Managing Rumours: Guidance to Country Offices

SIDA SRH/HIV COVID-19, Intervention Set n. 1

October 2020

UNICEF Eastern and Southern Africa Regional Office

TABLE OF CONTENTS

GLOSSARY OF TERMS

Disinformation: false information shared on purpose to deceive and harm.

Misinformation: false information shared inadvertently.

Rumour: unverified information.

Rumour management: process of preventing, tracking or responding to unverified information in order to improve accuracy of the discourse around a certain issue.

Social listening: process of monitoring and analysing information from selected communication channels (e.g. social media platforms) to understand concerns, questions, or beliefs expressed by a certain audience around an issue.

OVERVIEW

This guidance document provides suggestions on how to translate insights from social listening into rumour management activities. The document draws on the "[Eastern and Southern Africa Regional Social Media Monitoring and Rumor Management Strategy to Support Risk Communication and Community Engagement (RCCE) on COVID-19](#)" developed by UNICEF Communication for Development section in the Eastern and Southern Africa Regional Office in collaboration with the Communication section.

The long-term goal of rumour management activities is to build trust with the communities served and mitigate the impact of false content (misinformation/ disinformation) on access to services and health outcomes. While the recommendations in this guide are applicable to rumour management in general, the examples and resources are focused on the COVID-19 response and, in particular, Sexual and Reproductive Health and Rights (SRHR) issues.

This guide focuses on three types of strategic actions to address rumours and mis/disinformation: preventive strategies, early detection and associated indirect actions, and direct engagement.

BACKGROUND

The COVID-19 pandemic has been closely tied with what has been called an *infodemic*, a "second disease" that occurs when massive information volumes (particularly with a high prevalence of false information) hinder the public health response[1]. The threats of false information are real, even more so in an outbreak type of situation where people are concerned, constantly looking for information on how to handle emerging challenges and where scientific evidence is evolving rapidly.

Social listening tracks questions and queries, as well as concerns, complaints and suggestions shared by the public. Social listening is the first step in rumour management, as it allows to identify rumours and draw insights for content development. Rumours are defined as information that has not been verified, as well as false information. The latter category encompasses both misinformation, when false information is shared inadvertently, and disinformation, when false information is shared deliberately. While the scholarly debate around definitions of these concepts is ongoing and additional categorizations have been proposed and discussed [2], this document refers to rumours as general unverified information and misinformation as content that has been proven to be false.

---

[1] World Health Organization. Infodemic management: infodemiology. https://www.who.int/teams/risk-communication/infodemic-management

[2] First Draft News (2017). Fake news, it's complicated. https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79

PREVENTING RUMOURS

Adopting a proactive approach to communication is a key component of effective rumour management. Factoring the risk of rumours and misinformation in activity planning can help prevent false content or limit its spread. Prevention activities work on two aspects: developing content that is "rumour proof" and increasing audiences' resilience to misinformation.

Rumour-proof content

While scientific evidence on the effectiveness of different rumour management approaches is still growing, there are a few key recommendations on content development that misinformation experts have pointed to:

- **Up to date and consistent.** While evidence evolves over time, providing the most recent information in a way that is consistent across owned platforms helps build trust and strengthen the position of the organization as a reliable source of information in the eyes of its audiences. For instance, if recommendations around a certain preventive measure have changed (e.g. as happened with masks), it is important to ensure old content from websites and social media posts that is now outdated is removed, changed, or labeled to clarify that new guidance is available. Failing to do so may rise confusion, questions and even rumours.

- **Strategically located.** Content around topics that are related to programmatic priorities should be placed strategically, for instance through search engine optimization or direct collaboration with digital platforms. When users ask a certain question in the search bar of a given platform, the top search results should provide accurate information and link them to helpful evidence-based resources. In past years, some social media platforms have worked with public health organizations to provide accurate content on critical issues. For example, when searching for vaccine-related keywords on Twitter in certain countries, the first resulting link is to a public health source. Unfortunately, these features are not always available in all languages or geographical locations, and partnerships with platforms may be needed to further improve the service.

Similarly, it's important that content that appears on featured snippets[3] (sometimes referred to as "answer boxes", like in the figure below) on search engines is reliable and aligned with the recommendations provided by UNICEF and its partners.
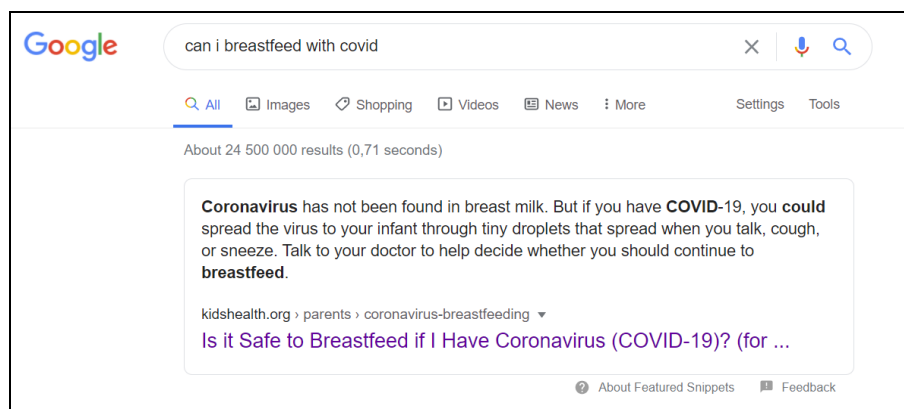
- **User-centered**. Users ask all sorts of questions on search engines like Google, and oftentimes they rely on answers from sources that are not accurate because that is what is available. When building content, it is key to think about the possible variations and questions that could emerge around the topic and try to incorporate specific key words. When content is not comprehensively addressing the issue at hand, there is room for "data voids", which emerge when few results are available to answer a certain search query[4]. Data voids can be easily filled with misinformation and rumours. Foreseeing critical occurrences or events that may generate information needs among certain population groups is key in this respect, as it allows to prepare content that addresses concerns that may arise (even indirectly). Sharing this type of content ahead of time may help mitigate the public's negative reactions to these occurrences. Building user-centered content also implies amplifying trusted voices and identifying spokespersons who are more likely to connect with a particular audience.

