excessive consumption of substances that create a high risk of developing NCDs remains largely common globally, from the ubiquitous overconsumption of salt and sugar to use of alcohol, tobacco or drugs. Dietary and nutritional information is prime terrain for misinformation due to the high virality of the topic and the large pool of nonexpert sources communicating on the topic. For example, a British study from 2012 found that around 70% of the reviewed dietary health claims made by newspapers in the United Kingdom had insufficient and unconvincing evidence. Misreporting of dietary advice by these newspapers was, therefore, considered widespread and deemed to contribute to public misconceptions about food and health (74). Another study, from 2013, on anorexia-related misinformation disseminated through YouTube videos, reveals that pro-anorexia content, despite being less common than informative videos, is more highly favoured and rated by its viewers (75).

### *Low ability to assess the credibility of information sources*

Somewhat correlated with a low ability to understand health content, another highlighted challenge was the low ability of audiences to select appropriate sources of information. This was framed as unrelated to low digital literacy levels, as young people, who tend to have higher digital literacy levels, exhibited similar difficulties in recognizing credible from non-credible online information sources.

The suggestion from the meetings was that audiences who receive training on how to appraise and access reliable sources of information critically are less vulnerable to online misinformation.

### *False speech is cheap*

A second point highlighted during the meetings was the low cost of producing false speech compared with reliable information. It is much easier to produce false information than factually accurate health information, which often requires study, research, confirmation and review. That difference alone poses a major practical challenge when discussing ways to contain the spread of misinformation.

The cheapening of speech (76) in general was both an exciting development in the democratization of access to larger, sometimes global, audiences and a direct cause of the information overflow that fosters the environment in which misinformation thrives.

The erosion of the role of professional intermediaries, such as journalists, doctors and teachers in distributing information has allowed unfiltered access to audiences and lowered incentives and pressure to deliver accurate, high-quality information.

The fact that false information is both easy to fabricate and easy to spread directly to audiences creates a practical challenge that is hard to tackle without attacking its fundamentally good foundation: it has never been easier for people to communicate with each other.

### ***Virality is fed by uncertainty***

Online content goes viral when it plays with strong emotions (77). Health topics are, by nature, prone to elicit feelings of uncertainty, fear, anger, unfairness, excitement, hope and curiosity about things that are very particular and important to individuals.

Topics such as cancer treatments and causes, diet implications, and NCD risk factors, including consumption of alcohol, tobacco or drugs, are fertile ground for misinformation to spread faster than most anti-misinformation strategies can cope with.

### ***Overabundance of information***

A difficult challenge to address is that of the overabundance of available information. The amount of peer-reviewed, scientific research on NCDs available to read online is already impossible to consume and process. When supplemented with accurate information, inaccurate information and outright falsehoods, this forms an impenetrably difficult informational landscape. If one wishes to increase awareness and build reliable information, one must acknowledge the need to filter, organize and present good information pedagogically.

### ***Freedom of speech and censorship***

Perhaps the biggest challenge regarding health misinformation is its constant balancing between conflicting values of contemporary democratic societies. As new technologies amplify free speech, its limitations are intrinsically more difficult to define. The most immediate and effective reaction to the spread of misinformation is suppressing false speech, by erasing it, drowning it, making it inaccessible or silencing the source. However, these strategies may collide with constitutional guarantees of freedom of speech, put in place to prevent censorship, which has historically always been based on narrative around higher purposes. Such constitutional protections limit the ability to suppress false speech as a policy against misinformation easily. Another side of this problem is that actions perceived as censorious may draw attention, victimize the sources, encourage conspiracy theories and ultimately not contribute to reducing the spread of misinformation. This is, however, a challenge with many solutions, as recognized by institutions (78) and scholars alike (79,80).



### ***Mistrust in public authorities***

Access to reliable information depends on the trust placed on public authorities, namely those who convey health facts and scientific data. Although trust in public authorities varies tremendously from country to country and from institution to institution within a country, latest studies claim that institutional trust is decreasing. According to the 2021 Edelman trust barometer (81) "a new era of information bankruptcy and a trust ecosystem unable to confront it" can be observed. Efforts by policy-makers to fight the spread of misinformation have inevitably to contend with this trust landscape.

### ***Cognitive biases***

We all tend to feel invulnerable and subconsciously use heuristics that filter the information presented to us in a way that tends to reinforce pre-existing biases rather than challenging or dismantling them. We also overestimate our ability to critically assess information while at the same time defaulting to the truth when we do not have a reason to believe otherwise.

### ***Speed, spread and scale***

The three concepts of speed, spread and scale reflect a structural change in how speech travels in contemporary times. Misinformation is not new, but it can spread faster, further and at a higher volume than ever before in this digital age. Any lie can be shared quickly through online platforms that spread through different continents and jurisdictions without borders. This structural change, allied to disinformation's inherent characteristic of spreading more quickly than accurate information, creates a further obstacle to debunking strategies (82).

### ***Health authorities are not listening carefully nor communicating effectively***

Public health bodies do not currently invest in "social listening". More effort is needed to hire qualified staff to analyse digital trends, segment populations and understand the different approaches required for different groups. Often, disinformation begins in small events and small communities and then spreads and grows into the larger digital sphere. If health authorities wish to contain this spread, they must act fast and directly on the source of the problem.

Moreover, the current ubiquitous approach is based on the outmoded and ineffective knowledge deficit model: assuming that simply providing people with information is sufficient. This model is flawed; confirmation bias and assimilation bias mean that new information commonly serves to reinforce pre-existing world views. This is a problem for people from all backgrounds and levels of education and evidence-based communication policies are needed.

### ***Experts are not communicating effectively***

Advertisers, headline writers, social media companies and even AI bots understand that humans are drawn to contentious, colourful, scandalous, emotionally charged content, rather than mathematics and statistics, yet most public health content is as dry and boring as it is factually accurate. A new way of communicating evidence-informed information must be found to connect experts and audiences in a disintermediated world.

### ***Profit and agnostic algorithms***

Health misinformation is a million-dollar business (83). There is money to be made from advertising, which is predicated on maximizing Internet traffic, clicks and views. In a surveillance capitalist model, Internet platforms profit from clustering communities into commercial interests and health misinformation provides an interesting connection between users. For the most part, online platforms initially chose to remain largely agnostic to the factual accuracy of the health content they host. Their algorithms funnel users towards content likely to attract and keep their attention and misinformation tends to fit that bill. This creates a perverse financial incentive to direct users to information that harms their health. The COVID-19 infodemic forced some policies to change, but this continues to be one of the main challenges: distributing false information is profitable.

### ***Underused co-creation***

Because misbelieving groups are not monolithic, some of the most effective strategies to tackle misinformation can come from working with affected communities to understand why they feel that way and co-create solutions. However, co-creation is expensive and requires skills and personnel that are not currently abundant in public health.

### ***Horizon clashes***

The need to address the problem in the short term drives resources to fast-action solutions, such as content moderation, fact-checking or labelling, and away from investment in longer-term goals of addressing the root causes through methods such as education, working with industry and developing effective legislation.

### ***Key demographics are hard to reach***

Although all social classes are affected by misinformation, evidence suggests that young people, minority groups and those with low incomes and levels of education may be more exposed and vulnerable to misinformation. Health deserts commonly overlap with credible journalism deserts, meaning those at greatest risk of NCDs also suffer from poor access to accurate and timely health information. These factors can exacerbate structural health inequalities.

Conspiracy theories also often flourish among those who feel left behind, powerless and alienated. They appeal because they help people to feel that they belong to a group, feel safe and understand the world. Deep socioeconomic structural causes underlie this issue.

### ***Inverse care law***

Tools to improve digital literacy and critical appraisal tend to be used the least by those who need them most. Similarly, existing social media fact-check warnings are disproportionately focused on English-language content. A far smaller proportion of misleading posts are tagged in other languages.

This fosters a paradigm where the most vulnerable audiences are those least reached by anti-misinformation strategies. Therefore, it is vital that such strategies are implemented locally and globally. That requires the involvement of local government, NGOs and traditional media, as digital platforms are less likely to direct attention to those communities. Nonetheless, their contribution, namely by making misinformation data available, is crucial to successfully tackle this issue.

### ***Multiple targets***

In conclusion, there is no silver bullet when it comes to misinformation. Audiences need direct action to improve their media literacy skills and to receive sound information from journalists, educators, fact-checkers and politicians. There are also different levels of action to trigger: individuals, businesses, government and wider society all have a part to play.

The motivations and reasons behind such beliefs differ among believers of conspiracy theories: a COVID-19 anti-vaxxer may have no problem with taking other vaccines. This is a moving puzzle made of problems with different solutions and powered by people with different incentives, making it particularly hard to find the right approach.

It is necessary to advance an all-encompassing strategy where all levels and multiple actors cooperate to fight health disinformation.

## A multilayered problem

The challenges above indicate that the problem of health disinformation and misinformation is a complex one. It is made of different sources, actors and means of communication, making fragmented regulatory measures ineffective. In order to tackle such a multilayered challenge, there is a dire need to foster cooperation between existing levels of regulation. A systematic, comprehensive approach is needed.

In analysing infodemics, Eysenbach (84) proposed a four-pillar framework to fight them: (i) science; (ii) policy and practice; (iii) news media; and (iv) social media. This underpins some of the most important challenges at stake.

The first pillar, focused on accurately translating information between levels (from science to policy, from policy to news, from news to the people), is challenged by all the externalities that compromise accurate translation: political biases, commercial interests, selective attention reporting and general misunderstandings. There needs to be a coalition of stakeholders, from academics to governments to the news industry, to ensure that certain norms are adopted (for example, preserving the chain of reference by promoting the previous source of information).

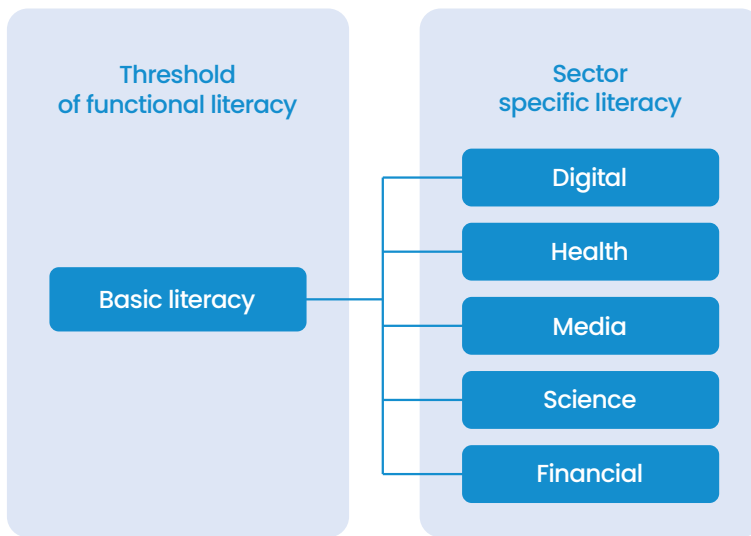
The second pillar pertains to knowledge refinement through filtering and fact-checking. This would include the view that the media industry and social media should adopt transparent and cross-market norms, such as standardized labelling of false content and similar filtering and algorithmic practices. The EU's 2022 Strengthened Code of Practice on Disinformation contains important guidance on this matter. (78) Transparency and consistency across platforms are key, as fact-checking and filtering could otherwise be inaccurately perceived as partisan attempts to hamper free speech in certain jurisdictions or contexts. The legal and constitutional considerations here should be further explored.

The third pillar is e-health literacy, where governments and civil society have a crucial role to play at two different levels: first, digital literacy, which can be understood as the specific skills and tools necessary to exchange ideas, assess information and to make informed actions in an online environment; (85) and secondly the more specific health literacy dimension. According to WHO, health literacy represents the cognitive and social skills that determine individuals' motivation and ability to gain access to, understand and use information in ways that promote and maintain good health (86).

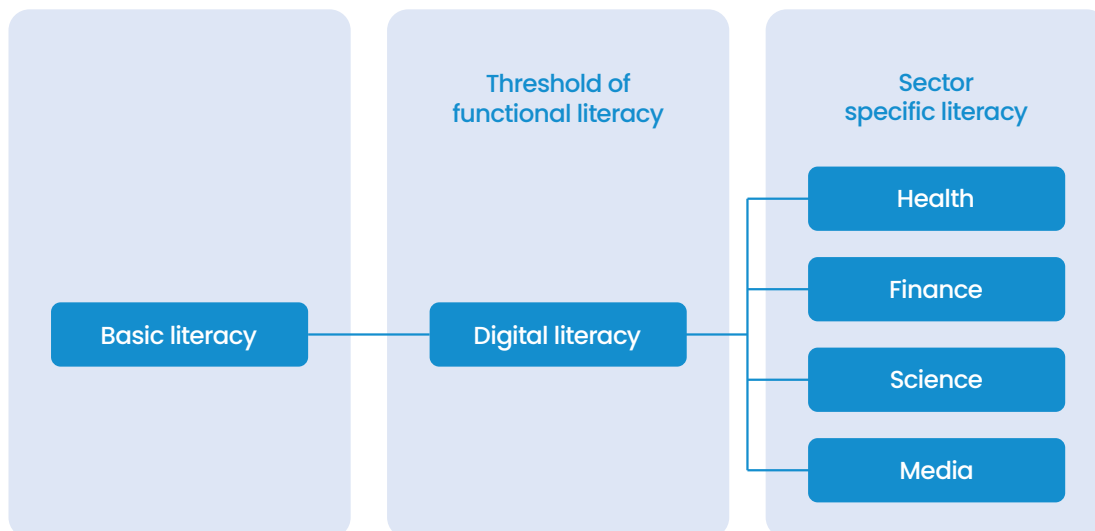
In fact, WHO identified health literacy as playing a central role in determining inequities in health in both rich and poor countries (87). Responsiveness to health education, the use of disease prevention services and poor self-management of NCDs have all been linked to low literacy levels (88).

Nowadays, however, attaining health literacy, as well as other field-specific literacies such as financial literacy, media literacy or science literacy, requires an ever-growing level of digital literacy. Although digital literacy could formally be considered just another field-specific literacy, it is now almost as essential as the basic literacy skills of reading and writing: the threshold of functional literacy, encompassing what can be considered as sufficient basic skills enabling an individual to function effectively, (89) has moved in the past decade from the classical threshold (Fig.2) to a modernized novel threshold (Fig.3).

**Fig. 2.** Classical literacy threshold



**Fig. 3.** Novel literacy threshold



This change requires more than mere acknowledgement: it must be an absolute priority to ensure that individuals have the sufficient skills and tools to be considered digitally literate. This is the fulcrum of any serious effort to tackle misinformation, mainly health misinformation.

Finally, the fourth pillar pertains to what Eysenbach labels "infodemiology and infoveillance": continuous monitoring and analysis of data and information-exchange patterns on the Internet (90). This requires research, data collection and mining and a better understanding of how information is being exchanged online.

This framework is an excellent response to the question "what should we focus on?". However, an answer is also needed for the question "who should be doing what?".

## Consulting stakeholders: outcomes

### Member States

This chapter looks at positive initiatives carried out by Member States and international organizations aimed at countering the spread of misinformation in general and, when available, health misinformation in particular. These initiatives differ in their methods and scope, but all aim to diminish the impact of the spread of misinformation in general.

The efforts have been of different nature, ranging from adopting anti-misinformation laws, creating working groups or task forces, and soft-law approaches such as responding to the gravest infractions with network shutdowns. An interactive map of these initiatives has been published by the Poynter Institute (91).

Several of the following initiatives relied on some degree of cooperation between government, civil society and industry and were promoted by policy-makers attempting to target an increasing problem.

The EU has been the most active voice in the WHO European Region regarding outlining a pathway to fight online misinformation. As it unequivocally advocates, the importance of cross-society cooperation cannot be underestimated: "large-scale disinformation campaigns are a major challenge for Europe and require a coordinated response from EU countries, EU institutions, social networks, news media and EU citizens" (92).

In 2017 a High-level Expert Group was appointed to issue recommendations on a European strategy to counter misinformation. In the report prepared by that group, it was recognized that online misinformation required a multidimensional approach sustained by five pillars (93):

- promotion of media literacy
- protection of competition and sustainability of the European news media ecosystem
- promotion of digital transparency
- development of new ways of engaging with newsreaders
- continuous monitoring of the status and effectiveness of the strategies.

The group report was focused on disinformation, i.e. the intentional spread of false information, and should be understood in that context: as a reaction to a problem caused by bad actors rather than as a solution to a problem where victims are often unwilling perpetrators.

It is, notwithstanding, a valuable effort to guide Member States interested in targeting misinformation via public policy and legislation.

The Council of Europe has also contributed meaningfully to this discussion by commissioning the report *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy-Making* (10). This report maps the problems of information disorder and information pollution and offers solutions.

The European Commission (EC) has consolidated its anti-disinformation strategy in the Action Plan against Disinformation (94). This Action Plan looks at disinformation as a weapon used by foreign powers to undermine peace and democracy for geopolitical gain (Table 1).

**Table 1.** The EU Action Plan on Disinformation

Pillar	Actions
I. Improving the capabilities of Union institutions to detect, analyse and expose disinformation	<ol style="list-style-type: none"> <li>1. Strengthen the StratCom task forces and EU Delegations with additional resources (human and financial) to detect, analyse and expose disinformation activities</li> <li>2. Review of the Task Force South and Task Force Western Balkans mandates</li> </ol>
II. Strengthening coordinated and joint responses to disinformation	<ol style="list-style-type: none"> <li>3. Establish by March 2019 a Rapid Alert System that works closely with other existing networks (EP, NATO and G7)</li> <li>4. Step up communication pre-EP elections</li> <li>5. Strengthen strategic communications in the neighbourhood</li> </ol>
III. Mobilising private sector to tackle disinformation	<ol style="list-style-type: none"> <li>6. Close and continuous monitoring of the implementation of the Code of Practice, including push for rapid and effective compliance, and a comprehensive assessment after 12 months</li> </ol>
IV. Raising awareness and improving societal resilience	<ol style="list-style-type: none"> <li>7. With Member States, organise targeted campaigns for to raise awareness of the negative effects of disinformation, and support work of independent media and quality journalism</li> <li>8. Member States should support the creation of teams of multidisciplinary independent fact-checkers and researchers to detect and expose disinformation campaigns</li> <li>9. Promotion of media literacy, including through Media Literacy Week (March 2019) and rapid implementation of the relevant provisions of the Audio-visual Media Services Directive</li> <li>10. Effective follow-up of the Elections Package, notably the Recommendation</li> </ol>

Note: EP: European Parliament; NATO: North Atlantic Treaty Organization; G7: Group of Seven.

Source: European Commission, 2018 (94).

Some countries have recognized the importance of the topic by establishing a clear connection between the fight against misinformation and human rights such as Canada in their Digital Charte (95), Portugal in their Charter of Human Rights in the Digital Era (96), and Spain in their Charter of Digital Rights (97).

The same ideas are now being proposed by the EC to the European Parliament and the Council under a Declaration of Rights and Principles (98) to guide the digital transformation of the EU. The document expressly acknowledges the risks of disinformation as well as the spread of illegal and harmful content, putting forth the idea that offline rules and principles should equally apply to the digital world.

The idea of these charters or declarations is to bring the discussion around disinformation to the realm of rights.

### **Platform regulation**

Online platforms are a key player in controlling the spread of disinformation. Consequently, it is no surprise that Member States and the EU have primarily focused on regulating their practices in this sphere.

Firstly, the EC proposed an ambitious European Democracy Action Plan (99), which aimed to set out measures to promote free and fair elections, strengthen media freedom and counter disinformation. The EC acknowledged the foundational role that platforms play in these objectives (100), especially regarding political advertisements. The Action Plan's objectives were then fully realized in the important package of legislation proposed by the EC, which is generally called the Digital Services Act (DSA) (101). The proposal included a general strengthening of the obligations of online platforms in many areas, ranging from accountability measures to content moderation. Notably, regarding disinformation, the DSA proposal equates the concept of disinformation with systemic risk for society and democracy (102). It also calls for a revision of the 2018 Code of Practice on Disinformation (103).

The Code of Practice is the result of EC efforts to co-regulate disinformation with the help of online platforms. It executes the objectives set out in the Joint Action Plan Against Disinformation, (104) namely its ambition to develop a code of conduct on the matter. This code was finally agreed in September 2018, described as a "voluntary, self-regulatory mechanism agreed on by representatives of online platforms, leading social networks, advertisers and advertising industry". The Code focuses on the role that the advertising industry and online platforms can play in combating the spread of online disinformation. It crafts an agreement over the transparency practices to be put in place regarding the scrutiny of advertisement placements and political advertisements, and the role of consumer empowerment to fight disinformation.

The Code of Practice has now been revised and a 2022 version was approved (78,105). This new version integrates some of the worries expressed in earlier evaluations of the 2018 version (such as the lack of meaningful key performance indicators) and – in line with the DSA – now strengthens its provisions on risk assessment and risk mitigation measures for large (more than 45 million average monthly active users) online platforms in the EU. This also includes empowering citizens and researchers to access more relevant data on the fight against disinformation. This new Code adds to other soft-law mechanisms such as the updated European External Action Service strategy on countering foreign information operations and the specific recommendations of the European Commission on combating COVID-19 misinformation.



## Fact-checking initiatives

In 2020 the EC created the European Digital Media Observatory (EDMO) (106): a European hub for fact-checkers, academics and other relevant stakeholders to support policy-makers with the following missions.

- Identifying and supporting fact-checking organizations across Europe by developing collaborative and cross-border activities as well as dedicated training modules.
- Creating and maintaining a global archive of peer-reviewed scientific literature on misinformation, as well as mapping, supporting and coordinating research activity on disinformation at the European level.
- Creating a public platform that provides information and materials to media practitioners, instructors and people in order to raise awareness, create resilience to Internet misinformation and promote media literacy efforts.
- Designing a framework (now published) that allows academics who wish to have a better understanding of disinformation to gain secure and privacy-protected access to platform data (57).
- Supporting public authorities in monitoring Internet platform policies aimed at limiting the transmission and impact of disinformation.

An essential aspect of EDMO is its governance, which was designed to be entirely independent of government authorities despite being funded by the EU's budget.

In Florence, Italy, the European University Institute leads a consortium that manages EDMO. The consortium includes the Greek company Athens Technology Center, the Danish Aarhus University and the Italian fact-checking organization Pagella Politica.

The governance structure of EDMO is entirely independent of public authorities, including the EC. The governance structure consists of an advisory board in charge of defining EDMO's operating rules and strategy and an executive board in charge of implementing the contract in consultation with the advisory board.

