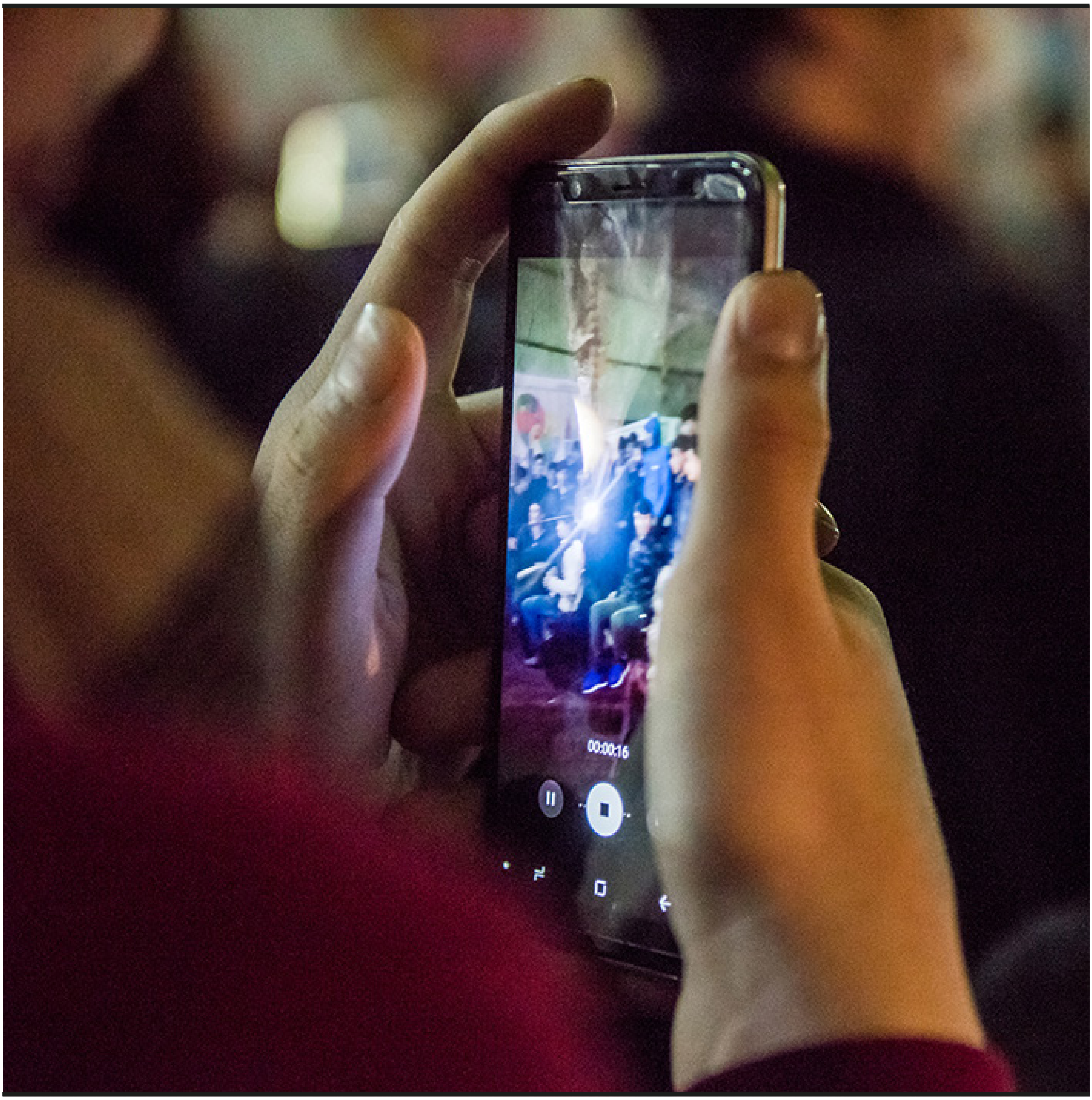
the
Humanitarian Leader

(Mis)communication? Social listening and the exclusion of marginalised voices

IRENE SCOTT



CENTRE FOR
HUMANITARIAN
LEADERSHIP

MARCH 2022

THE HUMANITARIAN LEADER:
(Mis)communication? Social listening and the exclusion of marginalised voices