Audiences' resilience to rumours

Another key approach to preventing rumours is supporting communities in developing the skills to become more resilient to false content. This shift requires reaching out to a range of audiences including community leaders, journalists and local media. It is worth noting that different groups of the population may have varying digital media literacy levels and different behaviors when it comes to misinformation sharing[5], so it's worth reflecting on which priority

---

[3] A featured snippet is a selected search result that appears at the top of the search results page in the form of a box. The goal of featured snippets is to provide a quick answer to the search/question asked by the user.
[4] Data & Society (2019). Data voids: where missing data can easily be exploited. https://datasociety.net/wp-content/uploads/2019/11/Data-Voids-2.0-Final.pdf
[5] Older users share more misinformation. Your guess why might be wrong. Humans and Technology, https://www.technologyreview.com/2020/05/26/1002243/misinformation-older-adults/

groups are being considered to build resilience against rumours and how to best approach them. Interventions that help increase emotional skepticism[6] include:

- **Promotional campaigns.** Developing targeted messages that encourage users to think before sharing a piece of content they are not sure is accurate can be an effective way to raise awareness around how digital behaviors can impact people's decisions. An example of this is the Share Verified campaign launched by the UN to promote healthy information behaviors in times of COVID-19.[7]

- **Media and information literacy programs.** Training individuals to approach the news with skepticism and unpack it to identify content at risk for misinformation is key. An example of a curriculum to help journalists develop these skills is offered by the handbook "Journalism, fake news and disinformation" developed by UNESCO.[8] Internews also has several resources for journalists that are specific to COVID-19.

- **Weight of evidence reporting.** Weight of evidence reporting accounts and weights scientific evidence on a certain issue, rather than reporting "both sides" of a given issue[9]. This type of approach to journalistic coverage of news and issues has been advocated over the years particularly around issues such as global warming, where the vast majority of the scientific community agrees on the key evidence and recommendations. Promoting weight of evidence messages may influence attitudes towards certain key issues, for instance studies on immunization have suggested that it increases positive attitudes towards vaccines by reducing information uncertainty.[10,11]

- **Inoculation games.** These kind of pre-bunking interventions teach users about the techniques commonly used to produce disinformation, such as polarization and impersonation[12,13,14].

---

[6] CBS News, Interview with Clair Wardle. https://www.wcbi.com/emotional-skepticism-needed-to-stop-spread-of-deepfakes/
[7] UN Shared Verified. https://www.shareverified.com/en
[8] UNESCO (2018). Journalism, fake news & disinformation handbook. https://en.unesco.org/fightfakenews
[9] Dunwoody, S. (2005). Weight-of-evidence reporting: what is it? Why use it? https://niemanreports.org/articles/weight-of-evidence-reporting-what-is-it-why-use-it/
[10] Christopher E. Clarke, Brooke Weberling McKeever, Avery Holton & Graham N. Dixon (2015) The Influence of Weight-of-Evidence Messages on (Vaccine) Attitudes: A Sequential Mediation Model, Journal of Health Communication, 20:11, 1302-1309
[11] Christopher E. Clarke, Graham N. Dixon, Avery Holton & Brooke Weberling McKeever (2015) Including "Evidentiary Balance" in News Media Coverage of Vaccine Risk, Health Communication, 30:5, 461-472
[12] Roozenbeek, J., van der Linden, S. Fake news game confers psychological resistance against online misinformation. *Palgrave Commun* 5, 65 (2019).
[13] University of Cambridge. "Fake news 'vaccine' works: 'Pre-bunking' game reduces susceptibility to disinformation." ScienceDaily. ScienceDaily, 24 June 2019. https://www.sciencedaily.com/releases/2019/06/190624204800.htm
[14] Cook, J. (2017). Inoculation theory: Using misinformation to fight misinformation. https://theconversation.com/inoculation-theory-using-misinformation-to-fight-misinformation-77545

EARLY DETECTION

Establishing routine social media monitoring activities is key to detect early signals of rumours before they become "rumor trending events" and attracted significant social attention[15]. Early signals are defined as patterns that appear well before rumours reach their peak time (generally speaking, weeks or months before). Early detection can help guide proactive content development to address community concerns and questions ahead and fill data voids before they are filled by misinformation. Appendix I provides an overview of available digital tools for routine social media monitoring. This section provides some key highlights that can help specifically with early detection:

- **News alerts**. Tools such as Google alerts allow to create automatic notifications based pre-defined search criteria and can help pick up early signals of online conversations around key COVID-19 issues.

- **Virality**. Several social media monitoring tools include a functionality that allows to identify stories that have the potential to "go viral", such as trending scores in Talkwalker[16], and viral alers and overperfoming stories in Crowdtangle[17].

---
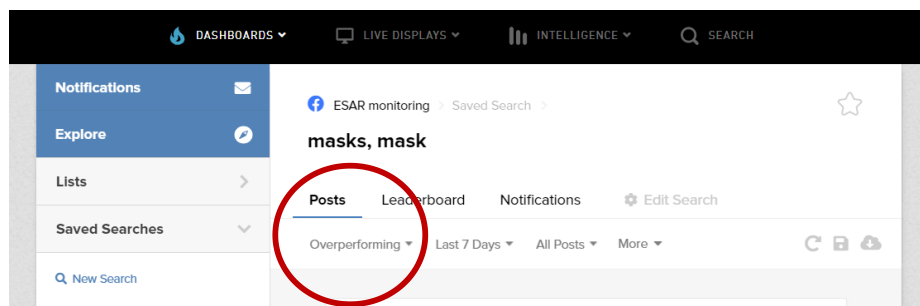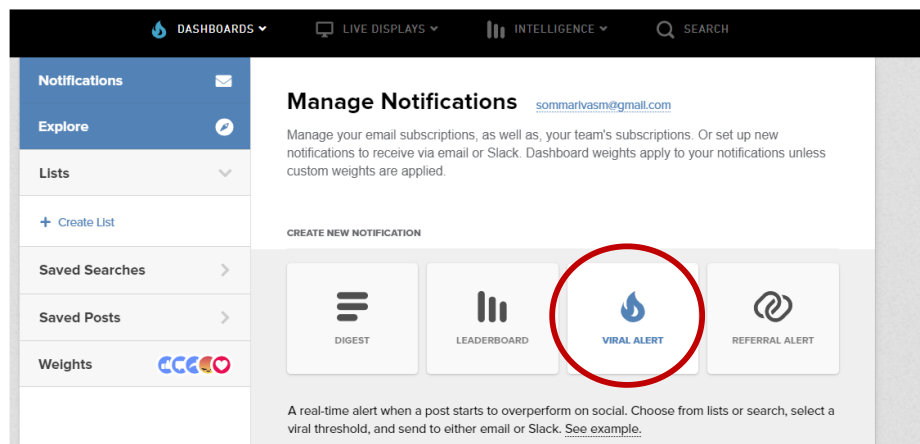
[15] Wang S., Mose I., Helbing D., Terano T. (2017). Early Signals of Trending Rumor Event in Streaming Social Media. 2017 IEEE 41st Annual Computer Software and Applications Conference

[16] UNICEF ESARO has raised a contract with Talkwalker for 12 months (June 2020 – May 2021)

[17] Crowdtangle. https://apps.crowdtangle.com/

- **User comments monitoring**. User comments below articles or social media post can provide insight into the main concerns, questions or pieces of misinformation circulating around a given topic. There are social listening tools that allow to monitor user comments in a systematic way. For example, UNICEF ESARO has developed a [dashboard](#) titled "Online conversation in Eastern and Southern Africa about Sexual and Reproductive Health and Rights in times of COVID-19" using Talkwalker to track online conversation around sexual and reproductive health and rights (SRHR) issues in the context of the pandemic, which includes a widget to monitor user comments. This dashboard is accessible to external partners through this [link](#).