In the electoral context, several fact-checking initiatives and hotlines have surfaced. Brazilian authorities have created a fact-checking platform where articles can be submitted for review (107). The Indonesian Government has something similar, having launched a website



where the population could dispel electoral disinformation (108). Australia created the Electoral Integrity Assurance Task Force to curb foreign interference on their elections via misinformation; similarly, Canada created the Critical Election Incident Public Protocol with the same goal.

The Ministry of Health in Bulgaria set up a national hotline to provide reliable information concerning COVID-19 and vaccines, contributing to countering misinformation by providing an accessible way for citizens to clarify their questions.

Many countries such as Croatia, Estonia and the Republic of Moldova developed official online information sources for COVID-19 information and launched social media campaigns that used celebrities to try to incentivize vaccination and dispel myths.

### Labelling

The United Kingdom's National Health Service (NHS) implemented a certification scheme – aptly named The Information Standard – that was intended to be a quality standard that encouraged organizations with a robust information production process based on best practices to produce high-quality information that meets the needs of its users. Institutions communicating about health could submit to certification under The Information Standard and, if passed, could wear a seal of information quality issued by the NHS. This is an example of a label arising from a governmental initiative, carrying the weight and goodwill of the NHS brand into the informational landscape. The Information Standard has been discontinued and eventually led to the creation of the PIF Tick, described below in the civil society section.

Labelling solutions so far have had low recognition from policy-makers as a viable complement or alternative to content moderation and fact-checking to prevent the spread of misinformation. In the EU Code of Practice, platforms have committed to "develop and apply tools or features to inform users, through measures such as labels and notices, that content they interact with has been rated by an independent fact-checker and work to implement them across all EU Member States languages".

Those attending the three meetings thought that this might be explained by the fact that specific solutions have not yet been the subject of academic research, and their effectiveness is yet to be scientifically tested, which is in part due to a lack of access to actual usage data by the tech platforms implementing them.

Despite this, labelling solutions are being promoted in the background by those weary of the government censorship risks involved in content moderation (review and deletion of content, suspension or cancellation of accounts, etc.) as a less harmful way of preventing audiences to come into contact with, and eventually spread, misinformation online.

However, important public policy concerns remain regarding the risks involved in governments supporting or creating labelling solutions that could be used to promote information deemed politically favourable to the detriment of politically inconvenient but accurate reporting. Recent examples are politicians labelling as "fake news media" certain mainstream outlets when reporting was not aligned with their or their party's political goals.

As such, the promotion of labelling solutions by governments can still be seen as a controversial move, but one that could, if well done, empower audiences with more information for them to make their own decisions on what to trust.

## Media and health literacy

Governments and international organizations have also explored the avenue of improving media literacy to empower citizens to combat misinformation.

WHO has been in the forefront of this effort, leading the international health community in producing information and reports for Member States to use in shaping their national policies and reaching out to their citizens. This was clear during the COVID-19 pandemic, where WHO materials directed at dispelling common myths were used and repurposed by countries all around the world. For example, Canada launched the Digital Citizen Initiative (109), a strategy to build resilience against online misinformation by supporting research, funding awareness campaigns and organizing projects on media and digital literacy and their role in anti-misinformation strategies.

In Belgium, the Government invited civil society experts and journalists to discuss options concerning topics such as information quality and solutions for online misinformation. The outcome was a website informing about the perils of misinformation and promoting participative democracy (110). The website had two goals: spreading reliable health information to the public and patients and debunking fake news on health in social media. The Belgian Government has reported that the website is well respected and widely used by the public and pointed to the following reasons behind its success.

- The site is hosted by the Centre for Evidence-based Medicine as this is trusted more than other government institutions.
- Content is reviewed by a panel of patients to check that it is comprehensible before it is shared publicly.
- Health information posts are linked to current events in popular culture. For example, content on Lyme disease was released when pop singer Justin Bieber was diagnosed with this condition.
- The Centre has hired a team to package information in appealing forms that work well online.

The Belgian Government's understanding (111) is that outsourcing the responsibility of ensuring information quality to industry, namely digital platforms, would lead to undesirable outcomes, and that media literacy policies were fundamental to correct the online misinformation problem.

Australia has launched the project HealthLit4Kids, a health literacy project involving schools, teachers, parents and students, applying the OPTimizing HEalth LiterAcY (Ophelia) approach (112). In an electoral context, the Australian Electoral Commission also launched a media literacy campaign in 2019 named Stop and Consider, an advertisement-based social media campaign encouraging audiences to assess election information and check the sources critically (113).

Scotland (United Kingdom) issued the campaign Making it Easy: a Health Literacy Action Plan for Scotland in 2014, focusing on developing health literacy-sensitive initiatives and structures; this has since been further broadened (114).

Some countries have formed health literacy alliances to encourage national agendas on the topic: this was the case in Austria (Austrian Health Literacy Alliance), Germany (Allianz für Gesundheitskompetenz) and Switzerland (AllianzGesundheitskompetenz).

In July 2021, Bulgaria, which has one of the highest rates of premature mortality due to NCDs in the EU, started implementing a new NCD-focused health literacy campaign.

### **Hate speech**

An important factor in combating misinformation and disinformation is the fight against online hate speech. In this field, governments and the EU have taken vital initiatives that aim at eradicate hate speech from online interactions.

Since the 2016 Brussels terrorist attack, the EC has enhanced the calls to fight hate speech, resorting to co-regulatory efforts with Member States and online platforms. Answering such a call, both Member States and industry have updated their frameworks against the problem. These joint efforts have been codified in the Code of Conduct on Countering Illegal Hate Speech Online (115).

The Code establishes the ground rules to deal with speech that entails hate speech or incitement to terrorism, notably the 24-hour take-down rule for certain types of content. It adopts the established definition of racist and xenophobic hate crime and hate speech (116). Facebook, Microsoft, Twitter and YouTube were part of the original signatories of the Code in 2016 but new members have joined, including Dailymotion, Instagram, Jeuxvideo and Snapchat, as well as Tik Tok and LinkedIn in 2021. The Code is an exercise of self-regulation whereby these companies adopt internal measures to fight hate speech.

The Code's latest assessment confirms the critical role of these online platforms in tackling hate speech. The last evaluation shows that on average, the companies are now assessing 81.0% of flagged content within 24 hours and 62.5% of the content deemed illegal hate speech is removed (117).

In Croatia, a long debate over the electronic media bill turned an initial draft proposing to penalize digital platforms for hate speech published by their users into a bill sanctioning the users instead (118,119).

In Germany, hate speech on social media platforms was also specifically targeted by the Network Enforcement Act, with large digital platforms (above two million users) risking fines of up to 50 million euros if they do not remove clearly illegal content within 24 hours and any other illegal content within seven days. The incentives posed by this legislation on digital platform have been subject to criticism, including by the United Nations Special Rapporteur for the Protection of Freedom of Opinion and Expression; however, research does not seem to entirely support this (120).

High fines and vague guidelines combined with short deadlines incentivize digital platforms to overzealously remove content that may be illegal, rather than applying the rules restrictively. However, as removing content often collides with the freedom of speech of the users, any such measures should be applied in a cautious and measured manner. The German Government has maintained that the law is necessary to curb the increase in hate speech and fake news.

France has also approved a similar law – the Avia law – which required digital platforms to remove hateful content within 24 hours. The same criticism was levied at this law, which was struck down, with the French Constitutional Court deeming it partially unconstitutional (121).

The Council noted that the absence of involvement of judicial courts in the determination as to whether specific content published is illegal, as well as the incentives to pre-emptively block speech out of caution, is a clear breach of the constitution. It also underlined that a conclusion on the illegal nature of content requires a level of analysis that is rendered impossible by the short time frames.

### **Truth regulation**

The above-mentioned EC report from 2018 discouraged countries from tackling disinformation via legislation aimed solely at curbing the spread of false information (94).

The reasons appear somewhat obvious: enforcement of such legislation often relies on interpretations of truth and falsity that can be subjective and tend to favour existing power structures, and it can easily be used to target political opponents or uncomfortable news reporting. This may lead to abuses, censorship and violation of free speech rights and guarantees.

Laws with such an aim were nonetheless approved and enforced worldwide, leading to individuals being prosecuted and/or arrested for sharing false information online. One such example comes from Bahrain, where an activist was convicted of spreading hatred and false news (122).

Bangladesh's parliament approved in 2018 the Digital Security Act, which charges with imprisonment those convicted of spreading false news and is largely considered by observers as an instrument to criminalize freedom of the press and freedom of speech (123). In Belarus, a law has been in place since 2018 targeting anyone who is found to spread false information online, including online platforms and websites (124). Burkina Faso, Cambodia and Cameroon have also chosen the route of penalizing those who spread false information with imprisonment (125-127).

More examples exist, demonstrating that there is a follow-up pattern to the implementation of "truth-regulation" legislation and illuminating the sensibility of the EC's High-level Expert Group's recommendation to avoid simple solutions such as criminalizing those who spread lies.

The line between protecting the population from fake news and censorship can be thin and must be walked with care, as freedom of speech and freedom of press are essential concepts within healthy democracies, resilient societies and educated populations.

## Industry

### *Social media initiatives*

Industry is an essential part of the tripartite co-regulatory efforts of a triple entente. By industry, this report generally means the market side of the fight against disinformation and misinformation. It includes companies that develop their activities using information society services and traditional media venues such as newspaper outlets, televisions or radio broadcasting. Such industry initiatives have taken different forms and it is essential to analyse them individually.

### *Setting Standards*

First and foremost, the industry has developed its own standards and practices to combat the spread of disinformation and misinformation. Different companies have adopted different strategies to contain the spread of malicious information, providing a good starting point to our analysis. In terms of setting standards, most online companies and social media platforms now include mentions to combat disinformation/misinformation in their terms and conditions and community guidelines.

Google/YouTube, for example, has developed its own misinformation policies, including the prohibition of "promoting dangerous remedies or cures: content that claims that harmful substances or treatments can have health benefits" (128). This policy has been taken seriously in the context of COVID-19, with YouTube banning Brazilian President Jair Bolsonaro from the platform for a week due to false medical claims. Likewise, Google policies do not allow adverts that potentially profit from or exploit a sensitive event with significant social, cultural or political impact, such as a public health emergency. The platform has also created policies that prohibit monetization of COVID-19 misinformation and pranks and challenges. The same could be said of TikTok under its community guidelines, which expressly reject advertising that contradicts COVID-19 medical information (129).

The same can be said of Facebook/Meta. The company aims at disrupting the economic incentives that might lead to the propagation of misinformation and ensures that its algorithms are constantly fed new data points that recognize misleading or malicious information (130).

The effectiveness and sufficiency of these initiatives is subject to review and debate, with commentators (131) and the EC (132) challenging, for example, lack of consistency in the application of standards. In fact, one of the reasons behind adopting the Code of Practice on Disinformation (133) was that these voluntary standards were judged to be insufficient, following an EC requirement that companies should report on their self-regulatory efforts. COVID-19 misinformation is a particular example where harm can be caused and more detailed analysis of platform policy continues to be conducted (134).



## Enforcing standards

Online platforms have also enforced those standards in different ways. These can be divided between prebunking and debunking strategies. Prebunking strategies aim at controlling disinformation or misinformation before the content is uploaded. Debunking refers to action taken once a piece of misinformation/disinformation is already online and potentially viral. Among the first category, there are practices such as labelling content that might contain disinformation or simply nudging people into reliable and trustworthy information. It would also include literacy efforts – such as investment in early education digital literacy or advertisements to raise awareness – and the use of cutting-edge technology such as blockchain-based applications. As examples of the latter, social and traditional media have adopted content takedown strategies and fact-checking campaigns.

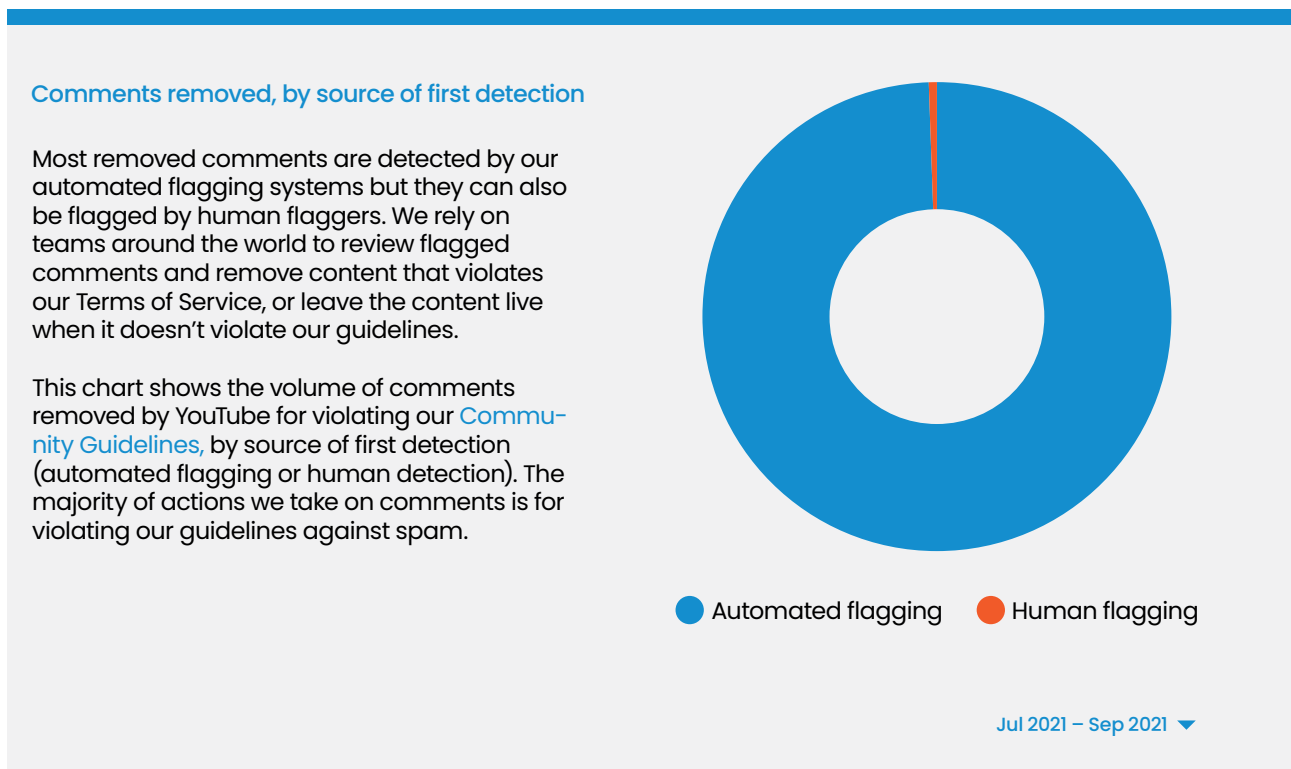
### Prebunking strategies (labelling and nudging)

#### Labelling and pre-upload filters

Prebunking strategies aim to control malicious content before it reaches viral status. They are preventive strategies that help combat disinformation and misinformation at their first stages.

An example of this is how the development of AI technology has reportedly contributed to tracking disinformation and misinformation campaigns before they reach viral status. According to Google, AI takedowns corresponded to 99.6% of all flagging of violations of YouTube Community Guidelines between July and September 2021 (Fig. 4).

**Fig. 4.** AI takedowns of content by YouTube, July–September 2021

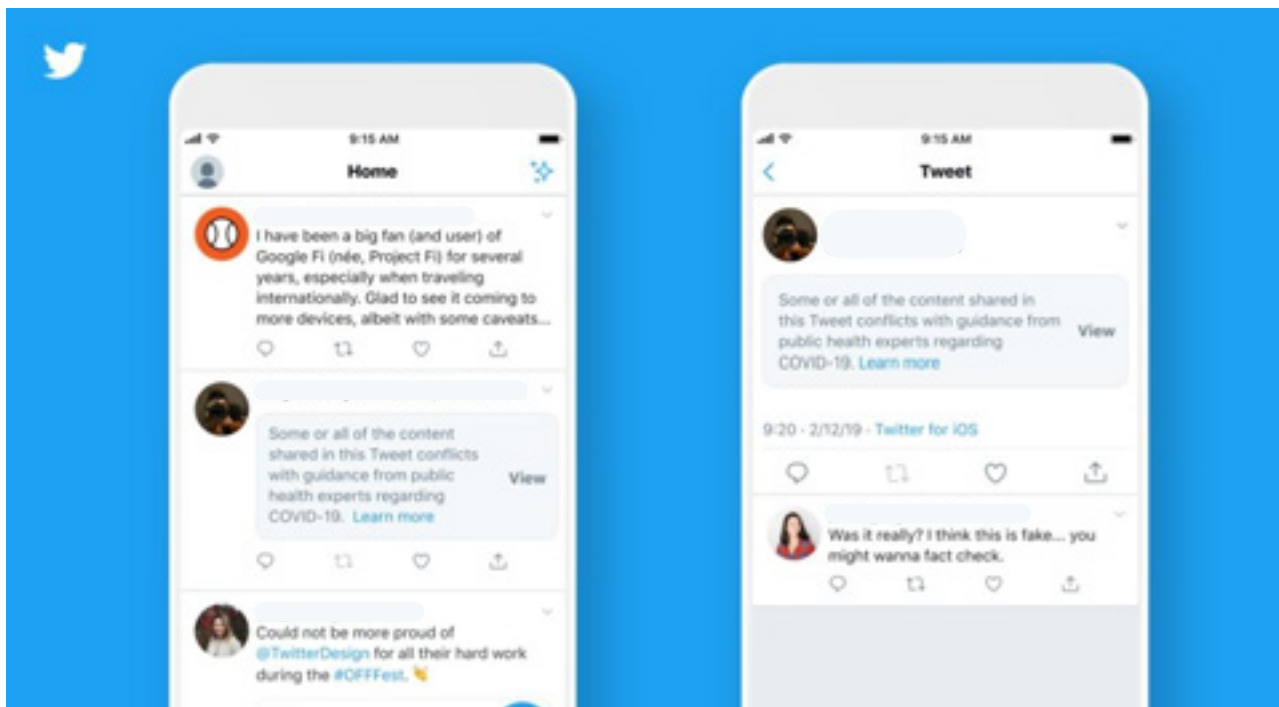


Source: YouTube, 2021 (135).

This means that the industry is increasingly relying on AI systems to combat disinformation and misinformation, more than on human interaction.

Twitter has also made considerable efforts to label pieces of misinformation as such. During the COVID-19 pandemic, it upgraded its policy and revised its structures (136) including a two strike rule for authors who repeatedly post content flagged as disinformation/misinformation. As seen in Fig. 5, Twitter flags some content as contrary to existing expert analysis.

**Fig. 5.** Content flagged as misinformation in a Twitter feed



Source: Twitter, 2022 (137).

Labelling content as contradicting guidance from health officials creates additional friction in accessing content. This will ensure that users make a premeditated decision when reading a tweet. Labelling is often done in collaboration with government and nongovernmental entities, providing the expertise necessary to flag certain content as misleading.

### Nudging

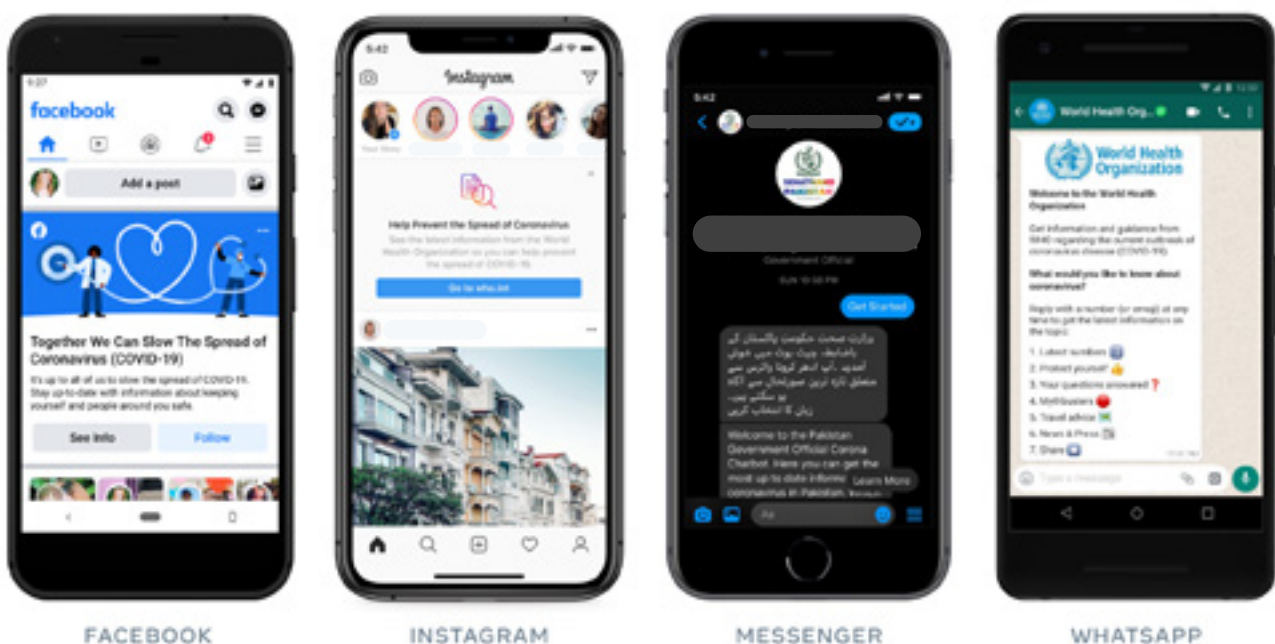
In other cases, platforms label certain content as disinformation or misinformation and also actively nudge users towards reliable information. For example, when ranking search results on their platform, Google Search elevates authoritative information from public health authorities. This means that Google algorithms now incorporate the fight against misinformation and disinformation as a critical element of their ranking methodology. The same concept is utilized by YouTube and has resulted in the reduction of watch-time of low-quality content by 70% on that platform. During the COVID-19 pandemic, Google



Search actively directed people towards reliable information by trending COVID-19 reliable topics. Google Trends also lets people explore what people are searching for, directing people towards specific content.

Likewise, Facebook has created its own centre for information about specific topics (138). During the COVID-19 pandemic, the platform would nudge its users towards this centre to ensure that reliable information was shown to the participants in certain conversations. Across its many platforms (Instagram, Messenger, Facebook and WhatsApp) the company developed strategies to alert and redirect users towards reliable health information such as information about COVID-19 by public authorities (Fig. 6).

**Fig. 6.** Nudging within Facebook, Instagram, Messenger and WhatsApp



Source: Meta, 2022 (139).