WORKING PAPER 023
MARCH 2022

Author Details

Irene Scott
Irene is a former journalist who specialises in building humanitarian information systems to connect crisis affected communities with quality information to inform decision making and influence humanitarian response efforts. She is currently the Senior Community Engagement Project Manager at Internews.

Acknowledgements

I would like to thank Em Winters for their constant support and guidance in refining and analysing social media data throughout the pandemic and the Internews Humanitarian team for encouraging me to continue to question our approaches and aim to be better. Big thanks also goes to the hardworking teams at both UNICEF and AIRA for sharing their social listening reports for review in this thesis.

Editorial Office

The Centre For Humanitarian Leadership
Deakin University
221 Burwood Highway
Burwood 3125 VIC
Australia

Editor in Chief: Dr Nazanin Zadeh-Cummings
Production Editor: Cara Schultz
ISSN: 2653-1011 (Online)

The Centre for Humanitarian Leadership acknowledges the Australian Aboriginal and Torres Strait Islander peoples of this nation. We acknowledge the traditional custodians of the lands on which we work. We pay our respects to ancestors and Elders, past and present.

Copyright

CENTRE FOR
HUMANITARIAN
LEADERSHIP

Abstract

This article aims to contribute to the growing scholarship on the use of social media by humanitarian organisations in a crisis. Although social media's role in times of crisis has been rigorously studied, much of this work looks at the distribution or collection of information by first-responders or relief organisations. However, there is a growing interest in the analysis of social media content to understand community perceptions and to guide public health and risk communication interventions. This article aims to explore some key limitations of data collected using Social Media Analytics (SMA) tools in fairly representing community-wide perceptions. Through a review of 'social listening reports' produced by UN bodies and international aid organisations, this article will explore whether these data deficiencies are fairly represented. This article concludes that while there are many well documented limitations in the use of social media discourse to holistically represent community perceptions, these limitations are not adequately discussed in the reporting produced from this data. Consequentially, users of this analysis cannot adequately weigh the quality of the data when using it to influence policy decisions.

Leadership relevance

This paper aims to fuel discussion in the humanitarian sector over the ethical use of technology in the sector. Far from condemning the use of technology, I aim to encourage practitioners to understand the benefits and limitations of these approaches and to foster transparency in the sector. Focusing on the emerging field of 'infodemiology', this paper comes at an important time in the COVID-19 response, when after two years of working at a breakneck speed, practitioners are looking back on their efforts to reduce health-related misinformation, responding to community information needs, and taking a critical look at the development and impact of emerging approaches and tools.

Introduction

In the last 20 years, social media platforms have grown from a novelty to a critical form of communication and engagement worldwide (Obar, 2015). As internet penetration grows and data costs drop, these user-centric platforms designed to help people connect and communicate have flourished. They are driven by the idea that by sharing our preferences, emotions, or even pictures of our lunch, we are building an increasingly important virtual community (Noveck et al, 2021).

In an emergency, users go to these virtual networks to request and share information, locate loved ones, and find community in crisis.

In an emergency, users go to these virtual networks to request and share information, locate loved ones, and find community in crisis (Appling et al, 2014). Increasingly, these virtual communities are being used by social science researchers to try and understand people's beliefs and perceptions. In this way, our online lives are directly influencing the policy decisions made for us in our offline lives. And while there is certainly merit in using these vast data sets for research, in this article I will explore the limitations of this approach, in particular, when using Social Media Analytics (SMA) tools. Understanding the limitations of any data set is vital in being able to weigh its relevance in any research or policy decision (Ross & Zaidi, 2019). My hypothesis is that by presenting this kind of data as an accurate depiction of community-wide insights, without a nuanced discussion of limitations, there is the potential to misrepresent community perceptions and to further silence and marginalise vulnerable groups.