- **Influencers.** Identifying who is contributing and leading the conversation is also a key part of rumour management, as influencers can be monitored and engaged to build a more accurate discourse. The figure below from the Talkwalker [dashboard](#) on "Online conversation in Eastern and Southern Africa about Sexual and Reproductive Health and Rights in times of COVID-19" shows an example of how to track influential authors and outlets posting content related to SRHR.

DIRECT ENGAGEMENT

Direct engagement with rumours can be warranted depending on the type of content being spread and its potential impact on outcomes such as adoption of preventive behaviors and access to services. This process of exposure of information that is false or misleading is also called de-bunking or fact-checking. There are two key decisions to be made around direct engagement with rumours: 1) does the rumour warrant de-bunking? 2) if so, what are key guidelines for engagement?
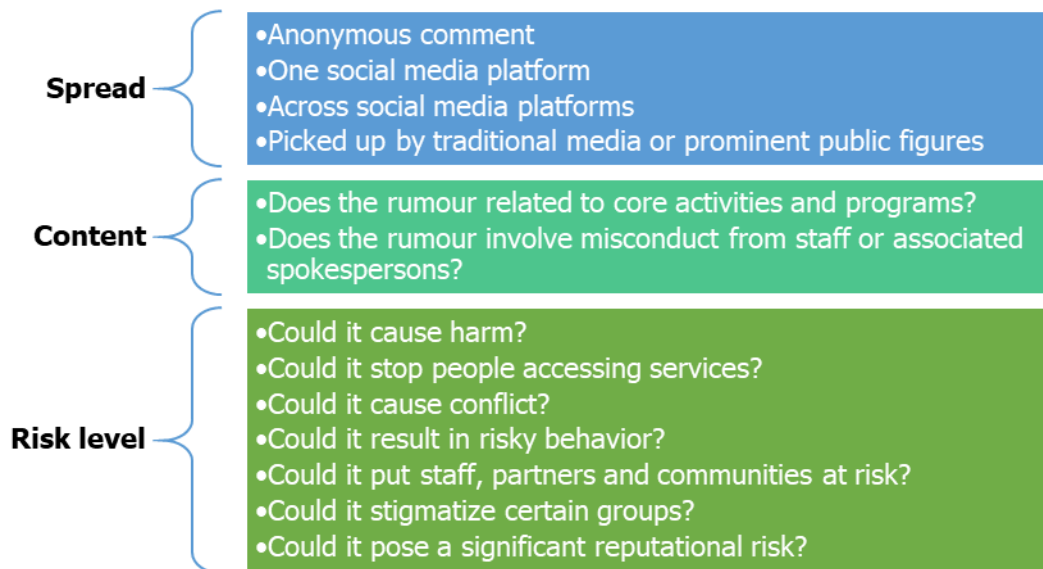
Deciding whether to engage with rumours

Majority of rumours may not warrant a direct response from UNICEF and its partners. Directly engaging with rumours to show they are not backed by evidence entails risks that need to be weighted, including the risk of alienating audiences or potentially reinforcing/amplifying the wrong information by giving it a platform[18]. Factors that come into play in deciding whether to engage with a piece of mis/disinformation include:

- **How widespread is the rumour?** A rumour that has been picked up by traditional media or being spread by public figures warrants considerations that can be different from a rumour that has only been shared by a few anonymous commenters online. However, even when a rumour has been seen only by a few users, it has the potential to create damage. Considerations around the risks associated with the rumour are therefore crucial to make an evaluation.

- **Is the topic related to key programs or activities?** Rumours that put the organization at reputational risk or directly refer to the activities and programs delivered by the organization warrant special consideration. For instance, a rumour about an issue that is within the scope of work (e.g. access to maternal care) may be addressed with higher priority compared to an issue where other partners are better positioned to intervene.

- **What risks are associated with the rumour?** Questions on whether the rumor poses concrete risks to different groups of the population are key. The practice guide "Rumour has it: a practice guide to working with rumours" by CDAC Network provides a good summary of these considerations[19].

---

[18] Phillips, W. (2018) Data and Society report. https://datasociety.net/wp-content/uploads/2018/05/FULLREPORT_Oxygen_of_Amplification_DS.pdf
[19] CDAC Network (2017). Rumour has it: a practice guide to working with rumours

Spread
- Anonymous comment
- One social media platform
- Across social media platforms
- Picked up by traditional media or prominent public figures

Content
- Does the rumour related to core activities and programs?
- Does the rumour involve misconduct from staff or associated spokespersons?

Risk level
- Could it cause harm?
- Could it stop people accessing services?
- Could it cause conflict?
- Could it result in risky behavior?
- Could it put staff, partners and communities at risk?
- Could it stigmatize certain groups?
- Could it pose a significant reputational risk?

Guidelines for engagement

Studies have shown that simply issuing a correction is often not effective to counter misinformation[20]. This is especially true when the correction clashes with individual values and belief system[21]. Evidence on what works and what doesn't when it comes to responding to misinformation is still growing, but a certain degree of consensus on key principles has already been achieved.[22,23,24,25,26]

- **Focus on the facts rather than on the rumour**. The correction should not reinforce the misinformation, and it should make very clear to the audience what the facts are. For example, using questions to introduce a factcheck is not considered a best practice as it instils doubts in users who may not click to read the full article. The example below uses a question in the title as a way to get the user to click on the link ("find out…"). This gives more prominence to the misinformation (e.g. "wearing a face mask reduces oxygen") and leaves the user wondering about whether there may be some truth to this argument.

---

[20] Nieminen, S., & Rapeli, L. (2019). Fighting Misperceptions and Doubting Journalists' Objectivity: A Review of Fact-checking Literature. Political Studies Review, 17(3), 296–309.

[21] Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation. Psychological Science, 28(11), 1531–1546.

[22] Cook J., Lewandowsky S. (2011) Debunking handbook. https://skepticalscience.com/docs/Debunking_Handbook.pdf

[23] Silverman C. (2020). Verification Handbook for Disinformation and Media Manipulation. http://verificationhandbook.com/book2/

[24] Paynter J., et al. (2019). Evaluation of a template for countering misinformation—Real-world Autism treatment myth debunking. PLOS ONE

[25] First Draft (2020). It matters how platforms label manipulated media. Here are 12 principles designers should follow. https://firstdraftnews.org/latest/it-matters-how-platforms-label-manipulated-media-here-are-12-principles-designers-should-follow/

[26] https://www.comminit.com/global/content/combating-misinformation-and-rumours-0

www.health.com › Infectious Diseases › Coronavirus ▾

Does Wearing a Face Mask Reduce Oxygen—and Can It ...