### Advertisements and questions and answers (Q&A)

Another method some platforms have used to reinforce the fight against health disinformation is to use traditional advertising channels to raise awareness. WhatsApp engaged directly with its users via television and radio to combat the spread of disinformation. In a famous advertisement in India, (140) the campaign Share Joy, Not Rumours aimed to alert communities to the dangers of spreading false information. Both television and radio advertisements were used to depict scenes in which misinformation affected relatives and friends' health and well-being.

Likewise, some platforms decided to address health disinformation by establishing a series of questions and answers for doubts arising from health policy measures. TikTok, for example, has launched a Q&A regarding COVID-19 and vaccines, the answers to which were provided by WHO (Fig. 7).

**Fig. 7.** TikTok COVID-19 Q&A service

Source: TikTok, 2022 (141).

### **Debunking strategies**

Debunking strategies are put in place once a given misinformation or disinformation content is already online and potentially viral. A fundamental tactic in fighting disinformation and misinformation is to debunk viral campaigns. Both online platforms and mainstream media are responsible for executing these strategies when their platforms and channels carry misleading information. The term refers to more than simply correcting false claims, appearing as an act of truth-telling exposing information previously provided as a sham or grossly exaggerated.

### **Fact-checking**

Fact-checking is an essential dimension of debunking disinformation campaigns. Social media platforms such as Facebook have helped gather the industry around fundamental principles such as the Code of Principles of the International Fact-Checking Network (IFCN). (142) Through Facebook's Third-Party Fact-Checking Program (143), many fact-checkers are called to attach facts and debunk conspiracies on content published on the platform.

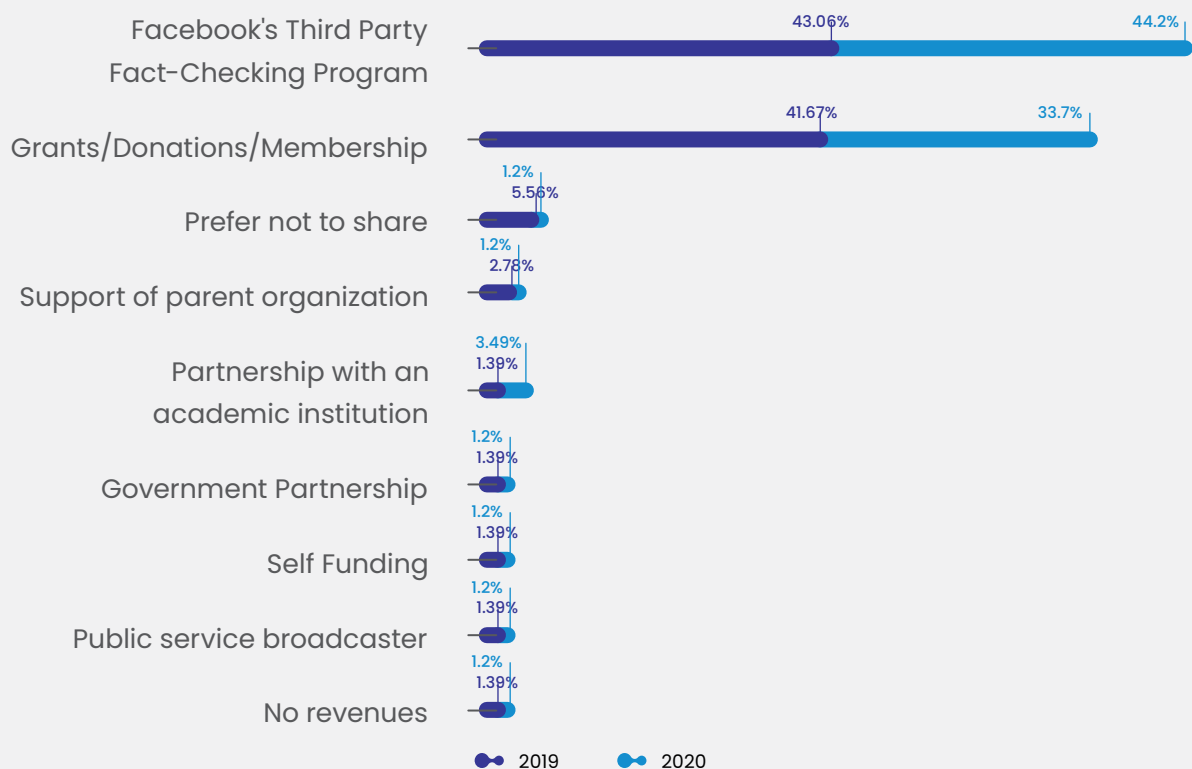
In fact, platforms such as Facebook continue to be the most important financier of fact-checkers around the world as shown by the 2021 Internet Fact-Checking Network report (Fig. 8).

**Fig. 8.** State of the fact-checkers 2021

## Revenue sources

At 44.2%, Meta's Third Party Fact-Checking Program remains the main revenue source for many organizations, at least since we started asking this question. Income from donations, grants or membership, though it declined by about 8% from the previous year, come in second.

### What was your largest source of revenue?



Source: The Poynter Institute, 2021 (144).

Large online platforms such as Twitter or Google also routinely cooperate with fact-checking associations worldwide (142) that abide by similar principles. In August 2021, Twitter decided to establish formal collaborations with two members of the IFCN: Reuters and the Associated Press (145).

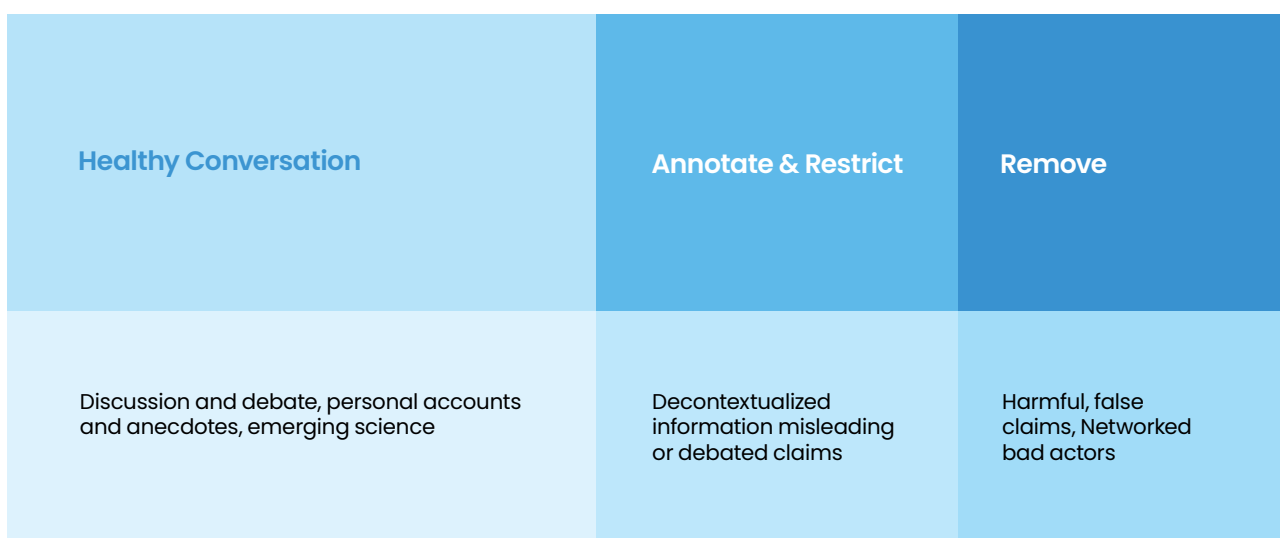
Twitter has also pioneered an innovative peer-to-peer fact-checking programme called BirdWatch (146). The feature is still being tested with a limited number of users but will allow each user to flag spreads of disinformation and provide her/his own views on how the information is incorrect. This will be then supported by independent fact-checkers who will help to debunk certain myths and falsehoods. TikTok has also started engaging with the fact-checking community, first regarding COVID-19 (147) but now also concerning other political matters (147).

### Content takedowns

Social media companies have also developed systems of content takedowns to fight health-related disinformation. Content takedowns are particularly invasive measures by which a given platform eliminates a piece of content or speech from its platform. They are often accompanied by measures such as suspension or expulsion from the platform.

Because these measures greatly impact users’ rights – their freedom of expression – platforms have developed internal guidelines to scale the response, ensuring proportional reactions to specific classes of content. For example, Twitter takes action within conversations according to the spectrum shown in Fig. 9.

**Fig. 9.** Spectrum of acceptable conversation used for Twitter moderation



Source: Twitter.


This spectrum highlights how the platform applies a proportional methodology to content, ranging from stimulating healthy speech, to labelling and annotating contextualizing information, to the complete removal of harmful or false claims. Misinformation often happens within both of the last two areas, rendering this case-to-case approach the only possible way to identify and act.

Within misinformation and disinformation themselves, Twitter adds increased nuance. The company decided to qualify information that is dubious in a tripartite way:

- **misleading information** is statements or assertions that have been confirmed to be false or misleading by subject-matter experts, such as public health authorities;
- **disputed claims** are statements or assertions in which the accuracy, truthfulness or credibility of the claim is contested or unknown; and
- **unverified claims** are information (which could be true or false) that is unconfirmed when it is shared.

These different categories are then treated in different manners, either no action, labelling, warning or, as the most extreme consequence, removal (Fig. 10).

**Fig. 10.** Twitter action consequences for mis- and disinformation dependent on harm and type



<b>Misleading Information</b>	Label	Removal
<b>Disputed Claim</b>	Label	Warning
<b>Unverified Claim</b>	No action	No action*
	Moderate	Severe
<b>Propensity for Harm</b>		

Source: Twitter, 2020 (137).

Facebook has also taken decisive action on countering disinformation and misinformation. The company also adopts a tripartite distinction between three types of action on misleading content: remove, reduce and inform (148).

Content that must be removed and content that must be reduced are both part of Facebook's debunking strategies. Regarding removal, the company removes content in exceptional cases limited to imminent physical harm (e.g. false claims about health practices), voting and elections, and videos that have been manipulated and cannot be identified as such by an average person (so-called "deepfakes"). The company perceives health disinformation as one of the most important categories of misleading content, and one that may lead to deletion.

In the same way as Twitter, Facebook has now established a strike system, in which users are alerted to continuous practices that breach community standards. If a continuous practice does not cease or is repeated over and over again, the three strike rule will apply and the user will face potential expulsion from the platform (149).

### **Traditional media initiatives**

Alongside social media and online platforms, traditional media plays a key role in the fight against health misinformation and disinformation. Many news outlets have embraced the combat against health disinformation by providing innovative ways to track down the sources of misinformation using cutting-edge technology. They have also increased the levels of literacy in order to empower users to be able to distinguish between good and reliable information and misinformation or disinformation campaigns.

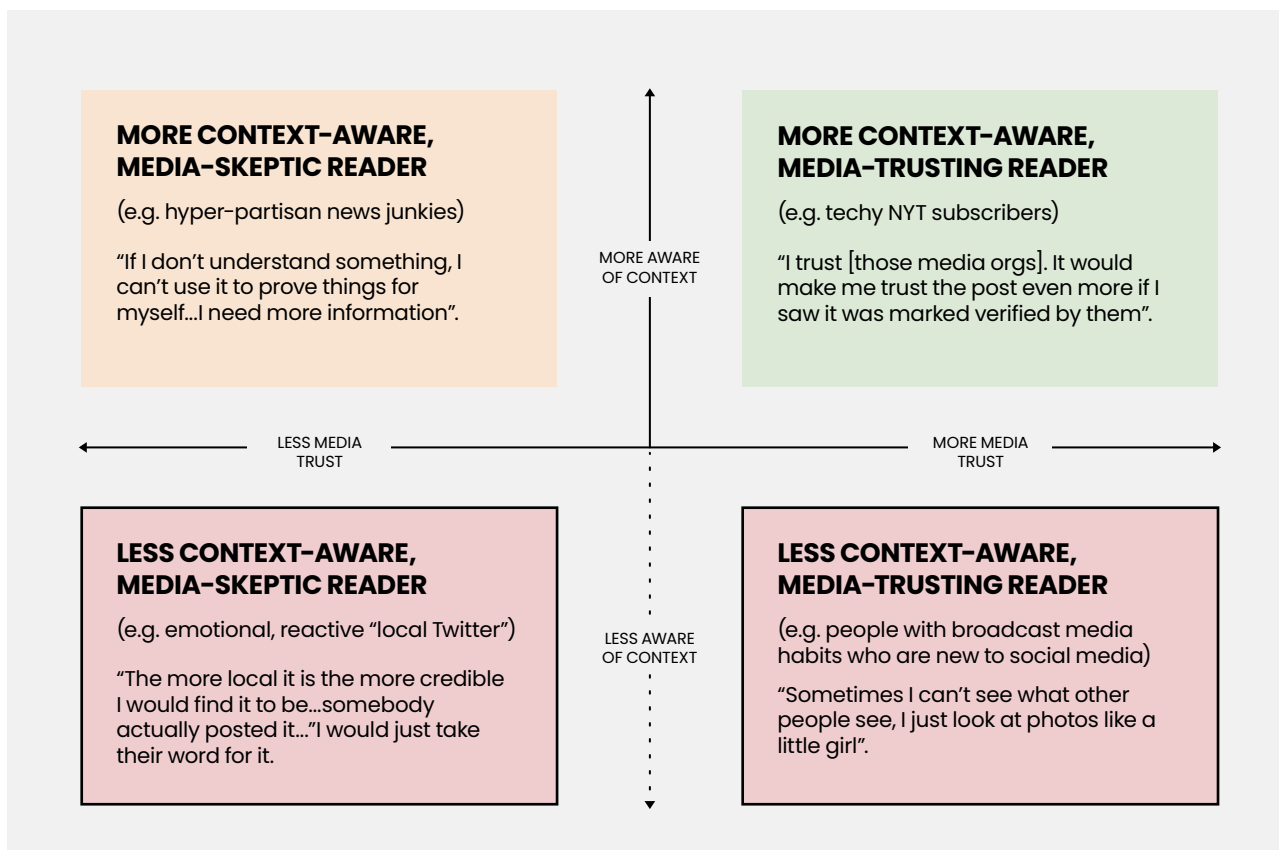
### Cutting-edge technology and news provenance

Traditional media can use cutting-edge technology to track down health disinformation and misinformation. Some news outlets are currently developing solutions based on blockchain technology that might track down the provenance of some pieces of content, allowing different stakeholders to access the origins of the disinformation.

For example, the New York Times (NYT) has developed an innovative project based on blockchain technology called the News Provenance Project (150). Marc Lavalee was one of the experts to present their work at the industry meeting, introducing the work of the NYT in providing credentials for information that travels around the Internet so that users can quickly and easily assess where it has come from. The News Provenance project stemmed from the observation that legitimate media is commonly used in misleading contexts: recycled, unsourced or modified. The paper takes the view that "knowing the origin and authenticity of information is a human right and the cornerstone of re-establishing trust on the Internet". They also believe that content publishers play a critical role in the information ecosystem.

According to the NYT, users can be segmented into four groups using two axes: those who are more or less trusting of mainstream media; and those who are more or less aware of veracity cues such as source and date (Fig. 11 below). Those who are less aware are likely to take content at face value. Those with low levels of trust are likely to be sceptical about all media institutions.

**Fig. 11.** Segmentation of users along trust and awareness axes

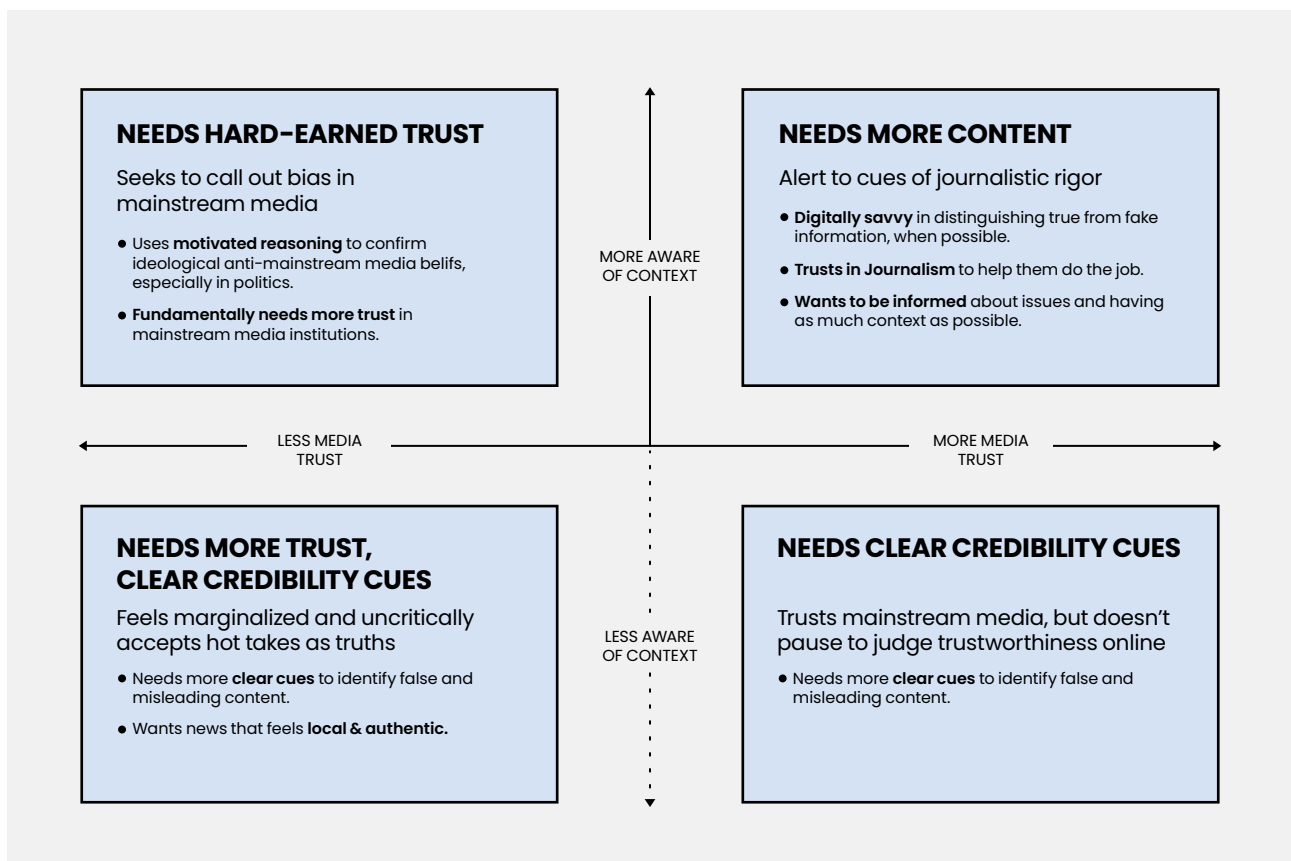


Source: New York Times.

Interestingly, NYT has found out that older users are seen to have higher levels of trust and lower awareness of cues. This is likely because the older generation grew up with trustworthy major broadcasters providing news, whereas younger users face competing sources online. This means that, to fight disinformation, different approaches are required to help users in each of the four groups to appraise content.

Based on this research, the News Provenance project team has developed several insights for credentialing visual content that answer the needs of those four groups (Fig. 12).

**Fig. 12.** Strategies for users in each of the four groups



Source: New York Times.

These insights are then transformed into working measures.

- Assess visuals for source information at the time that it is uploaded. This ensures that there is clarity about the source from the very first time that the content is viewed.
- Ensure that prompts induce a more critical mindset. Instead of flagging items as false, prompt users to "check for yourself – what does this photo show?" This is about introducing speedbumps rather than stop signs, as an unintended consequence of the latter can be making people curious, leading to the further propagation of misinformation.
- Highlight information that users can interpret for themselves.
- Provide multiple visual perspectives (i.e. include multiple photos of an event to help users build a stronger sense of what happened).



- Source editorial history from multiple publishers for a wider perspective: give access to the workflow from sourcing to publication.
- Use provenance to emphasize what is known, without discrediting all photographs that lack provenance information. This is to avoid disadvantaging local and small-scale news groups that do not have the resources to do the same.

Likewise, the Content Authenticity Initiative (151) – a collaboration between Adobe, Twitter and NYT to help creative tools like Photoshop to retain metadata and establish forward provenance for publishers and platforms to build a chain of custody and trust – has been working towards the creation and retention of metadata. This could result in a significant leap forward in the fight against health disinformation, as it would be possible to know the origins and authors of a given piece of false viral content.

Finally, Project Origin (152) is another initiative aiming at verifying the original source of content. This allows end users to see whether content has been altered or manipulated, which can help to identify deepfakes. In this work the United Kingdom's British Broadcasting Corporation (BBC) has partnered with Microsoft, the NYT and the Canadian Broadcasting Corporation to build a registry for platforms to perform two-factor authentication for media.

### *Standard-setting*

Many news outlets have adopted editorial guidelines that ensure their journalists actively avoid spreading disinformation on their networks. For example, the BBC has its own editorial guidelines (153) standards, with processes, policies, frameworks and checklists that journalists use to check the reliability of data and stories' reliability and whether presenting the information serves the public interest. Le Monde does the same through Les Décodeurs, its fact-checking system, explaining how its dedicated debunking team operates in fact-checking and providing reliable information (154).

Reuters has also adopted specific standards on the matter and a detailed methodology on how fact-checking should be conducted (155). Following Reuters classic Trust Principles, the fact-checking methodology follows several steps: (i) identify the content which is worthy of being fact-checked; (ii) summarize the key arguments of the content in question; and (iii) consult experts and look for evidence that supports or denies the claims made, naming sources and links which help to provide context to the debunking.

### *Fact-checking and debunking*

Another effective tool against health disinformation in traditional media is to have dedicated teams of debunkers. These are teams of journalists whose main task is to debunk conspiracies, flag misinformation campaigns or clarify the facts.

The French newspaper Le Monde, for example, has a dedicated team entitled Les Décodeurs, that constantly patrols websites on the Internet finding disinformation and misinformation campaigns (156). This team then produces reports and explanation pieces, in which myths are revisited and debunked. Moreover, Le Monde has developed its own fact-checking search engine, Decodex, where users can introduce websites and check

whether the information available is trustworthy. The BBC in the United Kingdom has a similar team, entitled Reality Check (157). The team uses a checklist to ensure that its news does not inadvertently amplify misinformation. Moreover, for more extreme cases, BBC has another team that looks at disinformation practices, especially those made to confuse and destabilize its audiences (the so-called anti-disinformation unit).

Likewise, in Portugal, Polígrafo newspaper (158) is a novel initiative fully dedicated to combating disinformation and misinformation in Portuguese media. The newspaper regularly advertises on TV and produces fact-checking reports on content that has gone viral on social media. Through a dedicated team of fact-check journalists, the programme presents their conclusions weekly, debunking disinformation and misinformation during prime time.