Methodology and limitations

I will begin by exploring the available literature related to the use of social media in disaster and crisis contexts by humanitarian agencies. I will then explore a non-exhaustive list of the main data quality limitations of SMA tools being employed during the COVID-19 pandemic by public health professionals and risk communicators. Finally, I will use these limitations as a metric to assess a selection of social listening reports created to inform the risk communication priorities of humanitarian agencies. Making the assumption that this kind of analysis aims to drive actionable intelligence, I will explore how a lack of transparency about data limitations could be misleading, or impact the ability of social listening reports and other outputs to meet this aim.

There are challenges in defining 'social media'. For the purposes of this paper, I will define social media as a web-based application designed to help two or more people communicate, connect and share user-generated content (Kietzmann et al, 2011). For ease, I will focus much of this paper on three commonly used social media platforms in humanitarian settings: Facebook, WhatsApp and Twitter (Walker, 2017). Facebook launched in 2004 and is the world's largest social networking site with an estimated 2.89 billion monthly active users, Twitter is a microblogging network which launched in 2006 and has an estimated 192 million users and WhatsApp is an instant messaging service launched in 2009 with an estimated two billion monthly users (Statista, 2021 and Albergotti et al, 2014). There are of course many other social media platforms available, but these will not be discussed due to the limitations of this paper. However, many of these platforms warrant further research, especially in contexts where apps such as YouTube, Instagram or Telegram may have greater community penetration and impact.

The scope of this research is limited by language and by the use of secondary data. I have only included studies, both academic and grey literature, produced in English. While this likely accounts for a significant portion of available research on this topic, it is expected that there may be relevant findings in research published in languages other than English that are excluded. This paper is primarily the result of secondary research—relying on already published papers, studies, and outputs. Key Informant Interviews (KIIs) would likely have contributed to a richer understanding of the processes and priorities that influence decision making in the creation of social listening reports and the data quality mitigation strategies employed. As this is an emerging field of practice, greater research is needed to properly document and analyse the challenges, successes and evolving methodologies used by practitioners in this field.

The use of social media in humanitarian settings

Disasters are socially experienced, and the increasingly prominent role social media plays in a disaster—as people share their experiences, advice and sometimes heartbreak—means that social media presents a goldmine of data for researchers.

Disasters are socially experienced, and the increasingly prominent role social media plays in a disaster—as people share their experiences, advice and sometimes heartbreak—means that social media presents a goldmine of data for researchers (Oh et al, 2013). Social media is a powerful tool in identifying a crisis; for example, in 2013 the first reports of the Boston marathon bombing (Cassa et al, 2013), and the Westgate Mall Attack in Kenya (Simon et al, 2014) were published first on Twitter, well before major news networks could share the information. It can be used to map the impacts of a natural disaster (Vieweg et al, 2010), or direct first responders towards victims (Lindsay, 2011).

The collection and analysis of social media data to influence public health interventions has gained popularity in recent years (see Dashtian, 2021; Hou, 2021; Hossain, 2016 and Broniatowski, 2018). There are a huge number of works aiming to understand the benefits of various approaches—just one literature review from the World Health Organisation (WHO) uncovered more than 130 articles (Chanely, 2021). In contrast, there are well reported data quality issues, such as credibility and representative bias that may impact the use of this data in an emergency (see Duarte et al, 2018; Yang et al, 2021). However, there remains a gap in the research to understand how these limitations are being communicated to the humanitarian community during the COVID-19 pandemic.