May 14, 2020 - Does wearing a **mask reduce oxygen** levels, and does it increase CO2 levels? Find out what experts say about carbon dioxide toxicity.

A better approach is to state the fact directly in the title, so as to avoid any confusion. In the examples below, the fact is given prominence in the title (e.g. "ginger does not prevent or cure COVID-19") and the false claim the article is addressing is reported below in smaller font ("ginger can treat a coronavirus infection").
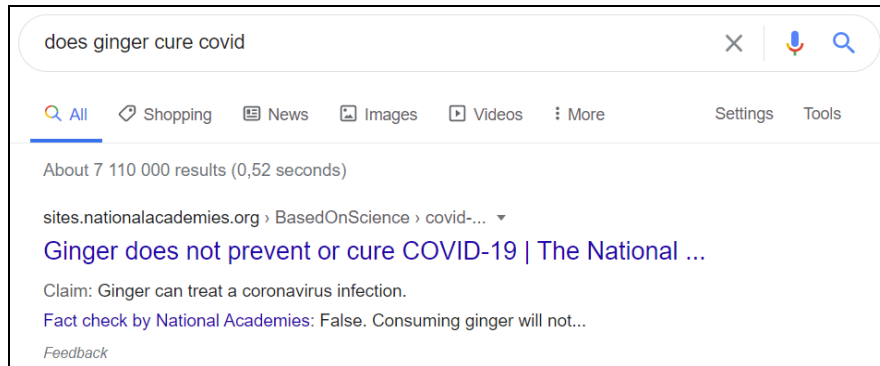

When possible avoiding negative sentences will reduce the likelihood of misunderstanding. In the example below, to respond to a rumour that COVID-19 spreads only below certain temperatures, the WHO Mythbuster provides the following fact "COVID-19 can be transmitted in areas with hot and humid climates". The negative sentence may be warranted depending on how directly one wants to engage with the rumour and there is mixed evidence as to whether repeating the misinformation in debunking is advisable in some cases[27].

---
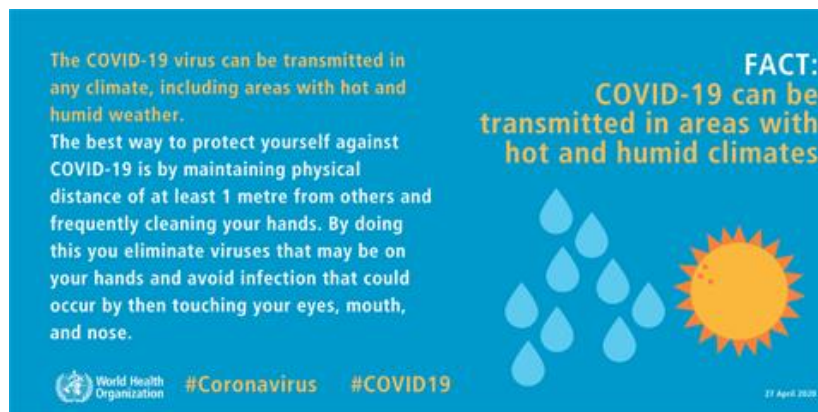
27   Ecker, U., Hogan, J., Lewandowsky, S. (2017). Reminders and repetition of misinformation: helping or hidering its retraction? https://www.sciencedirect.com/science/article/abs/pii/S2211368116301838

If there is an actual need to expose users to the misinformation in order to address it, then the best thing is to make sure that explicit warnings are set in place. Some considerations around labels can be found [here]. Experts have also called for a "truth sandwich"[28]: 1) start with the fact, as the first frame has the advantage, 2) indicate the misinformation, avoiding the language used to promote the misinformation, 3) return to the fact to make sure it gets more exposure compared to the misinformation.

- **Expose the misinformation techniques.** When false information is corrected, it leaves a gap in the mind of the user that will be filled with other information. To this end it is important to provide a line of reasoning of why the previous information was incorrect, showing the technique that was used to promulgate misinformation. The figure below from a study by Schmid & Betsch (2019) provides an overview of some techniques of science denial and associated rebuttal. In the example, the denial focuses on the fact that a certain vaccine is not 100% safe, and the rebuttal highlights that this is an "impossible expectation" that is placed on vaccines but not on other medical products, all of which can never be 100% safe[29].
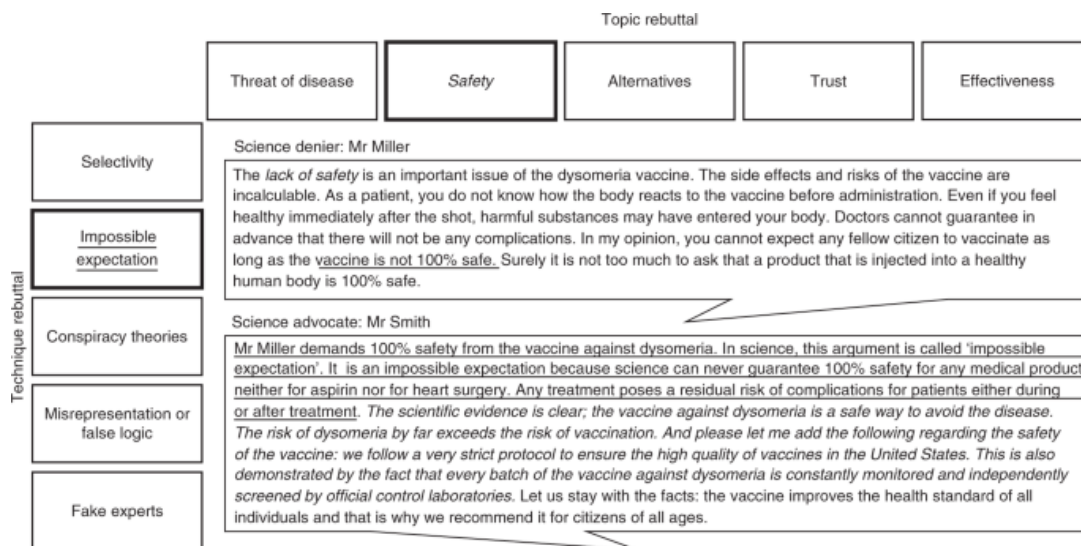
Source: Schmid, P., Betsch, C. Effective strategies for rebutting science denialism in public discussions. Nat Hum Behav 3, 931–939 (2019).https://www.nature.com/articles/s41562-019-0632-4?proof=true

- **Consider how to convey the message.** This aspect includes decisions around who should be offering the correction and how to frame the message. Identifying credible individuals and partners that are trusted by a certain audience can help