Reuters has also a dedicated team and page to debunk and fact-check important pieces of information. Reuters Fact Check is a programme dedicated to analysing social media posts that have gone viral, testing them and re-establishing the truth on the facts (155).

### *Fact-checking networks*

Another important initiative taken by traditional media is to engage in large networks of journalists that allow exchange of information about existing disinformation or misinformation campaigns. Associations such as the IFCN (142), under the Poynter Institute, group together several news outlets around the world and empower them with guidelines, criteria and platforms to exchange information.

### *Media literacy*

News outlets have also actively engaged in transmitting knowledge and increasing media literacy among their readers/listeners/viewers.

For example, the BBC's Young Reporter (159) is a partnership with schools to encourage children to think critically about how news stories are produced. It helps young people to develop content creation skills, find out about careers in broadcasting and share their own stories. The organization has also created BBC iReporter, an online interactive game to help young people to increase their digital literacy (160).

Likewise, Thomson Reuters, in partnership with the National Association of Media Literacy and Education, has launched a guide to identifying misinformation concerning the COVID-19 pandemic (161). The idea is to provide users with easy access to information that can help them make a conscious decision on existing pieces of information. The media outlet also recommends quick tips on spotting misinformation in your everyday interactions, as well as a teaching literacy tool for younger generations (162).

More than debunking, digital literacy empowers users to make their own decisions regarding content they see, turning readers and users into active players in the fight against disinformation and misinformation.

## Civil society

### Media literacy

#### *Type of initiatives*

Human-centred media literacy solutions often face scepticism about their effectiveness of the disinformation challenge. Despite evidence of the positive impact of information and media literacy programming on resilience to false and manipulative information and narratives, some policy-makers and donors favour technological or regulatory approaches to the disinformation crisis. In reality, technological, regulatory or human-centred approaches do not solve the problem alone; the most effective solutions require all three in cooperation.

Media literacy, unfortunately, is becoming politicized in some contexts. For example, in the United States of America, media literacy advocates report that just as fake news and disinformation have become political lightning rods, so has media literacy, as it is often perceived as a partisan issue. This could mean even further informational marginalization, deeper rabbit holes and wider conspiracy theory traps for those who reject it as an educational opportunity.

As with all education-based responses, continued funding and support over time is needed to sustain individual behaviour change and community information engagement norms towards becoming more empathy driven, responsible and critical. Media literacy initiatives are long-term solutions to build resilience against misinformation, and often struggle to stand out in the middle of alternatives promising more immediate results.

Policy changes that advance media literacy, cyber citizenship or healthy information engagement practices would benefit the goal of tackling misinformation. Specifically, the adoption of policies advocating for the integration of media literacy instruction in schools would be highly beneficial, as this would help to reverse the trend of deepening inequity in access to critical thinking and media literacy skills and subsequent vulnerability to manipulative information.

Media literacy initiatives are not just for schoolchildren. They must be a community-wide endeavour if the problem of global misinformation is to be counteracted at all levels. Further, it is important that when policy-makers make decisions, especially those that impact or require alignment with educators, those individuals are part of the conversation. We have learned during our meetings that a disconnect between policy-makers and the individuals who will work towards implementation at the ground level can sometimes jeopardize the success of the projects.

The lasting impression of the participants of the Forum is that media literacy is the least problematic solution against misinformation: one that does not involve censorship, platform regulation, content moderation or hierarchization of content. Media literacy targets audiences, inoculating them against the misinformation virus by providing them with skills and tools to make better information choices. Like a vaccine, it will not always work, and it will not end the problem; however, it has the potential to significantly reduce the spread and the impact of misinformation in our communities. It is, notwithstanding, a herculean solution to implement, requiring a lot of investment, coordination between government and the people on the ground, and particular focus on vulnerable communities (163).

With the COVID-19 infodemic, many media literacy projects gained a health component that should be supported and carried over to other health topics, such as NCDs. Health literacy related to NCD risk factors is already, to a certain extent, part of school curricula in certain countries. Some projects of the civil society have the potential to take government efforts to another level of impact (Case studies 1 and 2).

## Case study 1.

### Learn to Discern

One of the projects discussed in the high-level Civil Society meeting was the Learn to Discern (L2D) project, promoted by the International Research and Exchanges Board (IREX), an international, non-profit-making organization that specializes in global education and development. The project is in more than 20 countries globally and targets the public, ranging from secondary-age schoolchildren to adults.

L2D is an approach that recognizes the roots of vulnerability to disinformation and hate speech in human "operational systems" and in the incentives embedded into the social media infrastructure. The goal of the project is to equip those who consume information with critical thinking skills to navigate the polluted information space in a healthy, responsible and empathy-driven way. Since its original impactful performance in Ukraine where it reached over 90 000 people in nine months, L2D has created a long-lasting ability to recognize disinformation among participants; a year and a half after training, participants continued to be 25% more likely to check multiple news sources and 13% more likely to discern between disinformation and a piece of objective reporting (164).

L2D has since then been co-adapted with local partners and audiences in over 20 countries, most of which are conflict fragile, where it builds communities' resilience to state-sponsored disinformation, inoculates communities against public health misinformation, promotes inclusive communities by empowering its members to recognize and reject divisive narratives and hate speech, improves young people's ability to navigate increasingly polluted online spaces and enables leaders to shape decisions based on facts and quality information. Globally, L2D initiatives today include peer-to-peer and play-based online training led by young people in Georgia, Jordan, Serbia and Tunisia; it is integrated into education systems in Ukraine through secondary school and teacher training and in Estonia, Latvia, Lithuania and Jordan through higher education it provides effective locally trusted and influencer-led training for citizens of all ages in Jordan, Montenegro, Sri Lanka and Ukraine; and social media and online game/interactive content platforms in Albania, Indonesia, Jordan, Serbia and Ukraine.

Across these diverse geographies, contexts and participant demographics, L2D programmes demonstrate positive impact.

For example, in Jordan, through IREX's young people-led peer-training model for L2D skill-building, young people improved their skills and abilities to analyse information in their traditional and social media streams by 97% and their confidence and sense of control in navigating these spaces by 41%. In Serbia, young people participating in a similar training model improved their information analysis and evaluation skills by 43% and their sense of control over their information environments by 17%. In Ukraine, IREX worked with the Ministry of Education to integrate L2D competencies into the instruction of history, language and other subjects in 420 schools and among more than 50 000 high school students, who increased their ability to analyse and engage with media and information critically by 29%. Learners who have used IREX's online course on media literacy, Very Verified, supplemented with practical workshops led by a facilitator increased their awareness and knowledge of media and information structures by 69%, their ability to analyse and evaluate information by 31% and more than doubled their scores on an assessment

testing their ability to identify hate speech. A 2020 randomized control trial conducted by the RAND Corporation tested IREX's media literacy social media materials in the United States and found them to be effective in reducing engagement with foreign government propaganda among even the most partisan news consumers.

Impact measurement of IREX's L2D programmes varies but generally includes an analysis of skills learned and mastered, including skills such as the ability to analyse and evaluate information, the ability to identify hate speech and/or disinformation, and the likelihood of checking multiple news sources.

The L2D project is funded by donors such as the Joint United Nations Programme on HIV/AIDS, the United States Department of State and the United Kingdom Foreign, Commonwealth & Development Office, as well as other governments, such as those of Canada and the Netherlands. Their primary funding mechanisms include grants and cooperative agreements.

## Case study 2.

### **International Media Literacy Research Symposium**

The International Media Literacy Research Symposium is a forum to bring together new and established researchers from all areas studying media literacy education around the world, targeting international researchers, practitioners, community leaders and parents.

The Symposium is sponsored by non-profit-making organizations such as the International Council for Media Literacy, universities such as Sacred Heart University, Emerson College and Fairfield University and other donors from various community entities.

Founded in 2013, it aims to shorten the present media literacy gap by filling it with works from current scholars, new researchers, graduate students, educators and others from all over the world who have a vested interest in opening this field and moving it forward.

The issues covered are wide-ranging, examining through both a current and historical lens misinformation and disinformation in all areas including news, health and digital technologies and their influence on individuals. The most recent focus is on the process of filtering information through algorithms and understanding the impact of this, as well as the role of AI in how information is consumed.

The Symposium promotes the idea that media literacy education in all sectors, from education, health or business, is fundamental to society. It highlights the room for growth in research in this area and the importance of providing funding to projects that research and promote media literacy solutions, as well as to projects providing training to communities.



## *Source credibility schemes and labelling*

### *Type of initiatives*

Audiences are increasingly searching for health information online, but the multiplicity of sources of varying accuracy makes it increasingly difficult to discern who should be trusted.

There are certain heuristics that are commonly used by individuals when deciding whether or not a source is credible enough to be trusted: the reputation of the source or of the person sharing that source, familiarity with the source or with the content, or a confirmation bias towards the conclusions presented by the content, among others. In an ocean of sources it is likely that audiences will not be able to effectively assess most of them.

To solve this problem, some third parties (both corporate and non-profit-making) have created solutions that indicate to audiences whether a given source of online health information is credible. These solutions can be referred to as source credibility labels or schemes. As seen above, these are also being acknowledged by policy-makers as one solution to be implemented to help to curb the spread of misinformation.

Existing research focuses on how people assess the credibility of a source. However, there is room to learn more about how impactful credibility labels are, particularly when applied to online health information, in producing a reaction that counters misinformation spreading. This research is much needed and should be encouraged.

Notwithstanding, interesting initiatives are taking place in this field and there are growing calls for national governments and health departments to get involved and endorse schemes that independently verify producers of trusted health information. Similar calls suggest that major digital and social media platforms should recognize these solutions and prioritize the content of certified organizations, which would be a major algorithmic change away from virality (Case studies 3 and 4).

## Case study 3.

### NewsGuard

NewsGuard is a company created by journalists and focused on providing human-curated news reliability ratings to mitigate false news. They do so via a browser extension that identifies when a user visits a website that has been rated by NewsGuard journalists and identifies using a small green or red label whether the website is considered credible or not; it also has labels for satirical sources or platform sources.

This website score is achieved by reviewing it using nine criteria pertaining to credibility and transparency.

Credibility:

- does not repeatedly publish false content
- gathers and presents information responsibly
- regularly corrects or clarifies errors
- handles the difference between news and opinion responsibly
- avoids deceptive headlines.

Transparency:

- website discloses ownership and financing
- clearly labels advertising
- reveals who is in charge, including possible conflicts of interest
- provides the names of content creators, along with either contact or biographical information.

These scores are unobtrusively displayed alongside websites and allow users to explore the ratings in the form of a nutritional label. The labelling approach obviates many of the problems that come with content moderation and other forms of content removal approaches. By providing a journalist-made review about the information source, NewsGuard provides users with information for them to make a better decision, ultimately empowering them (Fig. 13).

**Fig. 13.** Newsguard ratings on a Facebook search page, with detail expanded

The image shows a Facebook search results page for the query "vaccines ods". The top result is a post from "cdc.gov" titled "Vaccine Effect Vaccines Work? | Influenza vaccine effe" with 6 shares. To its right is a Newsguard warning: "Proceed with caution: This website severely violates basic standards of credibility and transparency." Below this is a "Nutrition Label" for the "NaturalNews.com website network" with a score of 5/100. The label lists several issues under "CREDIBILITY" and "TRANSPARENCY".

CREDIBILITY	TRANSPARENCY
✗ Does not repeatedly publish false content	✗ Website discloses ownership and financing
✗ Gathers and presents information responsibly	✗ Clearly labels advertising
✗ Regularly corrects or clarifies errors	✓ Reveals who's in charge, including any possible conflicts of interest
✗ Handles the difference between news and opinion responsibly	✗ The site provides names of content creators, along with either contact or biographical information
✗ Avoids deceptive headlines	

Other search results include "CDC video PR rushed to market verification" and "CDC Expert: Vaccines Cause Autism".

Newsguard has created a health-specific tool, Healthguard, to provide the same service but focused on health information sources. Throughout the epidemic, Newsguard has been partnering with WHO by sending reports outlining the top COVID-19 and vaccination falsehoods circulating across major social platforms, the sources behind those claims and the groups that live on those claims.

Newsguard is a paid service (by monthly subscription). Healthguard was made freely available during the COVID-19 infodemic.

## Case study 4.

### PIF TICK

The project PIF TICK (Patient Information Forum Trusted Information Creator) is a non-profit-making source credibility scheme based in the United Kingdom and promoted by the Patient Information Forum (PIF).

It targets primarily organizations that create health information (from public, private and voluntary sectors), helping them to be better and more responsible health information broadcasters and, as a second impact layer, reaches patients, citizens and health-care professionals, helping them to recognize and share trustworthy health information.

PIF started working on the scheme in 2018 when the former NHS England Information Standard was discontinued. PIF is a membership organization representing information producers. Its members wanted a scheme to demonstrate the quality and trustworthiness of their information, and to ensure their processes would produce information products in line with best practice principles.

With their members, PIF created 10 key criteria to assess the trustworthiness of health information:

- information is created using a consistent and documented process
- staff are trained and supported to produce high-quality information
- information meets an identified consumer need
- information is based on reliable, up-to-date evidence
- patients are involved in the development of health information
- information is written in plain English
- print and digital information is easy to use and navigate
- users can give feedback on information
- information is promoted to make sure it reaches those who need it
- the impact of information is measured.

PIF then devised an assessment process which was piloted in 12 cross-sector organizations for six months across 2019 and 2020. Major charities including Cancer Research United Kingdom, Macmillan Cancer Support and Mind were involved in the pilot along with smaller charities, NHS Trusts and private companies. The insurer BUPA provided grant funding for the pilot.

The criteria were checked for suitability and common sense (sense-checked) with the public, who rated as crucial the criteria "evidence-based", "easy to use and understand" and "produced by trained staff". The public felt a tick was a simple symbol to represent the scheme. Evaluation of the pilot was undertaken, and members found the assessment process was rigorous and supportive; this helped to improve and maintain information production processes.

The process of certification starts with an online application by the organization that wants to be certified. After a first assessment, the organization's information creation practices are assessed against the criteria to prepare an action plan aiming to address any fragilities.

After that action plan is completed, the Tick is awarded and can be used by the organization on any content that is produced using the approved process.

PIF subsequently engages in annual assessments and spot checks to ensure that the process is maintained and provides training to members to build expertise. The organizations pay an annual membership fee scaled to the organization's turnover and type.

The goals of the PIF TICK project are to:

- develop a scheme to demonstrate the quality and trustworthiness of health information to patients and citizens;
- ensure information producers produce information products in line with best practice principles;
- give health-care professionals and other organizations confidence to signpost information produced by PIF TICK organizations;
- raise public awareness of the scheme, build health literacy and contribute to wider media and information literacy; and
- develop a self-financing scheme which is equitable to access to organizations of all sizes in all sectors.

There are certain challenges in building a credibility scheme such as PIF TICK. Devising a flexible scheme that meets the needs of small charities and large corporations and be financially secure is a primary challenge; another is ensuring a transparent and robust review process for the criteria to ensure they align to best practice in a rapidly evolving field.

To raise public awareness of the scheme on a small budget, PIF relies on the reach of member organizations and partner with other organizations to ensure that trusted health information is signposted by professionals.

Recruitment of new assessors with the correct skill set and implementing knowledge management systems to allow for the scaled growth of the scheme has also been difficult.

Finally, language is the main barrier to PIF operating the service in other WHO European Member States, although the model could be franchised to local operations.

As of January 2022, the scheme had 100 members with more waiting to join. The PIF TICK trust mark is currently appearing on a diverse range of health information published in all formats and is being used as a requirement by a growing number of information signposting systems such as the Information Prescriptions used by Cognitant Group with NHS patients. PIF has joined the Media and Information Literacy Alliance to further work in this area.

## AI

### *Type of initiatives*

AI is a powerful technology with tremendous potential. It can be used to produce hyperrealist misinformation, such as the videos popularly known as deepfakes that have circulated to showcase the potential danger of this technology. It is also highly susceptible to abuse by misinformation actors, that can take advantage of its mechanics to supercharge their efforts.

However, AI can also be used to tackle the spread of misinformation online.

Some tech companies are working on ways of automatically detecting misinformation through natural language processing, machine learning and network analysis. The goal is that an algorithm would recognize content as misinformation and rank it lower so that audiences are less likely to come across it. Regular exposure to the same information makes it more likely for someone to trust it, so a tool that recognizes misinformation and commands an algorithm to stop showing content that reinforces it could be effective.

This does not come without challenges, however; a lot goes into defining what constitutes misinformation, how to identify it, and how to tell it apart from factual information and build this knowledge into a tool that does not flag legitimate information as false and thus demote it. The unreliability of this evaluation raises concerns, but it is a use of the technology worth perfecting and pursuing for the potential it shows in curbing the spread of false information online. Pairing AI tools with human intervention may be a way forward (Case study 5).

## Case study 5. **Logically.ai**

Logically.ai was founded in 2017 with the purpose of protecting democratic debate and process and provide access to trustworthy information. To attain that goal there is a suite of products and services to reduce and eventually eliminate the harm caused by the spread of misinformation and targeted disinformation campaigns.

With a team of over 120 data scientists, engineers, analysts, developers and investigators, Logically works with governments and organizations around the world in the fields of national security, public safety, election integrity and public health.

It combines cutting edge AI and human expertise to apply both scale and nuance to the problem of misinformation and disinformation, identifying issues before they become widespread.

Logically's AI technology monitors, identifies and disarms problematic content at scale, and highly trained human fact-checkers and investigators provide complex research and analysis.

While human analysts and fact-checkers help to continuously improve the algorithms and check they are working as intended, AI enables the scalability, timeliness, efficiency and consistency of Logically's expert fact-checkers.

Logically.ai is an example of combining several ways of dealing with information; it shows that AI can be used to power many of the other solutions and approaches presented in this Toolkit.

One example is the platform Logically Intelligence, a threat intelligence platform that brings together Logically's capabilities in at-scale analysis, classification and detection to help governments and organizations monitor the online media landscape for the spread of damaging activity and narratives.

Another is the Logically app, a news platform sourcing 65 000 publishers into one streamlined news feed of fact-checked content. The technology clusters and prioritizes articles for the user, condensing the multitude of content available online into a news feed with objective headlines designed to inform, not sensationalize.

Finally, the Logically Browser Extension is a source credibility label that warns users of unreliable articles and sources while contextualizing the news, with reliable articles from over 100 000 publishers. It identifies toxic commentary on social media and highlights or hides it from the feeds based on the personal preferences of the user. Through the extension, it also facilitates access to Logically's express fact-checking service to help users to assess any suspicious claims.



## *Activism and campaigns*

### *Type of initiatives*

The driving power of civil society organizations comes from their ability to work alongside affected communities, reflecting the voice of end users and striving to meet their needs.

The need to work with diverse online and geographical communities to understand their needs, communication preferences and then to co-develop solutions with them is a fundamental step in the fight against misinformation. Given the vast number of different communities, no one public health organization could ever hope to work with every community. Quite simply, this task is impossible without civil society groups that are engaged and motivated to spread the word, do the groundwork and raise awareness. Acknowledging their role is crucial.

Participants of the WHO Forum noted that organizations like WHO should not seek to get its own messaging to all people directly, but provide solid information that local groups and communities can then "repackage" to share in appropriate ways for locals and members of the online tribe. NGO representatives told the Forum that the aim should be to engage and empower communities and that one should expect to see the best results when community leaders share information with their peers.

Activists and campaigners are often at odds with industry players and governmental institutions, but there is a clear agreement when it comes to misinformation and specifically health misinformation: a better-educated and informed audience will be more resilient against false information.

Activists have proved essential in filling in accountability and supervision gaps, by tracking and exposing online misinformation and monitoring the effectiveness of measures implemented by governments and industry.

The role of activists and campaigners, both online and offline, in raising awareness and creating engagement is, therefore, essential to generate enthusiasm, attention and critical thinking (Case study 6).

## Case study 6.

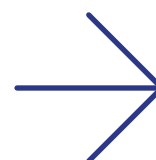
### Avaaz

The Avaaz Foundation runs an anti-misinformation project globally, with a focus on Brazil, the EU, and the United States. Avaaz is a global civic movement and advocacy organization with over 69 million members worldwide. Avaaz is 100% funded by members, and, therefore, only accountable to its members and not to major donors or foundations; it does not receive money from corporations or governments.

The targets of Avaaz initiatives are policy-makers, regulators, citizens and big technology leaders. Having started in 2018, the aim of the project is to uncover the scale of the disinformation problem on social media platforms as well as how it is adversely affecting public health, democracies, societies and (action on) climate change. Additionally, the aim of the project is to develop and propose effective policy and legislative solutions to big technology platforms and law- and policy-makers in order to pass legislation in key jurisdictions. Avaaz has developed a set of policy proposals, also serving as legislative principles, to effectively and efficiently tackle the disinformation crisis on online platforms in a human rights-based way.

As a result of this project, Avaaz has become a global leader in disinformation research and policy-making, including law-making. Their team of researchers has published ground-breaking reports using innovative methodologies, some of which are listed here.

- The scale of the COVID-19 disinformation problem on Facebook and how Facebook can flatten the curve of the COVID-19 infodemic. Its research found that the pieces of COVID-19 misinformation content sampled and analysed in early 2020 were shared over 1.7 million times on Facebook and viewed an estimated 117 million times – representing only the tip of the iceberg.
- How public health can be massively affected by Facebook's algorithm. In 2020 Avaaz uncovered health misinformation spreading networks with an estimated 3.8 billion views in the preceding year. Content from the top 10 websites spreading health misinformation had almost four times as many estimated views on Facebook as equivalent content from the websites of 10 leading health institutions, such as WHO and the Centers for Disease Control and Prevention. At the time the research was published, only 16% of all health misinformation analysed had a warning label from Facebook.
- How fake news may be making us sick. Its research indicated that misinformation may be reducing vaccination rates in Brazil. This report was the result of a joint study between Avaaz and the Brazilian Society of Immunizations, a member of the Vaccine Safety Net – the global network established by WHO.



- How health experts and scientists became targets of misinformation on social media; not only do false claims and online attacks, with their high interaction rates, have the potential to spill into offline violence (as in the case of Belgian virologist Marc van Ranst, who was forced into hiding with his family after receiving death threats, United States immunologist Anthony Fauci, who needed an armed security detail to protect him, or German virologist Christian Drosten, who received death threats), they also pose a challenge to global efforts to end the pandemic and could have a chilling effect on the wider scientific community.