The prevalence of social listening during COVID-19

In an effort to coordinate communication and engagement activities happening in response to the pandemic in humanitarian settings, a network of Risk Communication and Community Engagement (RCCE) working groups and taskforces were launched at the local, regional and global levels. Chaired by the WHO, the United Nations Children's Fund (UNICEF), and the International Federation of Red Cross and Red Crescent Societies (IFRC), these platforms group together like-minded bodies, including governments (ministries of health, etc), local and international NGOs, and civil society to address the "infodemic" (Zarocostas, 2020). While the term "infodemic" has been in use for more than 20 years, it has grown to prominence during the pandemic, notably through its use by the WHO to refer to "an overabundance of information—some accurate and some not—occurring during an epidemic" (WHO, 2020).

Prevention measures such as government mandated curfews and other restrictions on movement and gathering have made it difficult for communities to engage in ways they previously did, and it has also often put humanitarians at a distance from the communities they serve. For the safety of their teams, many field-based activities have been limited (Nutbeam, 2021; Plexico-Sinclair, 2020) and so it is natural the sector should turn to social listening; a form of listening that can be conducted remotely (Gilmore et al, 2020).

Social listening in a humanitarian context can be defined broadly as the process of monitoring and analysing community conversations in online spaces (such as social media) to understand needs and inform humanitarian responses (Stewart, 2018; Hou, 2021). There is a growing body of social science research that aims to better understand how social media data can be used to study people's sentiments and attitudes as an alternative to self-reported surveys (Appling et al, 2014). For example, researchers look to social media to understand how communities share information in an emergency (Simon et al, 2015; Cohen, 2013) or understand behaviours related to the spread of misinformation (Pasquetto & Jahani, 2020; Bowles et al, 2021). During the pandemic, social listening data is being used to understand people's public health perceptions.

During the pandemic, social listening data is being used to understand people's public health perceptions.

These RCCE fora have become the natural platforms for the sharing of these social listening insights. Not every organisation may have the resources to perform social listening, and so these coordination mechanisms allow insights to be shared among member agencies, usually with accompanying risk communication guidance. The communal nature of these reports intensifies the importance of a transparent discussion on the limitations of data so that members can make informed policy decisions.

Why social listening data is problematic

It is estimated that there are 500 million tweets sent every day (Rao et al, 2013). Such a huge volume of data would be impossible, if not impractical, to manually collect and analyse, so SMA tools are employed. These tools, designed to track brand insights and contribute to commercial marketing strategies, have been redeployed during the pandemic to understand sentiments related to COVID-19, vaccines, and trust in authority figures (Dashtian & Murthy, 2021). These automated systems work by using Artificial Intelligence (AI) to collect and categorise publicly available social media data in vast quantities (Gonçalves, 2017). The speed at which these tools can turn huge data sets into appealing visualisations has made them particularly attractive to time-poor humanitarian agencies. However, the rush to adopt this methodology may result in agencies not fully understanding, mitigating or being able to communicate the limitations of the data. A non-exhaustive set of limitations is explored below.

Digital divide: Who is represented in the data?

One well documented issue with the use of social media data to understand community-wide perceptions is equitable access (Ragnedda et al, 2013; Landers, 2017). In every country, there are people who either choose not to, or simply do not have access to social media. This disparity in access to the internet, mobile phones or computers across socioeconomic groups has been dubbed the "digital divide" (Brown et al, 1995). For example, even in America, which is one of the leading countries in digital innovation, more than 100 million Americans do not use social media at all (Perrin & Anderson, 2019; Wojcik & Adam, 2019). Compare this to countries with far greater challenges in achieving digital penetration. In Afghanistan, for example, just 9% of the population are social media users—predominantly young, urban and educated professionals (Orfan, 2020), and only 16% are women (Rai, 2019). Because of these stark limitations, any assessment of social media data could only include the perceptions of this elite, capital-centric, gender skewed portion of society. In social media metrics worldwide, women are underrepresented, as are elderly populations, people living with disabilities and low-income groups (Hargittai, 2015). As a consequence, if this data is used to inform the design of humanitarian responses, we risk designing responses based on the needs of the privileged, while further marginalising and disproportionately censoring vulnerable groups.