---

[28] https://twitter.com/georgelakoff/status/1068891959882846208?lang=en
[29] Schmid, P., Betsch, C. Effective strategies for rebutting science denialism in public discussions. Nat Hum Behav 3, 931–939 (2019).https://www.nature.com/articles/s41562-019-0632-4?proof=true

convey messages and clear doubts. A source's perceived trustworthiness is often more determinant than their expertise in changing users' misperceptions.[30]

Additionally, a main barrier to corrections is that, when corrected, individuals feel often compelled to argue in favor of the misinformation as a way to protect their perceived integrity. No one likes to be told they are wrong, but individuals are less likely to feel threatened by a correction when their self-confidence is reinforced. Therefore, using self-affirmation techniques when presenting audiences with information that challenges their beliefs may lead them to more openly accept that they were wrong.[31] The UN Share Verified initiative has produced content with example of conversations to correct misinformation shared by friends and relatives in a way that shows empathy and does not place blame on the individuals who shared the false content.
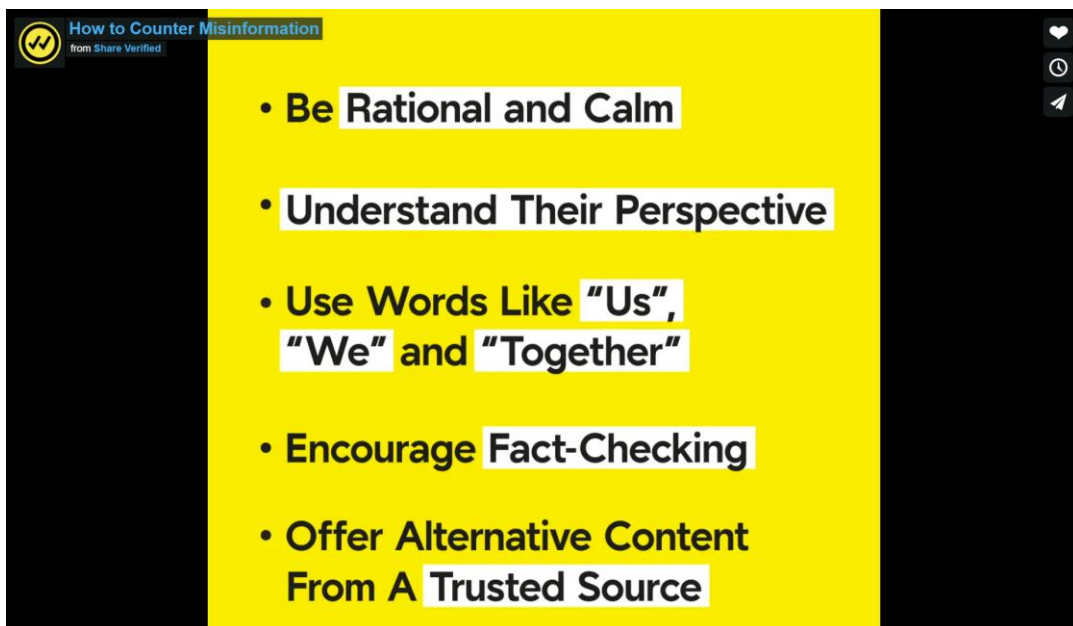
- **Make the content visually appealing.** Visuals are typically more effective to attract attention and help individuals process the information. Making sure the accurate content is more engaging and visually appealing than the misinformation it tries to address is key to increase its spread and impact.

---

[30] Guillory JJ, Geraci L. Correcting erroneous inferences in memory: The role of source credibility. J Appl Res Mem Cogn. 2013 Dec 1;2(4):201–9.

[31] Sherman, D. K., & Cohen, G. L. (2002). Accepting Threatening Information: Self–Affirmation and the Reduction of Defensive Biases. Current Directions in Psychological Science, 11(4), 119–123.

ROADMAP TO RUMOUR MANAGEMENT

The following are some easy-to-implement recommendations on how to enhance rumour management activities around SRHR and COVID-19 issues in ESAR:

RUMOUR PREVENTION

To prevent rumours from arising consider implementing the following actions:
- Check that the main recommendations around access to SRHR services and COVID-19 prevention strategies are up to date and consistent across all owned digital platforms. Availability of content in different languages is also key.
- Use findings from social listening activities (conducted at regional or national level) to inform decisions around content on social media and other digital platforms (e.g. sharing information on breastfeeding during COVID-19 when related search queries are rising on Google in the country).
- Check UNICEF and partners' content positioning in search engine results is prominent when it comes to issues on which the organization is leading the response.
- Collect/review data on media and information literacy and information sharing behaviors (e.g. how many times have users shared information on SRHR and COVID-19 without knowing if it is accurate?) to identify key priority groups for activities that build users' resilience against misinformation.

DETECTION

To quickly detect rumours and concerns use the "Online conversation in Eastern and Southern Africa about Sexual and Reproductive Health and Rights in times of COVID-19" dashboard to identify trending stories, key influencers and outlets that are sharing misinformation to further monitor their content.

CONTENT DEVELOPMENT

When developing content to address rumours and concerns (e.g. mythbusters):
- Ensure that facts are prominent and, if the rumour is repeated, that it is clearly labeled and cannot be confused with the fact.
- Carefully consider using questions in titles or posts that could potentially reinforce a rumour or raise additional confusion.

APPENDIX I: DIGITAL SOCIAL LISTENING TOOLS

AUTOMATIC NEWS ALERTS

Certain tools and platforms allow users to set automatic alerts to receive notifications when content meeting a certain set of criteria is published. One of these platforms is Google alerts, which allows to create a series of alerts for selected key words and receive notifications on the results with custom frequency and delivery times (6). Inserting combinations of key words that can signal COVID-19 misinformation, such as "Coronavirus AND garlic" or "Plandemic", within a region of interest helps surface problematic content. News aggregators such as Feedly.com (which has a free function) also allow to set automatic alerts for newly published content meeting certain characteristics.
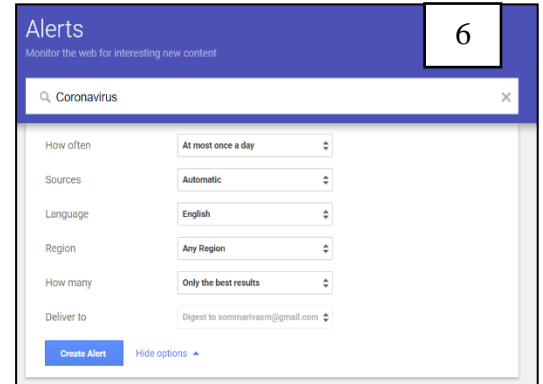

SOCIAL MEDIA MONITORING

The number of tools available to monitor social media platforms is constantly growing and existing applications continue to expand their functionalities. Most of these services do not require coding expertise and are extremely user friendly.