These reports have been featured widely in the global media and, in combination with extensive advocacy efforts, have put significant pressure on the CEOs of big technology companies, law-makers, policy-makers and regulators to take action to protect people and societies from the harms of viral online disinformation. In part as a result of Avaaz research and advocacy efforts, online platforms like Facebook and YouTube have already taken some important steps to curb the spread of disinformation and inform users who have seen disinformation. Law-makers have also started to develop important regulations in this respect, but much more needs to be done.

## Fact-checkers

### *Type of initiatives*

Fact-checking is usually associated with industry initiatives lead by journalists or media organizations. However, there are also NGOs and other civil society forces working towards building tools and solutions based on fact-checking principles or aimed at improving the fact-checking that is done by the industry.

In topics where expert knowledge is needed (for example, health misinformation), civil society organizations with a remit focused in those areas are often better equipped to bring together experts and produce fact-checking content. What is lost by not following a journalistic method may be gained by using experts who master the topic and can contribute meaningfully to dispel myths and confirm facts.

These are also important complements for communities who are underserved by journalism, or live in places where freedom of the press is limited. In these contexts, the work of NGOs is particularly important, such as:

- GhanaFact, a project by FactSpace West Africa
- Teyit, an independent fact-checking platform based in Türkiye
- Tech4peace, a fact-checking social media platform in the Middle East
- PesaCheck, in East Africa
- Factnameh, in the Islamic Republic of Iran.

Case study 7 outlines the activities of Meedan, which builds software to support global journalism and information access.

## Case study 7. Meedan

Meedan is a technology non-profit-making organization that builds software and initiatives to strengthen global journalism, digital literacy and accessibility of information for the world. The Health Desk by Meedan's Digital Health Lab is a tool for journalists and fact-checkers.

Meedan's Digital Health Lab is an applied research initiative working towards equitable access to health information. The Digital Health Lab is focused on combining evidence-informed responses and rigorous user research to address digital health information inequity.

Over the last two years in response to the COVID-19 misinformation crisis, the Health Desk project has delivered on-demand, on-deadline explainers to journalists and fact-checkers in over 70 countries. With an audience of 2.5 million viewers, their public health experts received questions from fact-checkers and journalists covering health misinformation topics and responded with:

- an expert science explainer on the topic requested, with background and context about the subject;
- a set of glossary terms and definitions to understand the science in greater detail; and
- a recommended source list that the fact-checker can consult for more information.

The Meedan team has worked on questions shared by different organizations and individuals, distilled the research and translated it in fact-checked responses. They have partnered and worked with teams and organizations such as Africa Check, Berkman Klein Centre, India Today, Nigeria Health Watch, Reuters, Speak Up Africa, Students against COVID-19, Suno India and VeraFiles.

Their three research streams are:

- community-specific case studies for public health content moderation policies;
- public health information risk categorization; and
- cross-language text similarity analyses to evaluate gaps between requested health information and information available to communities.

By building a model based on syndication, where fact-checkers can use the content created directly for them and the content Meedan has developed in response to questions shared by other fact-checkers, health expertise can be provided at scale to reduce health misinformation more efficiently.

This project is carried forward by a team of scientists to produce research and evidence-informed responses that meets both the tone and the pace of newsrooms. This meant that the project management team had to set clear expectations for journalistic deadlines and schedules, as well as accessibility standards and training sessions to support the scientists through the initial launch of the production process.

## Case study 8.

### WHO infodemic education programmes

Since 2020, WHO has been organizing infodemic and infodemic management training programmes (165). Over 500 graduates across 120 countries have contributed to the global training programmes, the latest of which took place from 16 November to 9 December 2021, conducted online and cosponsored by the CDC, UNICEF and RCCE collective services.

The 3rd WHO training on infodemic management was a four-week programme designed to fulfil the needs of future infodemic managers, but also of the managers of infodemic managers. It emphasises the competencies described in the newly published competency framework for workforce response to infodemic management (166).

24 hours of live class sessions in total involved 35 lecturers across the four weeks, and trainees learned about topics such as:

- the emerging topics in infodemic management
- strategy development
- and policy implications in infodemic management.

Newly acquired skills were practiced in a simulation exercise. In groups, participants were virtually deployed in a prosperous fantasy country on the coast of the Narwhale Ocean, called the Kingdom of Great Wisdome, which had experienced a change in Ministry of Health leadership amidst growing pandemic fatigue and stagnating COVID-19 vaccine coverage and an infodemic that was running rampant.

Two previous infodemic management training courses in November 2020 and June 2021 operated on a similar basis, with participants learning about infodemic management skills that are needed to apply interventions, promote resilience of individuals and communities to infodemics, including misinformation, and to promote self-efficacy of individuals for self-protective health behaviours.

## Outcomes

By the end of the first course (167), infodemic managers were able to:

- measure and monitor the impact of infodemics during health emergencies;
- detect and understand spread and impact of infodemics;
- respond and deploy interventions that protect and mitigate the infodemic and its harmful effects;
- evaluate infodemic interventions and strengthen resilience of individuals and communities to infodemics; and
- promote the development, adaptation and application of tools for the management of infodemics.

Out of the 275 attendees, those that passed became part of the WHO roster of infodemic managers, supporting countries in infodemic management and response to health misinformation.

The second course continued to address the infodemic accompanying the COVID-19 pandemic, with participants from the first conference contributing from fields of speciality ranging across physics, law, behavioural science, epidemiology to user experience and design.

It is clear that a wide array of skills and expertise are required for infodemic managers to navigate the issues raised by infodemics, which go beyond the traditional boundaries of epidemiology, risk communication and community engagement and digital media. Through WHO and other parties such materials are increasingly available for free, and receiving feedback, improvement and development. One such example of a developed framework is the WHO competency framework: Building a response workforce to manage infodemics (166).

The infodemic education programmes also link into the WHO Information Network for Epidemics (EPI-WIN), which aims to make scientific information accessible, understandable and meaningful to all communities during emergencies so that their decisions, policies and actions are evidence-informed (168).



## The way forward: a triple entente against health disinformation and misinformation

The preceding chapters illustrate how the three levels of governance (governments, industry and civil society) are currently engaged in fighting the spread of health disinformation and misinformation.

First and foremost, governments have acted through different means (law, agreements, literacy, administrative agencies) and are actively fighting the spread of false information by placing renewed responsibilities on the many actors. Secondly, social media and traditional media – as natural vehicles of speech – have also, either explicitly or implicitly, acknowledged their role and responsibility in facilitating the creation, dissemination and amplification of misinformation. They have self-regulated in different ways, using technology to track unreliable information or to rank good information as more accessible and highly ranked. They have also developed their own internal structures to assess misinformation campaigns and committed to stopping the spread of damaging speech. Finally, civil society has played a vital role in ensuring that citizens/users remain aware of the perils of health disinformation/misinformation. Initiatives like fact-checking and debunking, allied to counter-disinformation campaigns, have allowed citizens to identify and refrain from spreading malicious information. Fig. 14 illustrates the scope of these measures.

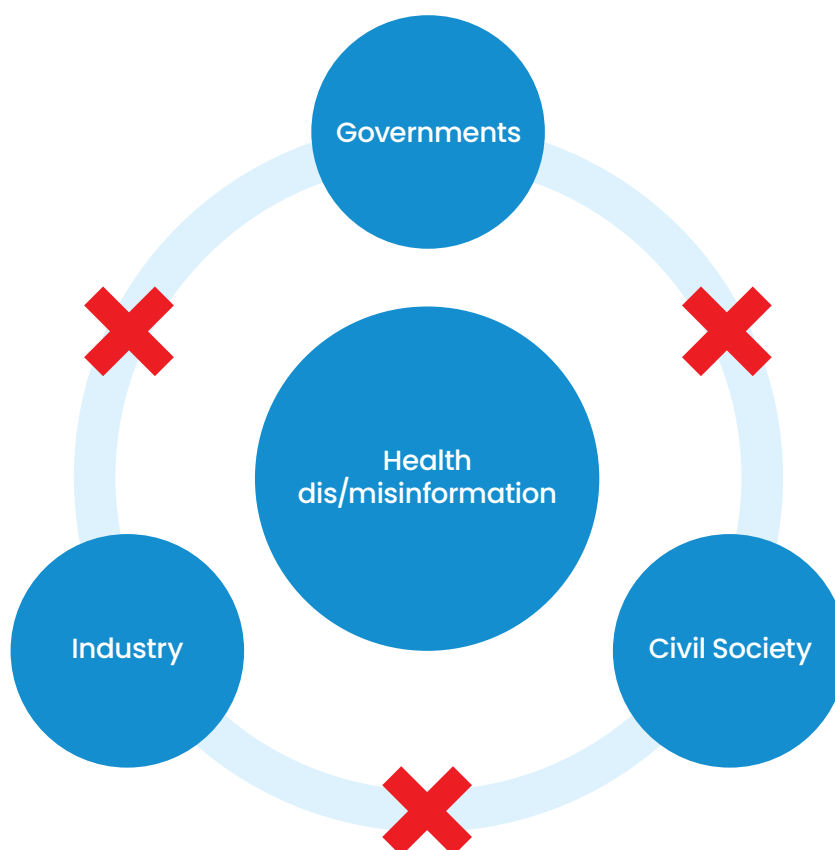
**Fig. 14.** Some of the measures implemented by the three levels of governance: governments, industry and civil society

Governments	Industry	Civil Society
<b>United Nations</b> <ul style="list-style-type: none"> <li>• Interagency cooperation</li> </ul>	<b>Media</b> <ul style="list-style-type: none"> <li>• Fact-checking tools</li> <li>• News literacy</li> </ul>	<b>Think Thanks</b> <ul style="list-style-type: none"> <li>• Forums of discussion</li> <li>• Academic writing</li> </ul>
<b>European Union</b> <ul style="list-style-type: none"> <li>• Rapid alert systems</li> <li>• European Digital Media Observatory</li> </ul>	<b>Social Media</b> <ul style="list-style-type: none"> <li>• Content labelling</li> <li>• Certification</li> <li>• Redirecting users towards reliable information</li> </ul>	<b>NGOs</b> <ul style="list-style-type: none"> <li>• Counter-misinformation campaigns</li> <li>• News literacy</li> </ul>
<b>Member States</b> <ul style="list-style-type: none"> <li>• Content Laws</li> <li>• Anti-fake news laws</li> <li>• Media/News literacy</li> </ul>		<b>Civil Society</b> <ul style="list-style-type: none"> <li>• Independent fact-checking</li> <li>• Online fact-checking tools</li> </ul>

All of these initiatives are laudable but can only go so far. The complexity of health disinformation and misinformation phenomena renders any unilateral action difficult to fully achieve its purposes. For example, while it is true that governments have attempted to regulate the matter via law or regulations, it has become evident that they would need to create incentives for platforms to improve their self-regulatory initiatives. Likewise, social media platforms rely on the input of civil society organisations and fact checking organisations, whose work provides incentives for them to further act on this problem. International organizations themselves – such as WHO – require the help of Member States to execute and spread reliable medical information: information which can only be obtained through the help of networks of scientists working and sharing knowledge around the world.

The current paradigm is then one in which all laudable efforts are not achieving optimal results given the lack of effective cooperation between the different levels of governance in the fight against disinformation and misinformation (Fig. 15).

**Fig. 15.** Three levels of governance lacking cooperation

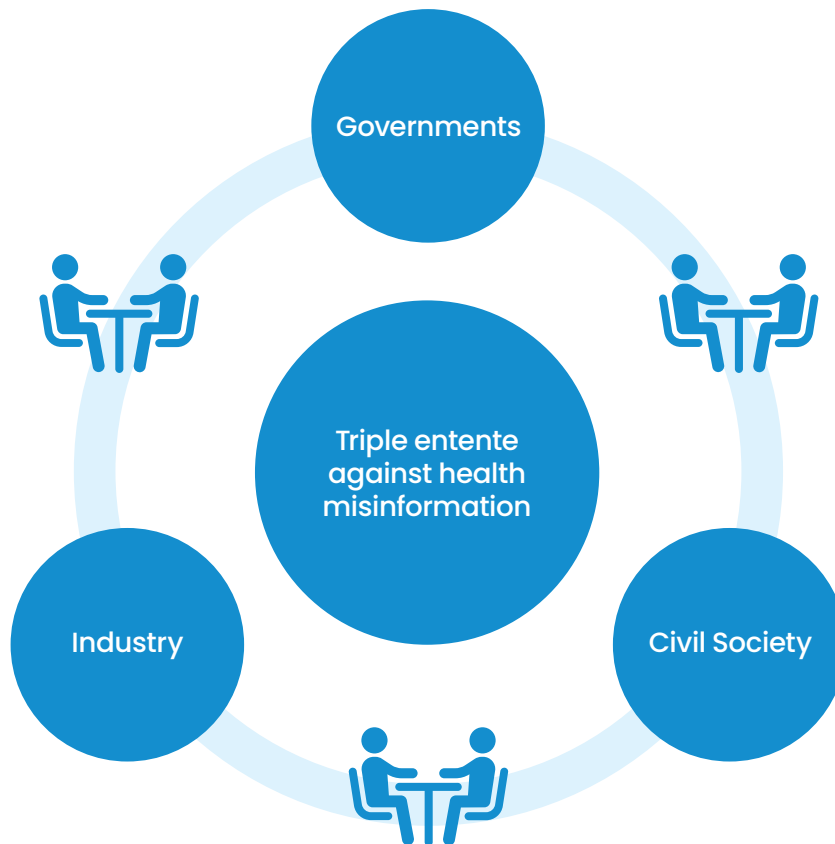


These approaches lack a fundamental element: effective cooperation. While each of the three levels of governance is actively working towards the same goal, it was not until the COVID-19 pandemic that growth in cooperation could be observed between the different levels. Health disinformation and misinformation have been previously tackled by these governance levels in a hermetic manner, considering the problem on platform X or on jurisdiction Y. These approaches miss the potential synergies between the different efforts, demanding repetition and duplication of efforts. The fact that sometimes different

stakeholders see each other as negative contributing forces for the problem may help to explain some resistance for extended cooperation, but it seems ultimately undeniable that no solution is complete without the buy-in from the three levels.

Hence, it is necessary to adopt an all-encompassing methodology in which the three levels of governance cooperate and take advantage of each other's potential. This report introduces the concept of a triple entente in the fight against disinformation and misinformation. Governments, industry and civil society must exchange information and collaborate exploring different venues and strategies of action (Fig. 16).

**Fig. 16.** Three levels of governance cooperating



This triple entente was briefly tried during the traumatic COVID-19 pandemic. However, learning from that difficult moment can help all levels of governance to harvest experiences and new synergies, contributing to long-lasting cooperation against health disinformation and misinformation.

## Building on the COVID-19 experience: a triple entente for the future

The COVID-19 pandemic was a moment of profound change in the fight against disinformation/misinformation. As the infodemic spread, the WHO Secretary-General stated: "We're not just fighting an epidemic; we're fighting an infodemic. Fake news spreads faster and more easily than this virus and is just as dangerous" (169).

This infodemic brought the different levels of governance finally together in the fight against disinformation and misinformation. This experience highlights the potential to cooperate in other areas of health disinformation/misinformation and serves as a critical test case to strengthening mutual exchanges between all levels of government. As announced in a joint statement by the International Federation of Red Cross and Red Crescent Societies, the Scientific and Cultural Organization, the International Telecommunication Union, the Joint United Nations Programme on HIV/AIDS, the United Nations, the United Nations Children's Fund, the United Nations Development Programme and United Nations Educational, the United Nations Global Pulse and WHO, the COVID-19 pandemic served as a call for cooperation:

We further call on all other stakeholders – including the media and social media platforms through which mis- and disinformation are disseminated, researchers and technologists who can design and build effective strategies and tools to respond to the infodemic, civil society leaders and influencers – to collaborate with the United Nations system, with Member States and with each other and to further strengthen their actions to disseminate accurate information and prevent the spread of mis- and disinformation.

Some examples highlight such collaboration and the potential of a triple entente.

### COVID-19 cooperation: government–industry

Governments engaged with social media during the COVID-19 and understood the fundamental role of online platforms in controlling the spread of health disinformation and misinformation.

Platforms such as Facebook, Google or Twitter welcomed the announcements of public authorities such as WHO or national governments. These platforms actually prioritize search results, redirecting people towards public sources of health information or even provided grants for public institutions to advertise at no cost on their platforms. This allowed public reliable information to expand to new audiences and for social media platforms to be used as channels of good information instead of hubs of disinformation.

Likewise, the BBC partnered with the United Kingdom Government to reach audiences beyond the United Kingdom, running the BBC World News and BBC.com COVID-19 campaigns. This was part of BBC Global News' commitment to donating free airtime to public health bodies and governmental organizations to promote messaging to combat the global coronavirus health crisis.

### **COVID-19 cooperation: industry–civil society**

Online platforms have also directly supported civil society efforts to fight COVID-19 disinformation by funding societal initiatives to control the spread of unreliable and malicious information. For example, Tik Tok partnered with several local and global organizations to share trusted information with the community. NGOs such as the American, British or Canadian Red Cross chapters have received financial help from the company to continue their health literacy works. Likewise, Tik Tok partnered with scientists from all over the world to incentivize the scientific community to post video updates on Tik Tok to show the world the progress being made on the vaccine (141). This team of public health experts answered all kinds of questions from our community, from what steps go into developing a vaccine to how they test them for safety – helping people to stay in the know.

The support to independent fact-checkers and the collaboration with entities such as the IFCN (142) and programmes such as Facebook’s Third-Party Fact-Checking Programme (143) are also examples of strict cooperation between different levels of governance against health disinformation and misinformation.

### **COVID-19 cooperation: international organizations–civil society**

International organizations and the EU have also directly exchanged efforts and best practices with civil society. For example, the EU created a learning centre for teachers worldwide to access relevant learning materials to share in class and develop health literacy. The same could be said of WHO, which provided citizens around the world with free courses on the facts around COVID-19 (170) and continued to improve its mythbusters around NCDs and other transmissible diseases (171).

### **COVID-19 cooperation: international organizations–government**

A crucial type of cooperation during COVID-19 was that of public authorities in different levels of public governance. The collaboration between international organizations and national governments during the pandemic proved essential for exchanging information about infected citizens, travel bans or the most successful measures used by other governments.

In that spirit, WHO provided a database compiling all of these experiences for Member States to access the different measures being implemented in different jurisdictions (172). The database included official reports commissioned by the Member States, documents in the public domain and other papers considered of scientific relevance for public authorities. This type of cooperation highlights the potential of exchanging information between different levels of governance, avoiding the duplication of efforts and ensuring full access to information by all players.

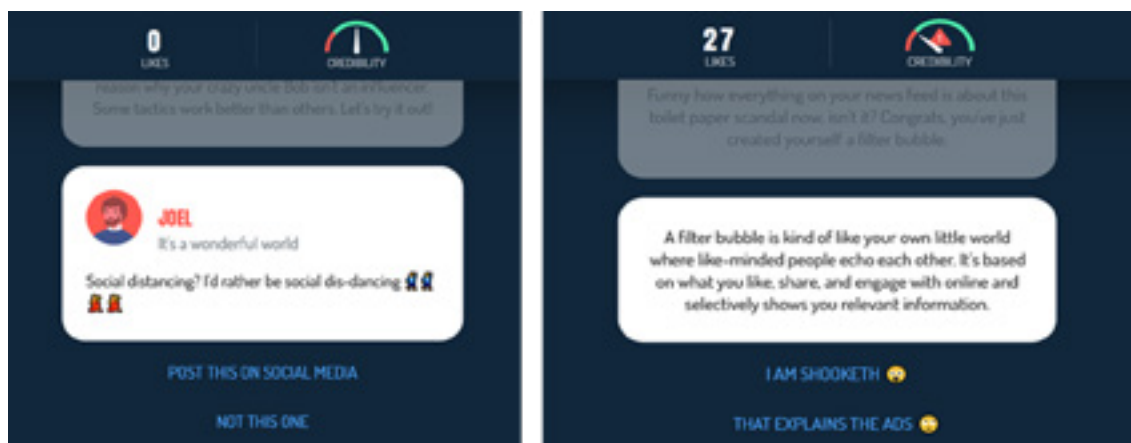
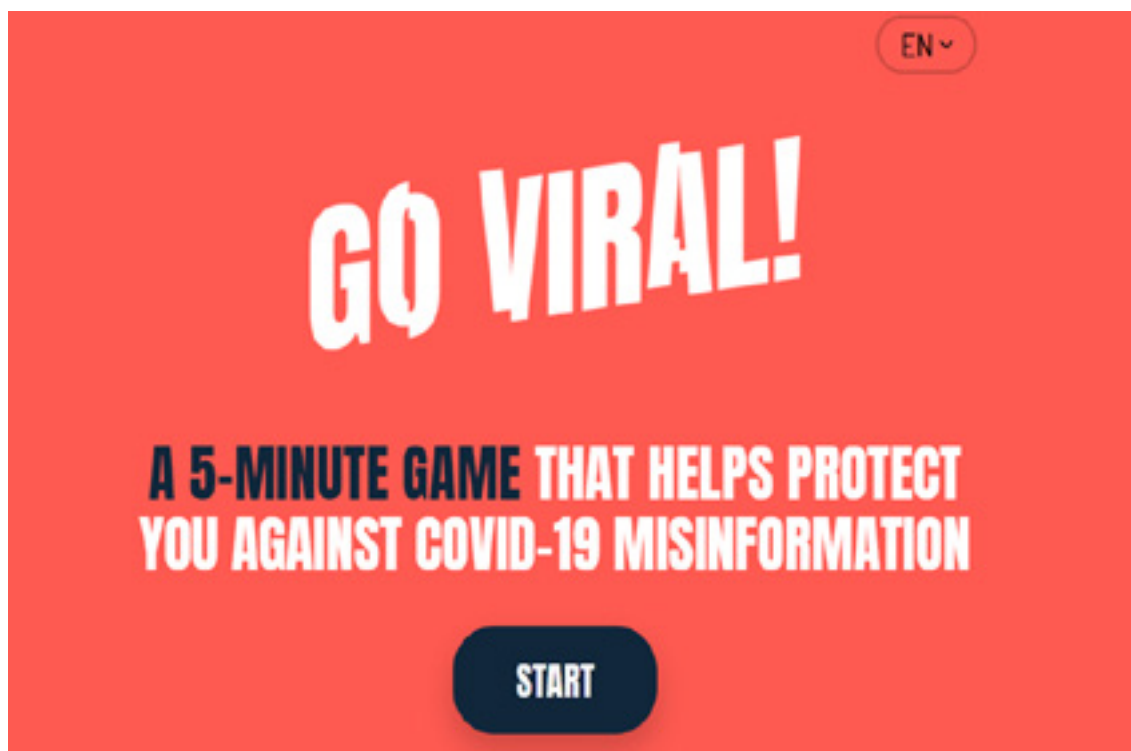
WHO also played an essential role in shaping national campaigns to build literacy for the COVID-19 pandemic. An example is how WHO cooperated with the United Kingdom Government to establish a campaign to spread the facts about the pandemic. The Stop the Spread campaign proved a success (173). The United Kingdom Government also offered a toolkit of the campaign assets to partner governments to translate and use in their countries, in another excellent example of governmental cooperation.