In social media metrics worldwide, women are underrepresented, as are elderly populations, people living with disabilities and low-income groups.

Language: Do dominant languages drown out marginalised voices?

Language can be another barrier in social media analysis. In most countries, the discourse on social media is held in the dominant language, or lingua franca (Hoffmann et al, 2017). English has become the default lingua franca for social media. Research into Twitter usage in Africa found that 77% of content originating from countries in Africa was in English, with Arabic and French featuring at only 10% combined (Winhill, 2018). When using AI technology, the user or researcher chooses the languages they wish to use to search and relies on the ability of that tool to understand the target languages needed for their analysis. This is done through Natural Language Processing (NLP), which involves a computer program 'learning' a language by absorbing millions of strings of data (for example, sentences) in that language (Johansson et al, 2016). Firstly, the user may choose to only search in dominant languages, so there is potential for selection bias. In addition, while these data sets are plentiful for the world's dominant languages, they may not exist for minority languages. This can mean that the tool may not identify posts in those languages at all, struggles to 'understand' or apply thematic categorisation, or has limited ability to comprehend the nuance of social media data in these languages (Duarte et al, 2018). This drop in accuracy can, for example, impact the recognition of tone that is required to understand if a post is a joke or threatening or dangerous content (Hirschberg & Manning, 2015).

This limitation is a distinct challenge for the analysis of Arabic social media text and the use of Arabizi; a form of Arabic which uses Latin letters and numbers to reproduce Arabic language that has been popularised by younger users (Darwish, 2014; Bies et al, 2014). In the context of NLP, this usually requires this data to be transliterated from Arabizi to Arabic script, or requires the system to be specifically trained to understand these complex mixed datasets (Guellil, 2021; Talafha et al, 2021). As a consequence, a SMA system that cannot understand or does not recognise this form of content risks excluding young voices.

It is evident that when working with SMA tools errors can occur—just as a human interpreter may misinterpret information. Policymakers must understand the capabilities and limits of these tools in regards to language, particularly for making decisions that could impact on the efficacy of a public health response.

What is captured: Public versus private posting

SMA tools work by scraping huge amounts of publicly available data from social media platforms. Their ability to pull in such immense amounts of data could distract some users from questioning exactly what kind of data is being captured. Publicly available data refers to posts and interactions that can be seen by anyone, without the need to 'friend', 'follow' or join a particular platform or group (Ravn, 2019; Markham, 2012). Twitter is a good example of a platform where the majority of the posts are public—just 13% of Twitter users in the United States choose to make their profile private (Remy, 2019). Twitter is a goldmine of data for researchers, and it accounts for a large portion of social media research due to the ease of extracting data, but it's important to remember that each platform may attract a specific demographic of users (Simon et al, 2015). In many nations, for example, Zimbabwe, South Sudan and South Africa, Twitter is dominated by political and social elites or the diaspora community (Windhill, 2018). In addition, the very public nature of the platform could make it intimidating for some users to openly engage with it (Salvatore et al, 2020). A study by the Pew Research Center in 2019 found that most users rarely tweet and a small group of prolific users (just 10% of accounts) were responsible for 80% of English language tweets. In addition, they found that Twitter users in the US were more likely to be young, highly educated, earn above average incomes and vote Democrat (Wojick et al, 2019). This combination of readily accessible data from a potentially narrow portion of the community can present a skewed picture of a society and its perceptions.

This combination of readily accessible data from a potentially narrow portion of the community can present a skewed picture of a society and its perceptions.

Despite its global popularity, there are distinct restrictions in terms of what SMA tools are able to capture from Facebook. In 2018, access to Facebook data was heavily restricted following the Cambridge Analytica scandal, when user profiles were used to direct political advertising (Zimmer, 2010). SMA tools can only capture information from posts made on a limited number of registered Facebook Pages (pages are a type of profile used by businesses, politicians, celebrities and media). Posts made on your personal profile, on friend's pages or within Facebook Messenger cannot be collected (Yang et al, 2021). Of the billions of posts made daily on the platform, only a tiny percentage could ever be analysed by SMA tools.