Crowdtangle allows to track information from Facebook, Instagram, Twitter and Reddit. The search function allows to select which platform to monitor and specify additional filters such as the timeframe, language, and whether the user has verified or unverified status. Note that the country filter is only available for Facebook Pages (7).
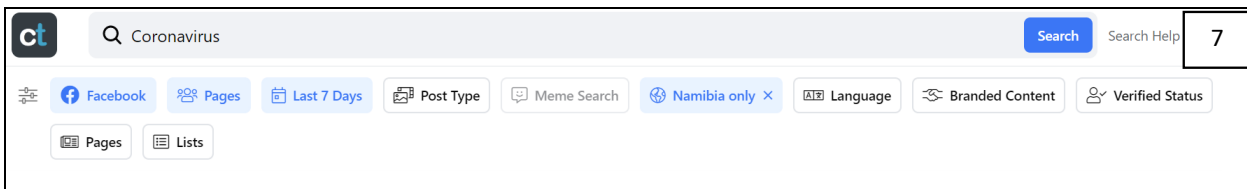

When using the "Search function" in Crowdtangle (8), the tool will visualize the number of interactions and posts relevant to the selected key words and criteria, showing a trendline over time, as well as top posts ranked by default by the number of interactions.
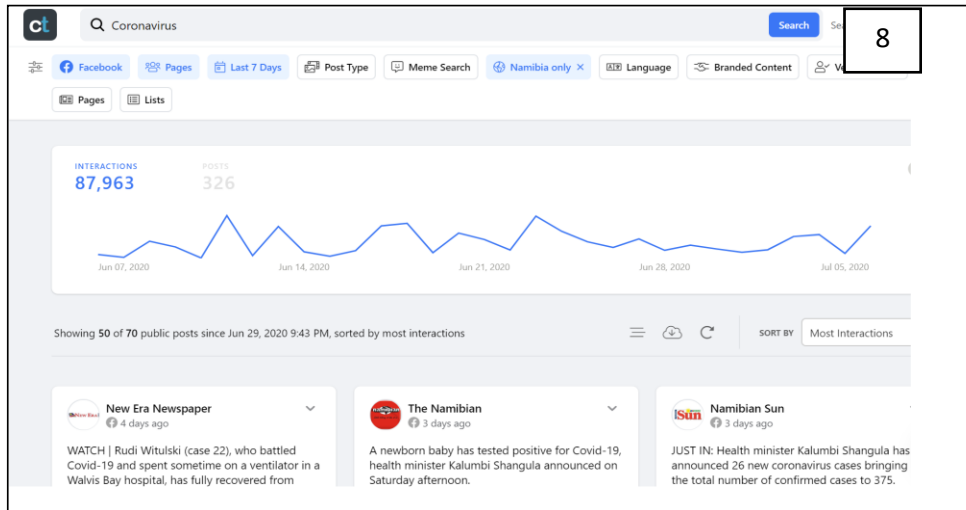
Under the "dashboard" section, the tool allows to create notifications for certain key words (9) with pre-defined frequency (e.g. daily, weekly etc.) This section also allows to save searches and select posts by weight. The latter functionality is particularly helpful when looking for posts that are generating highly positive or negative sentiment, as those may be likely to go viral.
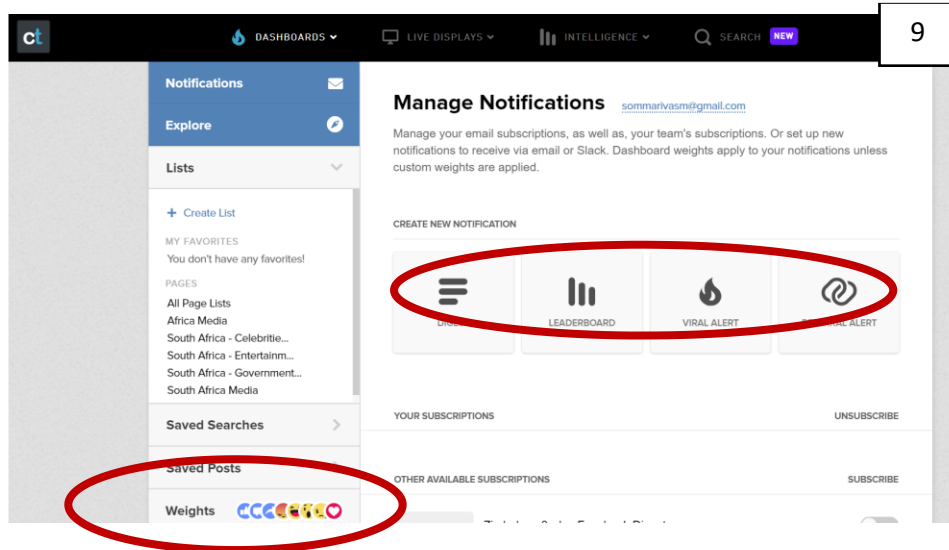
The Crowdtangle free plug-in is an additional tool that allows to check how often and by whom a certain link has been shared.
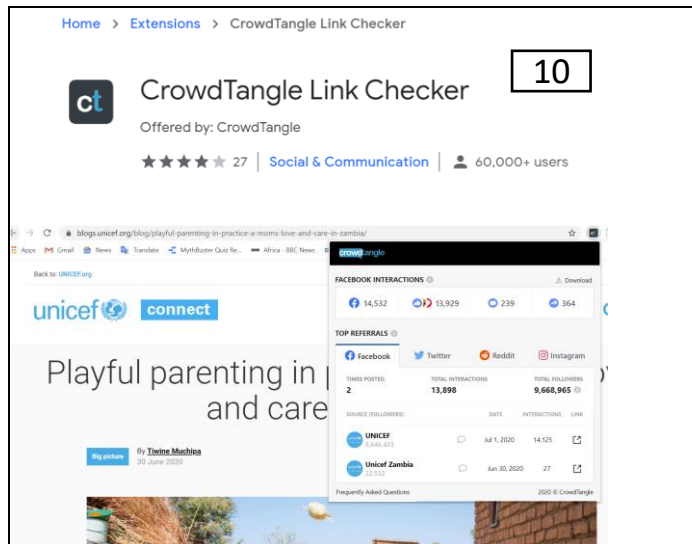
Additional tools to monitor content across platforms include:

- *Buzzsumo* is a social media content monitoring and analysis tool, which can be accessed with a free trial. The tool is particularly helpful to identify influencers and can be integrated with existing dashboards or internal reporting systems.

- *Mentionlytics* is a platform for social media monitoring, which can also be accessed with a free trial. The platform has an integration app that allows to combine it with Hootsuite functionalities.

- *Social Status* automates social media reporting and provides analysis and insights into owned social media accounts. Also accessible through free trial.

- *AgoraPulse* has a series of free tools available (the platform itself requires a subscription) such as the Facebook Page Barometer to analyze users' interaction with owned content, and the Twitter Report Card to compare performance to other partner accounts.