### COVID-19 cooperation: government–academia

Another excellent example of cooperation between different levels of governance was the collaboration between leading universities and governmental bodies across the world. A particular initiative by the name of Go Viral joined the United Kingdom Government and the University of Cambridge in developing a game that could increase the levels of health information literacy among young people.

The game aims at identifying the classic sources and practices of disinformation/misinformation through the user perspective in a social media context. The goal is to become the disinformant, playing the role of someone who spreads lies and misinformation. Once the students experiment such inverted roles, they fully understand the consequences of their actions (Fig. 17).

**Fig. 17.** Go Viral! online game



Source: University of Cambridge, 2021 (174).

## Conclusions and lessons learned

Based on the traumatic experiences of the COVID-19 pandemic, what can each of the three levels do to fight health disinformation and misinformation? The following comprise some practical suggestions to improve policy.

### The role of Member States in the triple entente

Member States play one of the most important roles in defining the terms of the cooperation between the different levels of the triple entente. Their motivation and initiative are crucial to gather the various stakeholders, providing a platform of discussion and exchange between public and private actors.

Member States highlighted some of the measures they had taken to combat health disinformation and misinformation, highlighting the need for further cooperation between all levels of governance. It became clear that the COVID-19 pandemic has had a great effect on accelerating existing initiatives and solidifying approaches.

#### *National practices*

A number of specific examples were provided to the meetings by Member States.

**Bosnia and Herzegovina** described the serious challenges around the speed and scale of the spread of misinformation during the pandemic. The Government is committed to the long process of helping the population to understand their role and responsibility in protecting their health and the health of others, focusing on traditional and social media channels.

**Bulgaria** held open national discussions and instituted new national media campaigns to counter prevailing myths, such as that 5G masts cause COVID-19. The Health Ministry shared WHO guidelines and set up a national hotline to provide reliable information around COVID-19 and vaccines. In July 2021 the Government started a new NCD health literacy campaign.

**Croatia** mounted multichannel campaigns to address health disinformation. Interestingly, the Government included the use of celebrities and influencers, as well as developing a new government website, to tackle the problem. The Government is looking to use these newly developed resources to tackle NCDs.

**Estonia** also developed a specific web page for COVID-19 information and launched social media campaigns that used celebrities to try to incentivize vaccination. During the meeting, the Estonian representative felt that a formal strategy was needed for addressing NCD misinformation, rather than using multiple but uncoordinated interventions.

**Montenegro** launched a new COVID-19 information campaign that included daily data briefings and a dedicated website. Dashboards on cases, hospitalizations and deaths were developed. Concerted efforts were made to engage with different media outlets to help the government to get its message out more effectively. New media content forms such as short videos were created to reach different population groups. New traditional printed materials were also produced. In terms of NCD response, there is already quite a lot of printed material available and NCD content is included in the national curriculum to try and improve health literacy among children.



**Portugal** developed a health information platform that has approximately one million followers. It produces simple, sharable infographics that have been widely disseminated on Instagram and Facebook. There has been success with a "true or false" concept where a fact is presented and users are invited to guess whether it is supported by scientific evidence. The evidence-informed answer is then presented in plain language. The health ministry has also established new relationships with social media platforms to identify and tackle misinformation. Portugal founded a new behavioural insights team and are co-leading a health literacy initiative with the Russian Federation. The representative noted that new resources are required to fund these tools and ongoing funding will be required to redirect these tools towards NCDs. Portugal has also started monitoring risk perception among the population and using these data to design communication strategies.

The **Republic of Moldova** has been issuing daily press releases from the minister of health outlining mortality and vaccination rates. The National Agency of Public Health has also been issuing regular press releases to counter misinformation. In terms of NCDs, two campaigns have been launched that use the endorsement of a popular national celebrity – focused on reducing trans-fat and sugar consumption, respectively. The Moldovan government has also dedicated time to strategizing about how to communicate with industry and is now involving industry representatives in a new dialogue. A web page has been set up for prevention of NCDs with the support of the Swiss Government and new NCD surveillance efforts have been instituted during the pandemic. A national campaign on sugar reduction is also being launched.

**Romania** tried to combat COVID-19 fake news by following WHO guidelines and using social media channels to promulgate reliable information. Trusted doctors and researchers were used as spokespeople. Dedicated web pages were used to answer questions raised by the public. Romania already has a national health strategy that includes a suite of NCD policies and campaigns but is interested in developing its response further.

**Slovakia** has used videos and social media to tackle COVID-19 misinformation; however the situation has been very difficult: spokespeople and celebrities have been harassed for supporting public health measures. There have also been large public demonstrations against masks and other evidence-informed measures. Furthermore, Government ministers have been reticent in their support of vaccination. Given the lack of a robust state response, the Slovakian representative was concerned that it will also be difficult to tackle NCD misinformation.

**Slovenia** has used multiple different approaches to countering COVID-19 misinformation, leaning on WHO and the Centers for Disease Control and Prevention for mythbusting materials, although translation has been a persistent issue. A new website has been developed to provide reliable information about COVID-19 and vaccination. New public health intelligence tools have been developed to monitor hospitalizations, vaccination status and vaccination uptake. New print and social media materials have been created and new partnerships have been developed to aid disseminations. New Tik Tok and Facebook profiles have been set up to share up-to-date information and work through frequently asked questions. These accounts will be used to address NCDs in the future.

**Sweden** has focused on transparency – publishing all official data and analyses – and made sure that public authorities are available to the public with daily press conferences in order to build public trust. Active listening exercises and collection of new data on public attitudes have been used to design tailored communication campaigns that use “fact and feeling” to motivate and inform people. There are no active plans to translate these new approaches to address NCDs, but they will inevitably inform future strategies.

### *National examples of utilizing the triple entente*

Case studies 9–11 illustrate interactions between government, industry and civil society, both success and problems.

## Case study 9.

### Belgium: disseminating evidence-informed health information for the public

A good example of cooperation between the different levels has been the Belgian experience in combating health disinformation. Through a governmental website called Gezondheid en Wetenschap, Belgium has been able to spread reliable health information to the public and patients as well as to debunk fake news in (social) media (175). The website is well respected and widely used by the public.

There are several transferable lessons to be learned from this case study.

- The site is hosted by the Centre for Evidence-Based Medicine as this is trusted more than other government institutions.
- Content is reviewed by a panel of patients to check that it is comprehensible before it is shared publicly.
- Health information posts are linked to current events in popular culture. For example, content on Lyme disease was released when pop singer Justin Bieber was diagnosed with this condition, accompanied by carefully designed content that “pimped up” the minister of health (Fig. 18).
- The Centre has hired a team to package information in appealing forms that work well online.
- The team has worked closely with WHO in the past.
- A full-time journalist makes short videos for Instagram and TikTok.

**Fig. 18.** An example of a post from the Health and Science centre



**Translation:**

We have an ambassador!  
And according to our statistics  
he is better than Justin Bieber.

Read more: What do Jo Vandeuren  
and Justin Beiber have in common

## Case study 10.

### Finland: prerequisites for countering industry

Another excellent example of multilevel work on combating health disinformation and misinformation is the work conducted by the Finnish Institute for Health and Welfare. (176) Finnish representatives shared Finland's experience in countering NCD misinformation from tobacco, alcohol and food companies during the meeting.

Finland used to have the highest rates of cardiovascular diseases in the world. The North Karelia project, established in 1972, has led to a marked decline in the use of butter and tobacco, and in associated cardiovascular deaths. Today, sugar taxation is being challenged in international courts. The tobacco industry is regrouping through the use of e-cigarettes. The alcohol industry remains very active, seeking to influence WHO recommendations and EU legislation.

Several lessons from the Finnish "fat wars" of the 1980s remain instructive.

- There was consensus across the medical community about what needed to be done.
- There was a political consensus around the matter.
- Businesses came on board when they realized that they could sell profitable new products such as vegetable oils.

The Finnish representatives also highlighted the role of trust in fighting health disinformation and misinformation. Similarly to Belgium, it became clear that it was better to transmit information via the Institute rather than through political leaderships, which often have significantly lower levels of trust.

## Case study II.

### Armenia: building resilient communities through public education

A final exceptional example of the role of Member States in the triple entente is that of Armenia and the work of its National Institute of Health.

The representatives from Armenia highlighted how tackling NCD risk factors constituted a priority for the Armenian Government and the role that disinformation and misinformation played in preventing such an objective from being achieved.

Armenia has been very active in producing and disseminating accurate, evidence-informed NCD information for the general population. The representatives strongly felt that building strong NCD health literacy levels can help to reduce susceptibility to misinformation.

NCD information has been disseminated using the following forms:

- posters
- blogs
- dedicated mythbusting campaigns
- websites
- dedicated social media accounts
- SMS reminders with collaboration from telecoms companies.

Interestingly, the Armenian Government deliberately focused on intersectoral collaboration and directed segmented messaging towards children, pregnant women and older people. The Government also translated and disseminated WHO guidelines on healthy behaviour, such as physical activity and salt consumption.

The representative from Armenia also shared two important insights learned over recent years: efforts to tackle misinformation should be science based, and it is important to continually inform the population so that they can critically appraise disinformation in any sphere. For example, when the general population already knows that salt consumption is high, they are more likely to accept new targeted measures to tackle high salt consumption.

By comparison, Armenia experienced the problems of disinformation being used to undermine an initiative to support healthy behaviour with regard to tobacco use. Following the introduction of new tobacco laws that constrained the advertising and sale of electronic cigarettes, there was a wave of industry-fuelled myths and disinformation, as well as coordinated campaigns to undermine the credentials of public health authorities. This included the use of industry-backed puppets to whitewash all adverse health effects with claims that e-cigarettes were harmless, or even a public good to be used to help people to quit smoking. The tobacco industry also rehashed common economic arguments about the harmful effects of restricting tobacco use on the broader economy.

## The role of industry in the triple entente

Industry plays a fundamental role in this triple entente. Through social and traditional media, most relevant health disinformation and misinformation campaigns circulate widely; these outlets are the vehicles and gatekeepers of information. It is vital that the industry realizes the potential of collaboration with other stakeholders and fosters open communication channels.

### Algorithms

A potential pathway towards improving the fight against disinformation is to revisit the role of algorithms. This type of technology often funnels users towards harmful misinformation. Social media platforms should be more transparent about the criteria used to rank content, including ensuring that harmful information is demoted, and that good reliable information is privileged. Although stakeholders agree that industry actors should not be the arbitrator of truth, they are in the best position to address the matter technically. As they control the digital architecture, they must constantly contact other stakeholders – including public authorities and the scientific community – and curate content according to their reliable information.

Some platforms, such as Twitter, already let users "turn the algorithm off". Users will soon be able to select their own algorithms to use when browsing content on different platforms imported from third-party providers. While this ameliorates the problem of scientifically agnostic, revenue-focused proprietary algorithms (designed to maximize advertisement exposure), people may choose to plug in new algorithms that further restrict their exposure to opposing or scientific views and subsequently reinforce their prejudices or erroneous beliefs. A balance must be struck between curating good and reliable information and the excessive clustering of individuals under bubbles of understanding.

### Targeting vulnerable groups

The industry should increase the number of languages that fact-checkers process and focus efforts on at-risk groups. One of the big problems of content moderation and the spread of health misinformation and disinformation is the fact that most content moderation structures are focused on a limited number of languages. This has caused problems in the past in countries such as Ethiopia, Kenya and Morocco (177), as well as in India after the Myanmar crisis in 2018.

The industry must expand its moderation structures to take into account underrepresented languages across the many communities, being closer to the populations and their health concerns.

## *Improving transparency*

Industry should share information on the scale and nature of the problem and the actions they have taken to address it. Public health groups must understand how algorithms work, and social media companies should be able to explain certain content decisions easily. This is currently a proprietary black box.

Transparency is needed, requiring social media platforms to provide comprehensive reports on disinformation and misinformation, measures taken against it, and the design, operation and impact of their curation algorithms. Platforms' algorithms must also be continually and independently audited based on clearly aligned key performance indicators to measure impact, prevent public harm and improve design and outcomes.

## *Media provenance*

Another important change that the industry can adopt to enhance cooperation is checking media provenance. This requires two different approaches, as defended by Marc Lavallee from the NYT. On the one hand, publishers will want to use schemes that encourage web traffic to their sites, while small publishers and content creators will need cheap, user-friendly plug-ins that help them to do the heavy lifting of certifying provenance. This will require cooperation with stakeholders who can certify information as truthful and companies developing such a technology. On the other hand, governments will play an important part in regulating what the entire process of provenance accreditation will look like. The creation of a right to know the "provenance of information" will require the efforts of all three levels of governance: regulating through law (government), the technical platforms to access the provenance (industry) and the role of individuals and civil society in helping to identify the origins of the pieces of information shared (civil society).

## *Certifying sources*

Whereas censorship restricts civil liberties and can lead to entrenchment in beliefs, transparent accreditation browser extensions can provide users with tools to understand the position of the underlying sources and alert users to content coming from untrustworthy sources.

The browser add-ons Newsguard and Healthguard are examples of such certification. These browser extensions are produced by a team of journalists who rate and review the reliability of health information from websites, using nine criteria pertaining to credibility and transparency. These scores are unobtrusively displayed alongside websites and allow users to explore ratings. This approach obviates blocking, censoring or restricting freedoms – in fact, it empowers users by providing transparent, independent assessment of credibility. It is a form of prebunking as it rates domains (at source) rather than individual articles (178).



### ***Synthesizing complex scientific evidence***

Another important dimension of the cooperation between industry and other stakeholders is the transmission of complex information in easily accessible ways. This requires strong cooperation between traditional media and public authorities. For example, during the meetings leading up to this Toolkit, journalists expressed their desire for public health agencies to produce clear summaries of complex topics to help content producers to represent current scientific knowledge accurately. The United Kingdom's Science Media Centre (179) provides an excellent example of an organization executing this function, and WHO might be able to play a role in this space.

In any event, it becomes clear that public authorities and industry must be in close contact to provide easily understandable health information to the citizens. A strengthening of the ties between governments and news agencies and social media platforms (e.g. a space provided for health information fact-checking; a column for NCDs mythbusters; air time dedicated to health information; social media pages dedicated to debunking health-related myths; and so on) will greatly improve the effectiveness of anti-disinformation/misinformation campaigns.

### ***Data sharing and intellectual property***

Another key dimension of the cooperation under a triple entente is the sharing of relevant scientific data between all levels of governance. Public health representatives often express a desire for media platforms to share data on the size of the problem, the workings of proprietary algorithms and the (confidential) business plans that collectively underpin the spread of misinformation. As this might collide with industry secrecy and proprietary software, the three levels of governance should open secure and confidential communication channels that allow exchanges of information between public authorities, platforms/media and researchers.

Access to data for research purposes is fundamental. Civil society can only meaningfully and adequately engage with health disinformation and misinformation if it has the relevant data to extract conclusions and design awareness campaigns. Likewise, public authorities must better understand the range of the disinformation problem in order to tackle it. As industry constitutes the media channels through which the phenomenon mostly occurs (social media/traditional media), it has a pivotal point in creating bridges between all relevant stakeholders.

### ***Increase the resilience of the ecosystem through literacy***

Although the fundamental commercial architecture of media platforms is often the most discussed of all platform dimensions, there is also added value in investing in joint multilevel media literacy initiatives. This can take the form of critical appraisal training and improving the transparency of sources.

This means adopting multilayered literacy campaigns in which public authorities design the content, social media and media transmit and amplify that content and civil society engages with communities on the ground. This sort of layered approach requires the cooperation of all three levels of governance with the same objective: to combat health disinformation.

## The role of civil society in the triple entente

Civil society has an important role in monitoring the activities of Internet platforms (search engines, social media platforms and sites that host health articles) and governments: holding these actors to account and working to improve their conduct in a way that makes it easier for users to find accurate and reliable health information.

A number of industry-oriented actions were raised during the meeting.

### *Help detoxifying algorithms*

Civil society should work with industry to boost ethical standards and promote the end of algorithmic agnosticism. Avaaz, one of the participants in the experts meeting, was particularly adamant about this. NGOs and citizens must help the industry to ensure that science-based health messages appear first when people search for health information. The algorithm should not further actively promote known disinformation and systematic disinformers (so-called super-spreaders). This can only be done through a systematic detox of the algorithms.

- All platforms should stop accelerating any content that has been debunked by independent fact-checkers, as well as all content from pages, groups or channels that systematically spread misinformation.
- Fact-checkers should help to identify the super-spreaders of misinformation and media/social media outlets must ensure that they are demonetized. When an actor is systematically posting content that fails fact-checking, the platforms must ban these actors from advertising or from benefiting monetarily from the content.
- NGOs and other associations must help industries and governments to inform users and keep them safe. Users should be informed through clear labels when viewing or interacting with content from actors who have been repeatedly and systematically spreading misinformation and be provided with links to reliable information from public authorities.

### *Correcting the record*

In Europe, 87% of citizens across France, Germany, Italy and Spain supported corrections and 76% supported social media platform regulation to address misinformation. An essential part of the fight against health disinformation is to debunk mistakes and promote effective corrections to every person who sees a piece of disinformation.

Here, both social media platforms and civil society can play a role together. NGOs and associations of fact-checkers will correct the record – relying on scientific evidence or information from public authorities – and transmit this information to online platforms. These platforms must then expand these corrections to all the viewers of the false information, showing them the right side of the health query.

This is a fundamental part of countering disinformation. Research commissioned by Avaaz and conducted by leading experts proves that providing corrections to social media users who have seen false or misleading information can decrease belief in disinformation by half. Multiple other peer-reviewed studies have demonstrated that effective corrections can reduce and even eliminate the effects of disinformation.

Correcting the record should be at least a five-step process.

- Define: the obligation to correct the record would be triggered when: (i) independent fact-checkers verify that content is false or misleading; (ii) a significant number of people (e.g. 10 000) have viewed the content.
- Detect: the obligation of platforms to provide an accessible and prominent mechanism for users to report misinformation: for example by giving independent fact-checkers access to health content that has reached a certain threshold such as 10 000 or more people.
- Verify: once content is flagged, media and social media outlets must work with independent, third-party verified fact-checkers to determine whether reported content is misinformation.
- Alert: each user exposed to verified misinformation should be notified using the platform's most visible and effective notification standard (e.g. a push notification).
- Correct: each user exposed to misinformation should receive a correction that is of at least equal prominence to the original content and that follows best practices which could include: (i) offering reasoned alternative explanations, keeping the users' worldview in mind; (ii) emphasizing factual information while avoiding, whenever possible, repeating the original misinformation; or (iii) securing endorsement by a media outlet or public figure the user is likely to trust.

### **Using bots for good**

Another type of cooperation between civil society and industry is the creation of "bots for good". Instead of simply assuming that bots only spread misinformation – as is often done with current platforms – civil society could start using the same technologies to promote reliable information. In fact, during the COVID-19 pandemic, Facebook Messenger implemented a WHO chatbot that could automatically answer queries about the disease (180). The use of this technology requires the AI system to be fed with reliable information from trustworthy sources. This means that civil society and industry must work closely to ensure that the information conveyed by these machines is the best possible available at the time.

### **Supporting academia research**

It is also vital that academia continues to contribute to decisively shape government and industry action. Research on the causes, the sources and the potential solutions for spread of misinformation is needed to optimize approaches and ensure that time, motivation and resources are spent in an efficient and effective way.

Of particular importance is to continue researching how to better empower users to protect themselves from misinformation. Media literacy research, curricula and trainings are particularly relevant in this context and should be supported by governments.

Also important is to promote solutions that do not pose as much threat to freedom of speech as approaches underpinned in deleting content and removing users from platforms. Alternatives like source credibility labels, providing users with more information to decide better, may be of particular importance. It is fundamental that research

thoroughly supports the efforts behind information labelling, finding evidence that such labels are effective and useful to audiences, that they impact the decision to spread misinformation and that they are done without disturbing the freedom and balance of the information ecosystem.

### *Joining the conversation*

Civil society must also actively engage with the existing conversations online. This means that NGOs, associations and individuals must make an effort to participate in comments sections, online debates or viral discussions on health matters. Here online platforms and newspapers/televisions can help rank reliable information higher, shaping the discussions and providing the necessary guidance on a specific topic such as alcohol queries or nutrition myths.

## Afterword

The purpose of the Forum was to converse with the main stakeholders in the health misinformation topic: those that are creating projects, tools and policies to tackle the issue in the multiple fronts in which it currently presents itself.

There is a pre-infodemic and a post-infodemic world in the fight against health misinformation. The scale of this problem and the potential danger to human lives has become apparent with the COVID-19 pandemic and forced stakeholders to seek more creative, effective and collaborative solutions.

This wave of enthusiasm and dialogue must be maintained beyond the COVID-19 infodemic and put to good use in the coming years towards other health misinformation topics, especially those related to NCD risk factors.

NCD risk factors are inherently attractive topics of misinformation: they relate to day-to-day practices like eating, exercising, smoking or drinking alcohol; they address commonly devastating diseases that cause violent emotions; they are the lingering, permanent misinformation topics that have been largely overlooked by government policies, digital platforms, traditional media and the scientific community.

We shall take the urgency of action felt during the pandemic infodemic should be carried forward to continue pushing for collaborative, comprehensive and coordinated solutions for other health issues such as NCDs.

The view of those at the three meetings was that the perils of health misinformation are visible and reasonably documented, and a tackling strategy will demand an "all hands on deck" approach. Public authorities, the industry (including news media and social media companies and platforms) and civil society (through organizations and NGOs and on an individual level) should be engaged simultaneously to act together as a triple entente in the fight against misinformation. The scientific community should work to provide them with the roadmaps and the evidence needed for that.

This Toolkit is aimed at providing an overview of the problem, the challenges and the many types of solutions and approaches. There was no ambition to be comprehensive, as there are thousands of wonderful projects in this field, with more emerging every month on global, national and community scales, in all languages.

The main goal was to provide options and pathways and show how media literacy, content moderation, AI, good journalism, fact-checking, credibility labels and other solutions can and should coexist and be supported to keep the people safe and ready to face the overabundance of health information that exists and will be created online in the next decades.

Many people contributed to this Toolkit and to the Forum discussions. Lists of participants can be found together with the Meeting Reports, but an acknowledgement and special "thank you" is given to those mentioned in the Acknowledgements for their role in putting together a document that it is hoped may guide policy-makers from the Member States of the WHO European Region in their challenge to curb the spread of health misinformation on NCDs.