WhatsApp data is even more problematic for SMA tools and privacy restrictions mean they cannot access any data from this platform at all. However, WhatsApp accounts for a huge share of the social media market, especially in emerging markets, where its adaptability to low bandwidth and voice message features make it an attractive tool for users with unreliable data connections or low literacy (Berman, 2019). In addition, researchers believe the tool may be responsible for the spread of a significant amount of misinformation (see Broniatowski, 2021; Lazer, 2018 and Davies, 2020). Again, we see that the data collected by these tools presents an incomplete picture of the social media discourse that may be happening around the pandemic or other issues of interest to researchers or practitioners.

Who are you really: Unreliable demographic data

A challenge impacting all analysis of social media data, either via AI or through manual collection, is the difficulty in determining the authenticity of the users who post. A significant portion of the posts shared on social media are thought to come from either social or malicious bots or from troll farms (Dotto, 2020). Social bots are accounts controlled by autonomous software, designed to impersonate real users (Kenworth, 2019). Troll farms are organised operations, where workers are employed to manage fraudulent social media accounts to generate online traffic aimed at affecting public opinion (Snider, 2018). Malicious accounts have been found to post more often, with content that is more politically divisive than the average social media user (Broniatowski et al, 2021). Some research suggests that nearly half of the accounts posting about the pandemic on Twitter in the United States and the Philippines are bots (Uyheng, 2020) and that a minority of accounts and pages were responsible for the majority of pandemic related misinformation (Yang et al, 2021). While humans manually collecting social media content may be able to recognise an inauthentic account, SMA systems treat all content equally. This presents an opportunity for malicious actors to flood a particular context, influencing online discourse, and consequentially the social listening reports and policy decisions taken by humanitarian actors.

A challenge impacting all analysis of social media data, either via AI or through manual collection, is the difficulty in determining the authenticity of the users who post.

A further difficulty comes in trying to determine location and demographic data such as age or gender. Sex and age disaggregated data is important in all research to gain a comprehensive understanding on the most affected groups and their unique needs (O'Mathuna et al, 2017). When users post on most social media platforms, they may volunteer their location (through profile information or a location tag in the post), or geolocalise—that is, allow their device to share their location (Appling, 2014). However, users have the ability to tag the post as being anywhere in the world and in a crisis, it also allows users to 'pretend' they are in an affected area (Utomo et al, 2018; Wiegmann et al, 2021). For instance, research into the use of Twitter during ten elections in African nations in 2017 found that 53% of the most active posters were not even in the countries where the elections were contested (Winhill, 2018). SMA tools also struggle with determining location. For instance, when other location information is unavailable, many SMA tools automatically categorise a user location based on the language used in the post. For instance, a user posting in English who has not volunteered their location is categorised as being in the US, posts in Spanish are automatically considered to be from Spain and tools consider Arabic posts to be from Saudi Arabia (Talkwalker, 2022).

Other demographic data like age and gender is similarly problematic. AI systems determine these demographic identifiers by collecting information users include in their public profile. Again, while some users may simply prefer not to include this information, others may choose to include false information to intentionally mislead, or as a kind of practical joke (for example, a teenager posting their age as 100 years old) (Wiegmann et al, 2021). The challenges in accessing reliable disaggregated data from social media users limits what researchers can infer from this data and may contribute to generalisations about community perceptions, concerns and needs.