Besides Crowdtangle, tools to monitor **Facebook** include:
- *Ad Library* is a searchable collection of all ads running from across Facebook products, including Instagram. This guide produced by First Draft provides a good overview of how to use the library to monitor sponsored content on COVID-19. One suggestion is to select the country of interest in the homepage filter first, and then type keywords of interest (doing the other way round sometimes results in only a limited list of countries to select from) (11).
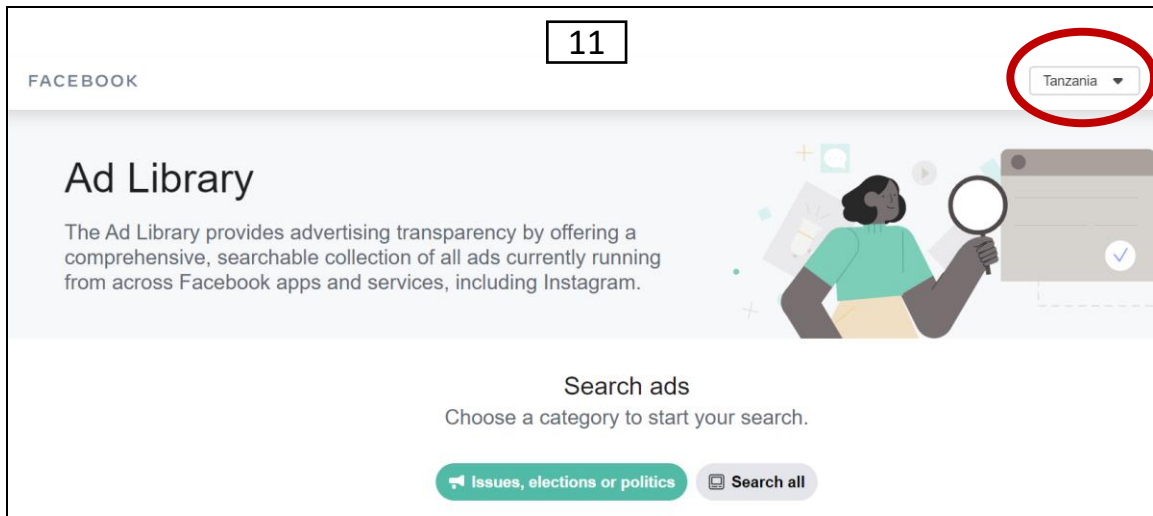
- *Graph.tips/beta* also allows to search Facebook public content by user, time, location.

- *Sociograph* can be used to monitor owned Facebook pages and groups.

- *Intelx.io* allows to search Facebook public content by key word and time, as well as search within posts of a specific page (12).

Tools to monitor **Twitter** include:
- *TweetDeck* is useful both to monitor owned accounts on the platform as well as content that is at risk of misinformation. The tool allows to visualize feeds of information for selected keywords or hashtags side by side and follow their evolution over time while filtering by engagement, language and media format.

- *Twitter list copy* is a tool that allows to merge Twitter lists. A list is a curated group of Twitter accounts. This tool is particularly helpful when it comes to misinformation monitoring because it allows to pull together accounts to follow into one feed.

- *Tweetbeaver* can be used to download information from Twitter, such as user's follower list or favorites. This functionality can be helpful in misinformation research if one has identified a suspicious account and wants to analyze which accounts that user is following and interacting with (which could lead to identify other suspicious accounts).

- *Hoaxy* (Beta version). This application (mainly drawing from US-based sources) visualizes the spread of claims and related de-bunks.

- *TweepsMap* tracks and analyzes conversations on Twitter (low monthly fee).

- *TrendsMapp* provides a free visualization of trending hashtags, words or accounts on a global scale (13).
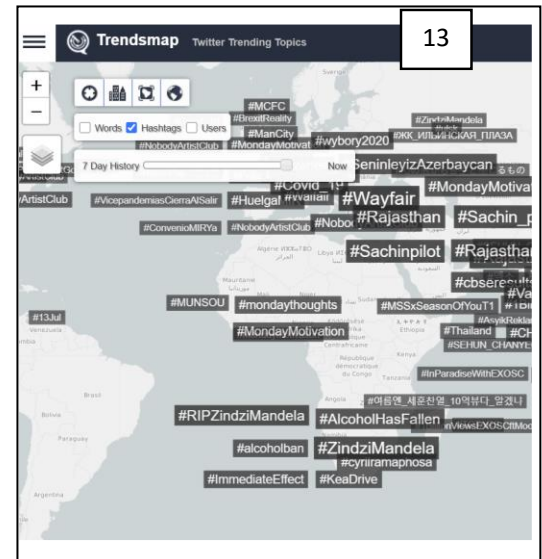
Tools to monitor **YouTube** include:
- *Channel subscription* allows to set up notifications on new videos from selected channels.

- *Google Trend Explore* has a YouTube search function that shows what users are searching on YouTube, by selecting "YouTube search" instead of "Web search".

Besides Crowdtangle, **Instagram** can be monitored using *Analisa*, which also allows to browse **Tik Toc** content (free account with limited functionalities). The hashtag search function can be helpful to track popular content on COVID-19 or related topics.
*Track Reddit* allows to monitor questions that are being asked on **Reddit** and provides the option to set up alerts when a specific search string appears in one of the question threads (14).
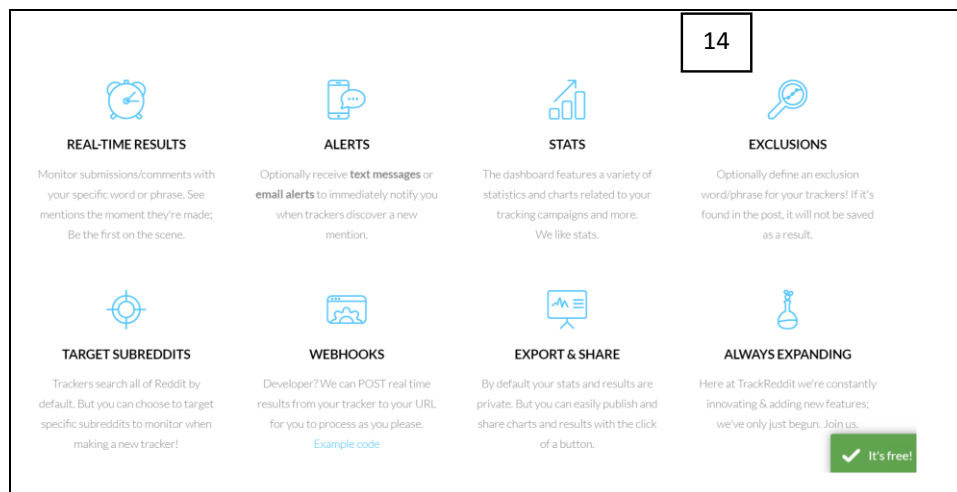
Note on Whatsapp monitoring. Whatsapp's end-to-end encryption makes it difficult to trace and monitor content. In collaboration with the International FactChecking Network, Whatsapp has launched a number of tip lines for users to forward content they believe could be false. PesaCheck and AfricaCheck are two organizations that have set up lines in ESA region and publish de-bunking articles to address misinformation circulating on the platform. Monitoring third-party reporting is a way to capture misinformation on this platform, although the time lag between the spread of the information and publication of fact-checks can be an issue.