## References<sup>1</sup>

- 1 Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med*. 2019;240:112552. doi: [10.1016/j.socscimed.2019.112552](https://doi.org/10.1016/j.socscimed.2019.112552).
- 2 Disinformation. In: Merriam-Webster dictionary [online database]. Springfield (MA): Merriam-Webster; 2022 (<https://www.merriam-webster.com/dictionary/disinformation>).
- 3 Health at a glance: Europe 2018: state of health in the EU cycle. Paris: Organisation for Economic Co-operation and Development; 2020 ([https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2018\\_health\\_glance\\_eur-2018-en](https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2018_health_glance_eur-2018-en)).
- 4 Acute health risks related to the presence of cyanogenic glycosides in raw apricot kernels and products derived from raw apricot kernels. Parma: European Food Safety Authority; 2016 (<https://www.efsa.europa.eu/en/efsajournal/pub/4424>).
- 5 DeStefano F, Price CS, Weintraub ES. Increasing exposure to antibody-stimulating proteins and polysaccharides in vaccines is not associated with risk of autism. *J Paediatr*. 2013 Aug;163(2):561–7. doi: [10.1016/j.jpeds.2013.02.001](https://doi.org/10.1016/j.jpeds.2013.02.001).
- 6 Is fake news making us sick? How misinformation may be reducing vaccination rates in Brazil. New York: Avaaz; 2019 (<https://avaazimages.avaaz.org/Avaaz%20-%20Is%20Fake%20News%20Making%20Us%20Sick%3F.pdf>).
- 7 WHO Director-General statement on the role of social media platforms in health information. Geneva: World Health Organization; 2019 (<https://www.who.int/news-room/detail/28-08-2019-who-director-general-statement-on-the-role-of-social-media-platforms-in-health-information>).
- 8 Understanding the infodemic and misinformation in the fight against COVID-19. Washington (DC): WHO Regional Office for the Americas; 2020 (<https://iris.paho.org/handle/10665.2/52052>).
- 9 Forum for Tackling Misinformation on Health and NCDs. Copenhagen: WHO Regional Office for Europe; 2022 (<https://www.who.int/europe/news/item/02-02-2021-forum-for-tackling-misinformation-on-health-and-ncds>).
- 10 Wardle C, Derakhshan H. Information disorder: toward an interdisciplinary framework for research and policy making. Strasbourg: Council of Europe; 2017 (<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>).

---

<sup>1</sup> All URLs were accessed on 16 June 2022.

- 11 On the European democracy action plan. Brussels: European Commission; 2020 (COM/2020/790 final, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A790%3AFIN&qid=1607079662423>).
- 12 McKernon E. Fake news and the public. by Edward. New York: Harpers Monthly; 1925: 528–536 (<https://harpers.org/archive/1925/10/fake-news-and-the-public/>).
- 13 Conspiracy theories and right-wing extremism: insights and recommendations for P/CVE, 2021. Brussels: European Commission; 2021 ([https://ec.europa.eu/home-affairs/networks/radicalisation-awareness-network-ran/publications/conspiracy-theories-and-right-wing-extremism-insights-and-recommendations-pcve-2021\\_en](https://ec.europa.eu/home-affairs/networks/radicalisation-awareness-network-ran/publications/conspiracy-theories-and-right-wing-extremism-insights-and-recommendations-pcve-2021_en)).
- 14 Douglas KM, Uscinski JE, Sutton RM, Cichocka A, Nefes T, Ang CS, Deravi F. Understanding conspiracy theories. *Polit Psychol.* 2019;40:3–35. doi: 10.1111/pops.12568.
- 15 Google sifting through one billion health questions each day. *The Telegraph.* 13 Mar 2019 (<https://www.telegraph.co.uk/technology/2019/03/10/google-sifting-one-billion-health-questions-day/>).
- 16 Jacobs W, Amuta AO, Jeon KC. Health information seeking in the digital age: an analysis of health information seeking behavior among US adults. *Cogent Soc Sci.* 2017;3 (1):1302785. doi: 10.1080/23311886.2017.1302785.
- 17 O’Dea S. Smartphone users worldwide 2016–2021. In: Statista [online database]. New York: Statista; 2022 (<https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/>).
- 18 Mansky J. The age-old problem of “fake news”. Washington (DC): Smithsonian Magazine; 2020 (<https://www.smithsonianmag.com/history/age-old-problem-fake-news-180968945/>).
- 19 Wu T. *The attention merchants: the epic scramble to get inside our heads.* London: Penguin Random House; 2016.
- 20 Zuboff S, Schwandt K. *The age of surveillance capitalism: the fight for a human future at the new frontier of power.* London: Profile Books; 2019.
- 21 Betsch C, Sachse K. Dr Jekyll or Mr Hyde? (How) the Internet influences vaccination decisions: recent evidence and tentative guidelines for online vaccine communication. *Vaccine.* 2012;30(25):3723–6. doi: 10.1016/j.vaccine.2012.03.078.
- 22 Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm: an overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine.* 2012;30(25):3778–89. doi: 10.1016/j.vaccine.2011.11.112.



- 23 Reyna VF. Risk perception and communication in vaccination decisions: a fuzzytrace theory approach. *Vaccine*. 2012;30(25):3790–7.  
[doi: 10.1016/j.vaccine.2011.11.070](https://doi.org/10.1016/j.vaccine.2011.11.070).
- 24 Xu Z, Guo H. Using text mining to compare online pro- and anti-vaccine headlines: word usage, sentiments, and online popularity. *Commun Stud*. 2017;69 (1):103–22.  
[doi: 10.1080/10510974.2017.1414068](https://doi.org/10.1080/10510974.2017.1414068).
- 25 Krishna A. Motivation with misinformation: conceptualizing lacuna individuals and publics as knowledge-deficient, issue-negative activists. *J Public Relat Res*. 2017;29(4):176–93. [doi: 10.1080/1062726X.2017.1363047](https://doi.org/10.1080/1062726X.2017.1363047).
- 26 Nicholson MS, Leask J. Lessons from an online debate about measles–mumps–rubella (MMR) immunization. *Vaccine*. 2012;30(25):3806–12.  
[doi: 10.1016/j.vaccine.2011.10.072](https://doi.org/10.1016/j.vaccine.2011.10.072).
- 27 Basch CH, Zybert P, Reeves R, Basch CE, 2017. What do popular YouTube™ videos say about vaccines? *Child Care Health Dev*. 2017;43(4):499–503.  
[doi: 10.1111/cch.12442](https://doi.org/10.1111/cch.12442).
- 28 Cuan-Baltazar JY, Muñoz-Perez MJ, Robledo-Vega C, Pérez-Zepeda MF, Soto-Vega E Donzelli et al. Misinformation on vaccination: a quantitative analysis of YouTube videos. *Hum Vaccines Immunother*. 2018;14(7):1654–9.  
[doi: 10.1080/21645515.2018.1454572](https://doi.org/10.1080/21645515.2018.1454572).
- 29 Porat T, Garaizar P, Ferrero M, Jones H, Ashworth M, Vadillo MA. Content and source analysis of popular tweets following a recent case of diphtheria in Spain. *Eur J Public Health*. 2019;29(1):117–22. [doi: 10.1093/eurpub/cky144](https://doi.org/10.1093/eurpub/cky144).
- 30 Tustin JL, Crowcroft NS, Gesink D, Johnson I, Keelan J, Lachapelle B. User-driven comments on a Facebook advertisement recruiting Canadian parents in a study on immunization: content analysis. *JMIR Publ Health Surv*. 2018;4(3):e10090.  
[doi: 10.2196/10090](https://doi.org/10.2196/10090).
- 31 Wood MJ. Propagating and debunking conspiracy theories on twitter during the 2015–2016 Zika virus outbreak. *Cyberpsychol Behav Soc Netw*. 2018;21(8):485–490.  
[doi: 10.1089/cyber.2017.0669](https://doi.org/10.1089/cyber.2017.0669).
- 32 Sharma M, Yadav K, Yadav N, Ferdinand KC. Zika virus pandemic—analysis of Facebook as a social media health information platform. *Am J Infect Contr*. 2017;45(3):301–2. [doi: 10.1016/j.ajic.2016.08.022](https://doi.org/10.1016/j.ajic.2016.08.022).
- 33 Seltzer EK, Horst-Martz E, Lu M, Merchant RM. Public sentiment and discourse about Zika virus on Instagram. *Public Health*. 2017;150:170–5. [doi: 10.1016/j.puhe.2017.07.015](https://doi.org/10.1016/j.puhe.2017.07.015).
- 34 Bora K, Das D, Barman B, Borah P. Are internet videos useful sources of information during global public health emergencies? A case study of YouTube videos during the 2015–16 Zika virus pandemic. *Pathog Glob Health* 2018;112(6):320–8.  
[doi: 10.1080/20477724.2018.1507784](https://doi.org/10.1080/20477724.2018.1507784)

- 35 Fung ICH, Fu K-W, Chan C-H, Chan BSB, Cheung C-N, Abraham T et al. Social media's initial reaction to information and misinformation on Ebola, August 2014: facts and rumors. *Public Health Rep.* 2016;131(3):461–73. doi: [10.1177/003335491613100312](https://doi.org/10.1177/003335491613100312).
- 36 Pathak R, Poudel DR, Karmacharya P, Pathak A, Aryal MR, Mahmood M et al. YouTube as a source of information on Ebola virus disease. *N Am J Med Sci.* 2015;7(7):306–9. doi: [10.4103/1947-2714.161244](https://doi.org/10.4103/1947-2714.161244).
- 37 Szmuda T, Syed MT, Singh A, Ali S, Özdemir C, Słoniewski P. YouTube as a source of patient information for coronavirus disease (COVID-19): a content-quality and audience engagement analysis. *Rev Med Virol.* 2020;30(5):e2132. doi: [10.1002/rmv.2132](https://doi.org/10.1002/rmv.2132).
- 38 Pennycook G, McPhetres J, Zhang Y, Lu JG, Rand DG. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychol Sci.* 2020;31(7):770–80. doi: [10.1177/0956797620939054](https://doi.org/10.1177/0956797620939054).
- 39 Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation, *J Med Internet Res* 2020;22(6):e19659. doi: [10.2196/19659](https://doi.org/10.2196/19659).
- 40 Li HO, Bailey A, Huynh D et al, YouTube as a source of information on COVID-19: a pandemic of misinformation?, *BMJ. Glob Health.* 2020;5:e002604.
- 41 Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z, Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study, *J Med Internet Res.* 2020;22(4):e19016.
- 42 do Nascimento IJB, Pizarro AB, Almeida JM, Azzopardi-Muscat N, Gonçalves MA, Björklunde M et al. Infodemics and health misinformation: a systematic review of reviews. *Bulletin of the World Health Organization.* 2022;9:544–61. doi: [10.2471/BLT.21.287654](https://doi.org/10.2471/BLT.21.287654). <https://apps.who.int/iris/handle/10665/362187>
- 43 Okuhara T, Ishikawa H, Okada M, Kato M, Kiuchi T. Assertions of Japanese websites for and against cancer screening: a text mining analysis. *Asian Pac J Cancer Prev.* 2017;18(4):1069–75. doi: [10.22034/APJCP.2017.18.4.1069](https://doi.org/10.22034/APJCP.2017.18.4.1069).
- 44 Chen L, Wang X, Peng T-Q. Nature and diffusion of gynecologic cancer-related misinformation on social media: analysis of tweets. *J Med Internet Res.* 2018;20(10):e11515. doi: [10.2196/11515](https://doi.org/10.2196/11515).
- 45 Chen H-M, Hu Z-K, Zheng X-L, Yuan Z-S, Xu Z-B, Yuan L-Q et al. Effectiveness of YouTube as a source of medical information on heart transplantation. *Interactive J Med Res.* 2013;2(2):e28. doi: [10.2196/ijmr.2669](https://doi.org/10.2196/ijmr.2669).
- 46 Kumar N, Pandey A, Venkatraman A, Garg N. Are video sharing web sites a useful source of information on hypertension? *J Am Soc Hypertens.* 2014;8(7):481–90. doi: [10.1016/j.jash.2014.05.001](https://doi.org/10.1016/j.jash.2014.05.001).

- 47 Qi J, Trang T, Doong J, Kang S, Chien AL. Misinformation is prevalent in psoriasis-related YouTube videos. *Dermatol Online J*. 2016;22(11):13030/qt7qc9z2m5. PMID:28329562.
- 48 Leong AY, Sanghera R, Jhajj J, Desai N, Jammu BS, Makowsky MJ. Is YouTube useful as a source of health information for adults with type 2 diabetes? A South Asian perspective. *Can J Diabetes*. 2018;42(4):395–403. doi: 10.1016/j.jcjd.2017.10.056.
- 49 Bessi A, Zollo F, Vicario MD, Scala A, Caldarelli G, Quattrocioni W. Trend of narratives in the age of misinformation. *PLOS One*. 2015;0(8):e0134641. doi: 10.1371/journal.pone.0134641.
- 50 Designing for disorder: Instagram’s pro-eating disorder bubble. Boston (MA): Fairplay; 2022 ([https://fairplayforkids.org/wp-content/uploads/2022/04/designing\\_for\\_disorder.pdf](https://fairplayforkids.org/wp-content/uploads/2022/04/designing_for_disorder.pdf)).
- 51 “Thinstagram”: Instagram’s algorithm fuels eating disorder epidemic. Washington (DC): Tech Transparency Project; 2021 (<https://www.techtransparencyproject.org/articles/thinstagram-instagrams-algorithm-fuels-eating-disorder-epidemic>).
- 52 Albarracin D, Romer D, Jones C, Hall Jamieson K, Jamieson P. Misleading claims about tobacco products in YouTube videos: experimental effects of misinformation on unhealthy attitudes. *J Med Internet Res*. 2018;20(6):e229. doi: 10.2196/jmir.9959.
- 53 Romer D, Reese Z, Jamieson PE. Misleading tobacco content is on the rise on YouTube. Cambridge (MA): The Harvard Kennedy School Misinformation Review; 2020. doi: 10.37016/mr-2020-010.
- 54 Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med*. 2019;240:112552. doi: 10.1016/j.socscimed.2019.112552.
- 55 Li, Y-J, Cheung CMK, Shen X-L, Lee MKO. Health misinformation on social media: a literature review. In: Pacific Asia Conference on Information Systems, 8-12 July 2019, Xi’an, China. Atlanta (GA): Association for Information Systems; 2019:194 (<https://aisel.aisnet.org/pacis2019/194>).
- 56 Pasquetto I, Swire-Thompson B, Amazeen MA, Benevenuto F, Brashier NM, Bond RM et al. Tackling misinformation: What researchers could do with social media data. *HKS Mis Rev*. 2020. doi: 10.37016/mr-2020-49.
- 57 Report of the European Digital Media Observatory’s working group on platform-to-researcher data access. Brussels: European Digital Media Observatory; 2022 (<https://edmo.eu/wp-content/uploads/2022/02/Report-of-the-European-Digital-Media-Observatorys-Working-Group-on-Platform-to-Researcher-Data-Access-2022.pdf>).
- 58 Infodemic. In: Health topics [website]. Geneva: World Health Organization; 2022 (<https://www.who.int/health-topics/infodemic>).

- 59 The fake news machine: inside a town gearing up for 2020. Atlanta (GA): CNN; 2-2-  
(<https://money.cnn.com/interactive/media/the-macedonia-story/>).
- 60 The disinformation dozen: why platforms must act on twelve leading online anti-vaxxers. Washington (DC): Center for Countering Digital Hate; 2021  
(<https://counterhate.com/research/the-disinformation-dozen/>).
- 61 Bay S, Fredheim R. How social media companies are failing to combat inauthentic behaviour online. Riga: NATO Strategic Communications Centre of Excellence; 2022 (<https://stratcomcoe.org/publications/how-social-media-companies-are-failing-to-combat-inauthentic-behaviour-online/33>).
- 62 Juneja P, Mitra T. Auditing E-commerce platforms for algorithmically curated vaccine misinformation. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). Assoc Comput Machinery. 2021;186:1–27. doi: 10.1145/3411764.3445250.
- 63 Lewin K. Frontiers in group dynamics: concept, method and reality in social science; social equilibria and social change. Hum Relat. 1947;5:145–6. doi: 10.1177/001872674700100103.
- 64 Lewin K. Frontiers in group dynamics II: channels of group life; social planning and action research. Hum Relat. 1947;1(2):143–53. doi: 10.1177/001872674700100201.
- 65 Shoemaker PJ. Gatekeeping. Newbury Park (CA): Sage Publications; 1991.
- 66 Soroka SN. The gatekeeping function: distributions of information in media and the real world. J Politic. 2012;74:514–515. doi: 10.1017/S002238161100171X.
- 67 White DM. The “Gatekeeper”: a case study in the selection of news’. In: White, DM, Dexter, LA, editors. People, society and mass communications. London: The Free Press of Glencoe; 1950:27.
- 68 2021 World Press Freedom Index: journalism, the vaccine against disinformation, blocked in more than 130 countries. Paris: Reporters Without Borders; 2021 (<https://rsf.org/en/2021-world-press-freedom-index-journalism-vaccine-against-disinformation-blocked-more-130-countries>).
- 69 Ellis R. Poll shows low trust in us public health systems. New York: WebMD; 2021 (<https://www.webmd.com/lung/news/20210514/poll-shows-low-trust-in-u-s-public-health-systems>).
- 70 The Digital Markets Act: ensuring fair and open digital markets. Brussels: European Commission; 2020  
([https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en)).
- 71 Diviani N, Van Den Putte B, Giani S, Van Weert JC. Low health literacy and evaluation of online health information: a systematic review of the literature. J Med Internet Res. 2015;17(5):e112. doi: 10.2196/jmir.4018.

- 72 Kickbusch I, Pelikan J, Apfel F, Tsouros A. Health literacy: the solid facts. Copenhagen: WHO Regional Office for Europe; 2013 (<https://apps.who.int/iris/handle/10665/326432>).
- 73 Van der Vaart R, Drossaert CH, de Heus M, Taal E, van de Laar MA. Measuring actual eHealth literacy among patients with rheumatic diseases: a qualitative analysis of problems encountered using Health 1.0 and Health 2.0 applications. *J Med Internet Res*. 2013;15(2):e27. doi: 10.2196/jmir.2428.
- 74 Cooper B, Lee W, Goldacre B, Sanders T. The quality of the evidence for dietary advice given in UK national newspapers. *Public Underst Sci*. 2012;21(6):664–73. doi: 10.1177/0963662511401782.
- 75 Syed-Abdul S, Fernandez-Luque L, Jian W-S, Li Y-C, Crain S, Hsu M-H et al. Misleading health-related information promoted through video-based social media: anorexia on YouTube. *J. Med. Internet Res*. 2013;15(2):e30. doi: 10.2196/jmir.2237.
- 76 Hasen RL. Cheap speech and what it has done (to American democracy). California (CA): Irvine School of Law; 2017 (<https://ssrn.com/abstract=3017598>).
- 77 Berger J, Milkman K. Emotion and virality: what makes online content go viral? *NIM Mark Int Rev*. 2013;5(1):18–23. doi: 10.2478/gfkmir-2014-0022.
- 78 2022 strengthened code of practice on disinformation. Brussels: European Commission; 2022 (<https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>).
- 79 Cobbe J, Singh J. Regulating recommending: motivations, considerations, and principles. *SSRN Electronic Journal*. Cobbe J, Singh J. Regulating recommending: motivations, considerations, and principles. *SSRN Elec J*. 2019;10(3):1–37. doi: 10.2139/ssrn.3371830.
- 80 Marecos J, Duarte FdA. A right to lie in the age of disinformation: protecting free speech beyond the first amendment. In: De Abreu BS (editor). *Media literacy, equity, and justice*. Abingdon-on-Thames: Routledge; 2022.
- 81 2021 Edelman trust barometer [website] Chicago (IL): Edelman 2021 (<https://www.edelman.com/trust/2021-trust-barometer>).
- 82 Vosoughi S, Roy D, Aral S. The spread of true and false news online. *Science*. 2018;359(6380):1146–51. doi: 10.1126/science.aap9559.
- 83 The big business of medical misinformation. San Francisco (CA): Healthline; 2021 (<https://www.healthline.com/health-news/the-big-business-of-medical-misinformation#4>).
- 84 Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. *J Med Internet Res*. 2020;22(6):e21820. doi: 10.2196/21820.
- 85 Turner J. The growing importance of digital literacy. Washington (DC): IREX; 2021.
- 86 Nutbeam D. The evolving concept of health literacy. *Soc Sci Med*. 2009;67:2072–8. doi: 10.1016/j.socscimed.2008.09.050.

- 87 Closing the gap in a generation: health equity through action on the social determinants of health: final report. Geneva: World Health Organization; 2008 (<https://apps.who.int/iris/handle/10665/43943>).
- 88 Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med.* 2011;155:97–107. doi: [10.7326/0003-4819-155-2-201107190-00005](https://doi.org/10.7326/0003-4819-155-2-201107190-00005).
- 89 Nutbeam D. Discussion paper on promoting, measuring and implementing health literacy: implications for policy and practice in non-communicable disease prevention and control. Geneva: World Health Organization; 2017 (<https://www.researchgate.net/publication/315091628>).
- 90 Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the internet. *J Med Internet Res.* 2009;11(1):e11. doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157).
- 91 A guide to anti-misinformation actions around the world. St Petersburg (FL): The Poynter Institute; 2018 ([https://www.poynter.org/ifcn/anti-misinformation-actions/?\\_\\_cf\\_chl\\_jschl\\_tk\\_\\_=pmd\\_B4wWdoXfmA0p0f8qsOwhvb0CSbAhUaVaioloOlnwYqw-1629884471-0-gqNtZGzNAfujcnBszQi9](https://www.poynter.org/ifcn/anti-misinformation-actions/?__cf_chl_jschl_tk__=pmd_B4wWdoXfmA0p0f8qsOwhvb0CSbAhUaVaioloOlnwYqw-1629884471-0-gqNtZGzNAfujcnBszQi9)).
- 92 Tackling online disinformation. Brussels: European Commission; 2022 (<https://digital-strategy.ec.europa.eu/en/policies/online-disinformation>).
- 93 A multi-dimensional approach to disinformation: report of the independent High-level Group on fake news and online disinformation. Brussels: European Commission; 2018 (<https://op.europa.eu/en/publication-detail/-/publication/6ef4df8b-4cea-11e8-b6d-01aa75ed71a1>).
- 94 Action plan against disinformation. Brussels: European Commission; 2018 ([https://ec.europa.eu/info/sites/default/files/eu-communication-disinformation-euco-05122018\\_en.pdf](https://ec.europa.eu/info/sites/default/files/eu-communication-disinformation-euco-05122018_en.pdf)).
- 95 Canada's digital charter: trust in a digital world. Ottawa: Government of Canada; 2022 (<https://ised-isde.canada.ca/site/innovation-better-canada/en/canadas-digital-charter-trust-digital-world>).
- 96 Carta Portuguesa de direitos humanos na era digital [Portuguese charter for human rights in the digital age]. Lisboa: Portugal Ministry of Education; 2021 (in Portuguese; <https://www.dge.mec.pt/noticias/carta-portuguesa-de-direitos-humanos-na-era-digital>).
- 97 Charter of digital rights. Madrid: Government of Spain; 2021 ([https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2021/SPAIN\\_Charter-of-Digital-Rights.pdf](https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2021/SPAIN_Charter-of-Digital-Rights.pdf)).