What do you think: How opinions are shared on social media

While we may be interacting online more than ever, research suggests that people interact and share thoughts and opinions differently in online spaces. Research by the Pew Institute suggests that when issues are particularly controversial, people may be less likely to share their opinion online than they would in person. They found only 42% of Facebook and Twitter users were willing to post about a sensitive issue, while more than 80% would have an in-person conversation (Hampton et al, 2014). In both online and offline contexts, people expressed that they were more likely to express their opinion if they felt their friends or followers might be likely to agree with them (Hampton et al, 2014). This aligns with both the social theories of 'group think' (which suggests that people will irrationally choose to adopt the opinion of the 'group' to support harmony) and with the 'silence spiral' (which suggests that group members will withhold a contradictory opinion to avoid being ostracised) (Noelle-Neuman, 1974). Determining public perceptions solely from social media data ignores the fact that people use these platforms in differing ways, and that social media may not be the fora they choose to share their opinions about challenging issues that are of interest to researchers such as political discourse, perceptions or behavioural insights (Hargatti, 2015).

Determining public perceptions solely from social media data ignores the fact that people use these platforms in differing ways, and that social media may not be the fora they choose to share their opinions about challenging issues that are of interest to researchers such as political discourse, perceptions or behavioural insights.

What are you looking for: The impact of researcher generated search terms

One challenge in social media analysis is the impact of the researcher themselves on the research. When working with AI tools, the researcher is asked to input a series of queries, or keyword searches, that allow them to narrow down the billions of data points available from social media platforms (Simon et al, 2015). This narrowing down is important to allow for data to be analysed in an efficient manner (you can't reasonably look at everything) but it also has the consequence of limiting the data to the researcher's priorities. For example, in Dashtion's research on the social discourse around the pandemic on Twitter, they collected 19 million tweets that contained the words 'coronavirus', 'covid' or 'mask' (Dashtion, 2020). This approach might yield a high number of data points, but does not capture the whole discourse on the topic. For example, if a social media post talked about the 'pandemic' instead of 'covid' that data would not have been collected. Keyword searching has the potential to miss content using local slang or common spelling mistakes (Appling et al, 2014). The selection of these terms introduces a natural bias where research is guided by the researchers, rather than the community's priorities. While arguably this is a limitation present in many approaches to research, when striving to understand 'community perceptions', we should aim to limit the mitigate impact of the researcher's priorities.

Analysis: Social listening reports

An analysis of current social listening reports will contribute to a better understanding of how humanitarian agencies are communicating the limitations of social media data collected using SMA tools. I will analyse two reports:

- COVID-19 *Infodemic Trends in the African Region* is a weekly social listening report created by the Africa Infodemic Response Alliance: a regional network hosted by the WHO that brings together fact-checking and media organisations, and non-governmental organisations (WHO, 2020). See Annex 1 for the report details.
- *Social Listening report on COVID-19 Vaccination in Morocco* is a weekly social listening report created by the UNICEF Communication for Development staff in the UNICEF Maroc office. See Annex 1 for the report details.

The reports will be analysed using the criteria discussed above:

- Demographic: how does the report address limitations in the demographic makeup of the data including age, gender and location?

- Language: What language/s does the report include? How does the report address limitations in the collection of that language?
- Source: What social media platforms are included in this data set? How does the report address limitations in the data included from these platforms?
- Search approach: How does the report address limitations in keyword search approach used?

These reports are likely to have been presented in RCCE coordination meetings (or similar) and may have included further verbal discussion of the limitations and benefits of the data. However, for the purposes of this research, only the published report will be reviewed.

COVID-19 Infodemic Trends in the African Region

Demographic

After a review of the content included in seven AIRA reports, it is clear that demographic information that might allow for a more actionable response to social listening data is missing. Neither the age, nor the gender of the data sources was mentioned in any of the reports analysed. Some location data is included—the report states in the introduction that it includes data from Kenya, Nigeria, South Africa, Ivory Coast, Burkina Faso, Senegal, Democratic Republic of Congo, Niger, and Mali[1]. A further breakdown of this location data, for instance whether the information has been collected predominantly from rural or urban populations is not provided. In addition, the report also includes country-level social media observations from Lesotho, Ghana, Cameroon, Benin, Namibia, Malawi, Mauritius, Tanzania, Zimbabwe, Reunion, and Uganda, despite these countries not being mentioned as "Target Countries" in the introduction.