MONITORING WEB SEARCHES AND DIGITAL NEWS

Search engines are one of the first places users go to when they have questions or concerns around a certain issue. Oftentimes, searches that are trending on engines can signal the presence of misinformation around a certain issue: users see something on social media, and they go on Google to verify whether what they read is true or find more information about it. _Google trends_ is a key tool that allows to track volume of searches for certain keywords on several channels including general web, Image search, New search and YouTube (as mentioned previously). It allows to compare key words (15) to explore whether searches on one topic are higher compared to another and understand trends in interest. Searches can be restricted by country and timeframe.
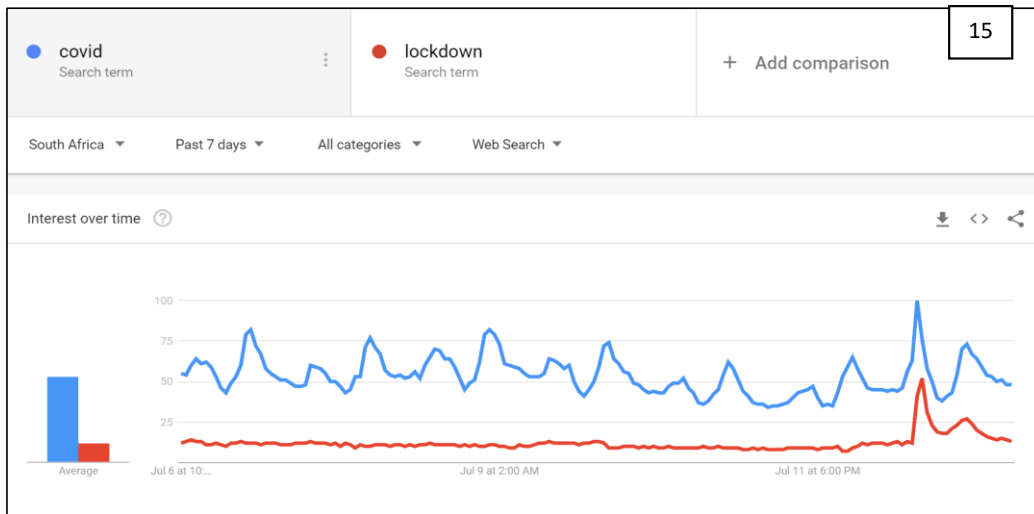
A helpful functionality of Google Trends is the "Related queries" section which shows searches related to a certain key word (in the example in figure 16, the keyword is "Covid" and was searched for Kenya, past month). Related queries can help understand which issues surrounding a certain topic people are interested in (in the example, rising searches have been related to symptoms). This can help understand information needs and inform content development strategies, as well as signal areas where misinformation can potentially arise.
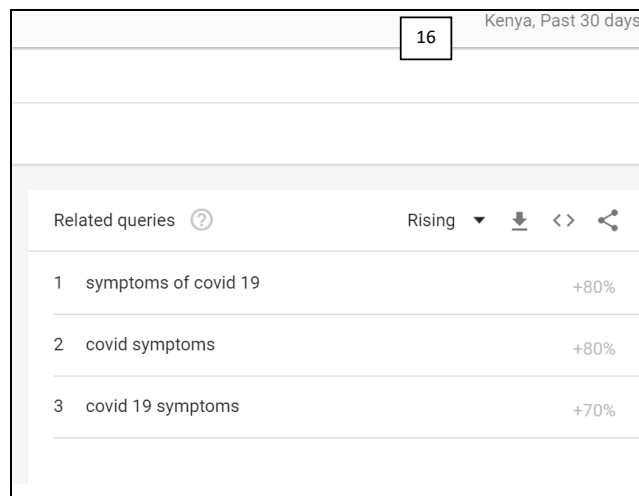
An additional tool available via Google is the *Dataset Search*, which allows to browse publicly available datasets that match keywords of interest.

TOOLS FOR GENERAL WEB MONITORING

There are tools that allow to download data from web pages for further analysis and identification of key insights. Two examples include *Data Miner*, a Chrome plug in that scrapes and downloads data from any site, and *Klaxon*, which tracks when a certain webpage is edited. Klaxon can be particularly helpful, for instance, to identify changes in recommendations from local public health authorities or other partners.
Manipulated images or real images taken out of context are a main channel through which misinformation is spread. Easy-to-use tools that allow for image verification include:

- *Google reverse image search* allows to upload images to track related pictures on Google Images.

- *Tineye* works similarly to Google reverse search. The Chrome extension allows to quickly see previous versions of the same photo.

- *Yandex* is a Russian search engine with a reverse image search function.

- *RevEye* is Chrome extension that combines Yandex, Tineye, Google Images for reverse image search.

Sorting results from these tools by date of publication allows to track, for instance, whether an older image is being used out of context (see an example fact-checked by AfricaCheck, where an image of President Ramaphosa attending a party in 2018 was being presented as recent to argue the President was violating social distancing regulations).

ADDITIONAL RESOURCES

- Cook, J., Lewandowsky, S. (2011). The Debunking Handbook. https://skepticalscience.com/docs/Debunking_Handbook.pdf
- Saltz, E., Shane, T., Kwan, V., Leibowicz, C., Wardle, C. (2020). It matters how platforms label manipulated media. Here are 12 principles designers should follow. https://medium.com/swlh/it-matters-how-platforms-label-manipulated-media-here-are-12-principles-designers-should-follow-438b76546078
- Schmid, P., Betsch, C. Effective strategies for rebutting science denialism in public discussions. Nat Hum Behav 3, 931–939 (2019). https://www.nature.com/articles/s41562-019-0632-4?proof=true
- First Draft. Free course: Too much information: How to separate the helpful from the harmful about coronavirus
- United Nations. Share Verified campaign and resources.
- AfricaCheck (2016). How to stop falling for fake news. https://africacheck.org/factsheets/guide-stop-falling-fake-news/
- UNESCO. "Journalism, fake news and disinformation'. Handbook for Journalism Education and Training.
- The Full Fact toolkit: simple practical tools anyone can use to identify bad information.
- International Federation of Red Cross and Red Crescent Societies (IFRC). Addressing mistrust and denial of COVID-19 in communities Guidance for African National Societies.