- 98 European digital rights and principles. Brussels: European Commission; 2022 (<https://digital-strategy.ec.europa.eu/en/policies/digital-principles>).
- 99 European democracy action plan. Brussels: European Commission; 2020 (<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0790&from=EN>).
- 100 More obligations and accountability for online platforms. In: European democracy action plan. Brussels: European Commission; 2020 (4.2; <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0790&from=EN>).
- 101 The Digital Services Act package. Brussels: European Commission; 2020 (<https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>).
- 102 Proposal for a regulation of the European Parliament and of the Council on a single market for digital services (Digital Services Act) and amending directive 2000/31/EC. Brussels: European Commission; 2020; para. 68 ([https://ec.europa.eu/info/sites/default/files/proposal\\_for\\_a\\_regulation\\_on\\_a\\_single\\_market\\_for\\_digital\\_services.pdf](https://ec.europa.eu/info/sites/default/files/proposal_for_a_regulation_on_a_single_market_for_digital_services.pdf)).
- 103 Code of practice on disinformation. Brussels: European Commission; 2018 (<https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation>).
- 104 Joint communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Action Plan against Disinformation. Brussels: European Commission; 2018 (<https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52018JC0036>).
- 105 The assembly of the signatories to strengthen the code of practice on disinformation. Brussels: European Commission; 2021 (<https://digital-strategy.ec.europa.eu/en/news/assembly-signatories-strengthen-code-practice-disinformation>).
- 106 EDMO at a glance. In: European Digital Media Observatory [website]. Brussels: European Digital Media Observatory; 2022 (<https://edmo.eu/edmo-at-a-glance>).
- 107 O tribunal da democracia [The court of democracy]. Brasília: Tribunal Superior Eleitoral; 2022 (in Portuguese and English; <https://www.tse.jus.br/>).
- 108 Sundaria EK. The fight against “fake news” in Indonesia. Sai Wan: Asia Times; 2019 (<https://asiatimes.com/2019/04/the-fight-against-fake-news-in-indonesia/>).
- 109 Digital citizen initiative. Ottawa: Government of Canada; 2022 (<https://www.canada.ca/en/canadian-heritage/services/online-disinformation.html>).
- 110 Mon opinion [My opinion]. Brussels: Federal Government of Belgium; 2022 (in Dutch and French; <https://monopinion.belgium.be/>).



- 111 Question écrite posée à Jean-Claude Marcourt, Vice Président et Ministre de l'Enseignement supérieur, de la Recherche et des Médias [Written question to Jean-Claude Marcourt, Vice President and Minister for Higher Education, Research and Media]. Bruxelles: Parlement de la Fédération Wallonie; 2008 (in French; <https://www.moinnet.be/document/FWB/QE/02-12-15%20QE%20FWB%20Statistiques%20ETALV.pdf>).
- 112 Nash R, Elmer S, Thomas K, Osborne R, Macintyre K, Shelley B et al. HealthLit4Kids study protocol; crossing boundaries for positive health literacy outcomes. BMC Public Health. 2018;690. doi: 10.1186/s12889-018-5558-7.
- 113 Stop and consider. Canberra, Australian Electoral Commission; 2022 (<https://www.aec.gov.au/elections/electoral-advertising/files/stop-and-consider-external-flyer.pdf>).
- 114 Making it easier: a health literacy action plan 2017–2025. Edinburgh: Scottish Government; 2016 (<https://www.gov.scot/publications/making-easier-health-literacy-action-plan-scotland-2017-2025/>).
- 115 Code of conduct on countering illegal hate speech online. Brussels: European Commission; 2016 ([https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online\\_en#theeucodeofconduct](https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en#theeucodeofconduct)).
- 116 Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law. Brussels: European Commission; 2008 (<https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32008F0913>).
- 117 Progress on combating hate speech online through the EU code of conduct. Brussels: European Commission; 2019 ([https://ec.europa.eu/info/sites/default/files/aid\\_development\\_cooperation\\_fundamental\\_rights/assessment\\_of\\_the\\_code\\_of\\_conduct\\_on\\_hate\\_speech\\_on\\_line\\_-\\_state\\_of\\_play\\_\\_0.pdf](https://ec.europa.eu/info/sites/default/files/aid_development_cooperation_fundamental_rights/assessment_of_the_code_of_conduct_on_hate_speech_on_line_-_state_of_play__0.pdf)).
- 118 Media Act and Act on amendments to the media Act (OG Nos. 59/04, 84/11 and 81/13). Zagreb: Government of Croatia; 2013 (<https://wipolex.wipo.int/en/legislation/details/16388>).
- 119 Electronic Media Act. Zagreb: Government of Croatia; 2021 (OG 153/09, OG 84/11, OG 94/13, and OG 136/13; [https://www.e-mediji.hr/repository\\_files/file/502/](https://www.e-mediji.hr/repository_files/file/502/)).
- 120 Tworek HJS. Fighting hate with speech law: media and German visions of democracy, J Holocaust Res. 2021;35(2):106–22. doi: 10.1080/25785648.2021.
- 121 Décision n° 2020–801 DC du 18 juin 2020: loi visant à lutter contre les contenus haineux sur internet [Decision n° 2020–801 DC of June 18, 2020: law to fight against hateful content on the Internet]. Paris: Conseil Constitutionnel; 2020 (<https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm>).
- 122 Prominent Bahrain lawyer gets suspended sentence for 'inciting hatred'. Reuters. 14 September 2016 (<https://www.reuters.com/article/us-bahrain-security-idUSKBN2652A7>).

- 123 Law Minister: Government to amend Digital Security Act. Dhaka Tribune. 2 March 2021 (<https://archive.dhakatribune.com/law-and-rights/2021/03/02/law-minister-no-one-can-be-arrested-under-dsa-until-police-investigation>).
- 124 Belarus passes legislation against ‘fake news’ media. Radio Free Europe. 14 June 2018 (<https://www.rferl.org/a/belarus-assembly-passes-controversial-fake-news-media-legislation/29291033.html>).
- 125 Burkinabe legislative threat to press freedom must be declared unconstitutional. Reporters Without Borders. 11 June 2021 (<https://rsf.org/en/burkinabe-legislative-threat-press-freedom-must-be-declared-unconstitutional>).
- 126 Cambodia “fake news” crackdown prompts fears over press freedom. The Guardian. 6 July 2018 ([https://www.theguardian.com/world/2018/jul/06/cambodia-fake-news-crackdown-prompts-fears-over-press-freedom?CMP=share\\_btn\\_tw](https://www.theguardian.com/world/2018/jul/06/cambodia-fake-news-crackdown-prompts-fears-over-press-freedom?CMP=share_btn_tw)).
- 127 In Cameroon, journalists are being jailed on charges of “fake news”. The Washington Post. 16 December 2018 ([https://www.washingtonpost.com/world/africa/in-cameroon-journalists-are-being-jailed-on-charges-of-fake-news/2018/12/15/80bcb5c6-f9ad-11e8-8642-c9718a256cbd\\_story.html](https://www.washingtonpost.com/world/africa/in-cameroon-journalists-are-being-jailed-on-charges-of-fake-news/2018/12/15/80bcb5c6-f9ad-11e8-8642-c9718a256cbd_story.html)).
- 128 Misinformation policies. In: Google/YouTube Help [website]. Mountain View (CA): Alphabet Inc.; 2022 ([https://support.google.com/youtube/answer/10834785?hl=en&ref\\_topic=10833358](https://support.google.com/youtube/answer/10834785?hl=en&ref_topic=10833358)).
- 129 Integrity and authenticity. In: TikTok [website]. Culver City (CA): ByteDance; 2022 (<https://www.tiktok.com/community-guidelines?lang=en#37>).
- 130 False news. In: Transparency at Meta [website]. Menlo Park (CA): Meta; 2022 (<https://transparency.fb.com/en-gb/policies/community-standards/false-news/>).
- 131 The virality project (2022): memes, magnets and microchips: narrative dynamics around COVID-19 vaccines. Stanford (CA): Stanford Internet Observatory; 2022 (<https://purl.stanford.edu/mx395xj8490>).
- 132 Assessment of the Code of practice on disinformation: achievements and areas for further improvement. Brussels: European Commission; 2020 (<https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>).
- 133 COVID-19 disinformation monitoring programme. Brussels: European Commission; 2020(<https://digital-strategy.ec.europa.eu/en/policies/covid-19-disinformation-monitoring>).
- 134 The virality project final report. Stanford (CA): Stanford Internet Observatory; 2022:122–35 (<https://cyber.fsi.stanford.edu/io/news/virality-project-final-report>).

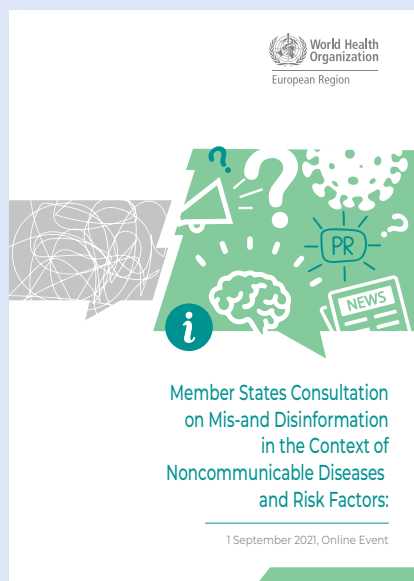
- 135 YouTube community guidelines enforcement. Mountain View (CA): Alphabet Inc.; 2022 ([https://transparencyreport.google.com/youtube-policy/removals?hl=en&comments\\_by\\_source=period:2021Q4&lu=comments\\_by\\_source](https://transparencyreport.google.com/youtube-policy/removals?hl=en&comments_by_source=period:2021Q4&lu=comments_by_source)).
- 136 COVID-19 misleading information policy. In: Twitter [website]. San Francisco (CA): Twitter; 2022 (<https://help.twitter.com/en/rules-and-policies/medical-misinformation-policy>).
- 137 Updating our approach to misleading information. In: Twitter [website]. San Francisco (CA): Twitter; 2020 ([https://blog.twitter.com/en\\_us/topics/product/2020/updating-our-approach-to-misleading-information](https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information)).
- 138 COVID-19 information centre. In: Facebook [website]. Menlo Park (CA): Meta; 2022 ([https://www.facebook.com/coronavirus\\_info](https://www.facebook.com/coronavirus_info)).
- 139 Combating COVID-19 misinformation across our apps. In: Facebook [website]. Menlo Park (CA): Meta; 2022 (<https://about.fb.com/news/2020/03/combating-covid-19-misinformation/>).
- 140 WhatsApp launches first TV campaign to fight misinformation in India. In: YouTube [website]. Mountain View (CA): Alphabet Inc.; 2022 ([https://www.youtube.com/watch?v=HWaVK2lpguw&ab\\_channel=HindustanTimes](https://www.youtube.com/watch?v=HWaVK2lpguw&ab_channel=HindustanTimes)).
- 141 Supporting our community through COVID-19. In: TikTok [website]. Culver City (CA): ByteDance; 2022 (<https://www.tiktok.com/safety/en-us/covid-19/>).
- 142 Commit to transparency: sign up for the International Fact-Checking Network's code of principles [website]. St Petersburg (FL): The Poynter Institute; 2022 (<https://ifcncodeofprinciples.poynter.org/>).
- 143 Meta's third-party fact-checking program. In: Facebook [website]. Menlo Park (CA): Meta; 2022 (<https://www.facebook.com/journalismproject/programs/third-party-fact-checking>).
- 144 State of the fact-checkers 2021. St Petersburg (FL): The Poynter Institute; 2021 ([https://www.poynter.org/wp-content/uploads/2022/01/IFCN\\_2022\\_StateFactChecking2021\\_v06.pdf](https://www.poynter.org/wp-content/uploads/2022/01/IFCN_2022_StateFactChecking2021_v06.pdf)).
- 145 AP expands access to factual information with new Twitter collaboration. The Associated Press. 2 August 2021 (<https://www.ap.org/press-releases/2021/ap-expands-access-to-factual-information-with-new-twitter-collaboration>).
- 146 Introducing Birdwatch, a community-based approach to misinformation. In: Twitter [website]. San Francisco (CA): Twitter; 2022 ([https://blog.twitter.com/en\\_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation](https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation)).

- 147 Combating misinformation and election interference on TikTok. In: TikTok [website]. Culver City (CA): ByteDance; 2022 (<https://newsroom.tiktok.com/en-us/combating-misinformation-and-election-interference-on-tiktok>).
- 148 Our approach to misinformation. In: Transparency at Meta [website]. Menlo Park (CA): Meta; 2022 (<https://transparency.fb.com/en-gb/features/approach-to-misinformation/>).
- 149 Counting strikes. In: Transparency at Meta [website]. Menlo Park (CA): Meta; 2022 (<https://transparency.fb.com/en-gb/enforcement/taking-action/counting-strikes/>).
- 150 The News Provenance Project [website]. New York: The New York Times; 2022 (<https://www.newsprovenanceproject.com/>).
- 151 The Content Authenticity Initiative [website]. San Jose (CA): Adobe; 2022 (<https://contentauthenticity.org/https://contentauthenticity.org/>).
- 152 Project Origin. London: British Broadcasting Corporation; 2020 (<https://www.bbc.co.uk/beyondfakenews/project-origin/>).
- 153 Material from the Internet and social media. In: Editorial guidelines. London: British Broadcasting Corporation; 2019:3.3.12 (<http://downloads.bbc.co.uk/guidelines/editorialguidelines/pdfs/bbc-editorial-guidelines-whole-document.pdf>).
- 154 La charte des Décodeurs [The charter of Les Décodeurs]. Paris: Le Monde; 2014 (in French; [https://www.lemonde.fr/les-decodeurs/article/2014/03/10/la-charte-des-decodeurs\\_4365106\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2014/03/10/la-charte-des-decodeurs_4365106_4355770.html)).
- 155 About Reuters fact check. London: Reuters; 2022 (<https://www.reuters.com/fact-check/about>).
- 156 Les Décodeurs [website]. Paris: Le Monde; 2014 (in French; <https://www.lemonde.fr/les-decodeurs/>).
- 157 Reality check [website]. London: British Broadcasting Corporation; 2022 ([https://www.bbc.co.uk/news/reality\\_check](https://www.bbc.co.uk/news/reality_check)).
- 158 Polígrafo [website]. Lisboa: Servidor de Apontadores Portugueses Online; 2022 (<https://poligrafo.sapo.pt/>).
- 159 Young reporter [website]. London: British Broadcasting Corporation; 2022 (<https://www.bbc.co.uk/teach/young-reporter>).
- 160 BBC iReporter [website]. London: British Broadcasting Corporation; 2022 (<https://www.bbc.co.uk/news/resources/idt-8760dd58-84f9-4c98-ade2-590562670096>).
- 161 Slowing the infodemic: how to spot COVID-19 misinformation [YouTube video]. London: Thomson Reuters; 2022 ([https://www.youtube.com/watch?v=jNe7SZgn000&ab\\_channel=ThomsonReuters](https://www.youtube.com/watch?v=jNe7SZgn000&ab_channel=ThomsonReuters)).

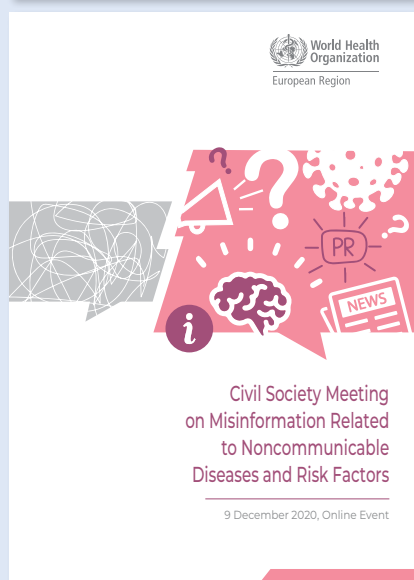
- 162 Slowing the infodemic: how to spot COVID-19 misinformation. London: Thomson Reuters; 2022 (<https://www.thomsonreuters.com/content/dam/ewp-m/documents/thomsonreuters/en/pdf/social-impact/name-reuters-classroom-guide.pdf>).
- 163 Livingstone S. Media literacy: what are the challenges and how can we move towards a solution? London: London School of Economics; 2018 (<https://blogs.lse.ac.uk/medialse/2018/10/25/media-literacy-what-are-the-challenges-and-how-can-we-move-towards-a-solution/>).
- 164 Murrock E, Amuya J, Druckman M, Liubyva T. Winning the war on state-sponsored propaganda: gains in the ability to detect disinformation a year and a half after completing a Ukrainian news media literacy program. Washington (DC): International Research and Exchanges Board; 2018 (<https://www.irex.org/sites/default/files/node/resource/impact-study-media-literacy-ukraine.pdf>).
- 165 Infodemic Management. Geneva: World Health Organization; 2022 (<https://www.who.int/teams/epi-win/infodemic-management>).
- 166 WHO competency framework: building a response workforce to manage infodemics. Geneva: World Health Organization; 2021 (<https://www.who.int/publications/i/item/9789240035287>).
- 167 1st WHO Infodemic Manager training, November 2020. Geneva: World Health Organization; 2020 (<https://www.who.int/teams/epi-win/infodemic-management/1st-who-training-in-infodemic-management>).
- 168 EPI-WIN: WHO information network for epidemics. Geneva: World Health Organization; 2021 (<https://www.who.int/teams/epi-win>).
- 169 Speech by WHO Director. In: Munich security conference, 15 February 2020, Munich, Germany. Geneva: World Health Organization; 2020 (<https://www.who.int/dg/speeches/detail/munich-security-conference>).
- 170 Coronavirus disease (COVID-19) training: online training [website]. Geneva: World Health Organization; 2022 (<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/training/online-training>).
- 171 Coronavirus disease (COVID-19) advice for the public: Mythbusters [website]. Geneva: World Health Organization; 2022 (<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters>).
- 172 WHO dashboard of COVID-19 related recommendations [website]. Geneva: World Health Organization; 2022 (<https://app.powerbi.com/w?r=eyJrIjoIO DgyYjRmZjQtN2UyNi00NGE4LTg1YzMtYzE2OGFhZjBiYzFjIiwidCI6ImY2MTBjMGI3LWJkMjQtNGl3OS04MTBiLTNkYzI4MGFmYjU5MCIslmMiOjh9&pageName=R eportSection729b5bf5a0b579e86134>).
- 173 Stop the spread: countering misinformation about COVID-19. Geneva: World Health Organization; 2020 (<https://www.who.int/news-room/feature-stories/detail/countering-misinformation-about-covid-19>).

- 174 Go viral! [online game]. Cambridge: Tilt Studio at the University of Cambridge; 2021 (<https://www.goviralgame.com/>).
- 175 Gezondheid en wetenschap [Health & science]. Brussels: Belgian Center for Evidence-Based Medicine; 2022  
(in Dutch and English; <https://www.gezondheidenwetenschap.be/>).
- 176 Institutet för hälsa och välfärd [Finnish institute for health and welfare]. Helsinki: Finnish institute for health and welfare; 2022  
(in Swedish and English; <https://thl.fi/en/web/thlfi-en>).
- 177 Facebook employees flag drug cartels and human traffickers. the company's response is weak, documents show. The Wall Street Journal. 16 September 2021 (<https://www.wsj.com/articles/facebook-drug-cartels-human-traffickers-response-is-weak-documents-11631812953>).
- 178 NewsGuard [website]. New York: NewsGuard; 2022 (<https://www.newsguardtech.com/>).
- 179 Science Media Centre [website]. London: Science Media Centre; 2022 (<https://www.sciencemediacentre.org/>).
- 180 WHO launches a chatbot on Facebook Messenger to combat COVID-19 misinformation. Geneva: World Health Organization; 2020 (<https://www.who.int/news-room/feature-stories/detail/who-launches-a-chatbot-powered-facebook-messenger-to-combat-covid-19-misinformation>).

## Annex



Member States consultation  
on mis- and disinformation in the context of  
noncommunicable diseases and risk factors:  
1 September 2021, online event  
(<https://apps.who.int/iris/handle/10665/363674>).



Civil society meeting on misinformation related  
to noncommunicable diseases and risk factors:  
9 December 2020, online event.  
(<https://apps.who.int/iris/handle/10665/363676>).



Media and online platforms meeting on misinformation  
related to noncommunicable diseases and risk factors:  
5 February 2021, online event  
(<https://apps.who.int/iris/handle/10665/363677>).



## **The WHO Regional Office for Europe**

The World Health Organization (WHO) is a specialized agency of the United Nations created in 1948 with the primary responsibility for international health matters and public health. The WHO Regional Office for Europe is one of six regional offices throughout the world, each with its own programme geared to the particular health conditions of the countries it serves.

### **Member States**

Albania	Lithuania
Andorra	Luxembourg
Armenia	Malta
Austria	Monaco
Azerbaijan	Montenegro
Belarus	Netherlands
Belgium	North Macedonia
Bosnia and Herzegovina	Norway
Bulgaria	Poland
Croatia	Portugal
Cyprus	Republic of Moldova
Czechia	Romania
Denmark	Russian Federation
Estonia	San Marino
Finland	Serbia
France	Slovakia
Georgia	Slovenia
Germany	Spain
Greece	Sweden
Hungary	Switzerland
Iceland	Tajikistan
Ireland	Türkiye
Israel	Turkmenistan
Italy	Ukraine
Kazakhstan	United Kingdom
Kyrgyzstan	Uzbekistan
Latvia	

**WHO/EURO:2022-6260-46025-66542**

### **WHO European Office for the Prevention and Control of Noncommunicable Diseases**

9, Leontievsky pereulok  
125009 Moscow, Russian Federation  
**Tel.:** +7 (495) 787 21 17

**Email:** [NCDoffice@who.int](mailto:NCDoffice@who.int)

**Website:** <https://www.who.int/europe/>