Language

Some language data is included in these reports. In the report introduction, a language code is used next to the names of some target countries which, it is assumed, refers to the target language of collection. For instance, we know that English language data has been collected in South Africa (the code "EN" is written next to the name "South Africa") and French data has been collected from Niger (the code "FR" is written next to the word "Niger"). No further information is provided in regards to the source data language. This is a limited sample—there are a potential 860 languages spoken in the target areas—though of course not all languages may be commonly used for social media discourse (Ethnologue, 2021). Dominant languages are often used in social media discourse, however the language chosen by the social media user may also signal other demographic traits such as whether the user has had access to secondary education or is considered of a higher social class. Because of this, it is important to disclose what languages make up social media data, and ideally provide demographic breakdowns, so that the reader of the social listening report is able to more clearly identify whether data is likely to have been collected from the average citizen, or the educated and elite classes.

Source

According to the methodology section of the AIRA reports, the reports are produced using "NewsWhip Analytics, TweetDeck, Crowdtangle, UNICEF Talkwalker dashboards as well as the WHO EARS platform" (AIRA, 2021). While this is an extensive list of mostly AI supported SMA tools, the report does not provide a clear breakdown of what social media platforms (or other online sources) are included in this analysis. However, report authors do provide a short sentence addressing limitations in the kind of data that can be extracted from each platform; "...data may be biased towards data emerging from formal news outlets/ official social media pages, and does not incorporate content circulating on closed platforms (e.g. Whatsapp) or groups (e.g. private Facebook groups)". It is arguable whether this brief description would be enough to inform a novice reader, but AIRA should be congratulated for making an attempt to address this key limitation.

Search approach

The report does not provide any information on the methodology used to search for this data other than listing the tools used (as explained above). There are other questions about their methodology that are also not addressed adequately. For instance, their definition of "trends" and the process for determining which trends are addressed in this report is not clear. Are the trends referring to the issues that have received the most individual posts, or do they refer to individual posts that have received a high level of engagement? One example post from Benin included in *Weekly Brief—September 13, 2021* had seen only 16 comments and six reactions as of 27 September, 2021. This could reasonably be considered a very low level of engagement with the post.

Review conclusion

While the AIRA team does take some steps to explain the limitations in their methodology in regards to data sources, a more nuanced approach to demographic data and in particular, language, is required to make this report a more practical tool to inform risk communication responses. While this element may be lacking, a positive and practical feature of these reports is the inclusion of guidance for practitioners

[1] Mali is mentioned as a Target Country in only one of the reports analysed: *Weekly Brief— May 17, 2021.*

near the end of the report. In a section titled, "Why is it concerning?", they provide brief insight into the social and behavioural impact of these perceptions, accompanied by another question: "What can we do?", where practical risk communication advice is offered. These sections contribute greatly to the likelihood of the report leading to practical policy and communication actions. However, questions remain surrounding how practical a report of this kind can be considering the diverse range of countries it attempts to address.

Social Listening report on COVID-19 Vaccination in Morocco

Demographic

Each report includes a breakdown (text and graphical) of gender, age and parental status. There is no mention of any limitations in the SMA tool's ability to determine these demographic features. In addition, "Key Takeaways" from the analysis are not presented with reference to the age, gender or parental status of the users that hold these beliefs. This gives the impression that these are the dominant beliefs across a homogenous community.

The analysis of "Family Status" is particularly unique. While it seems relevant in terms of UNICEF's mandate to advocate for and meet the needs of children (UNICEF, 2021), the SMA tool's ability to determine parental status is questionable considering this is not a key feature of data volunteered as part of social media user profiles.

At the beginning of the report, the authors state, "Location—Country of the search: Morocco", suggesting that this is the limitation they have put on the data collected through the SMA tool. While this should, in theory, exclude Moroccan diaspora or malicious actors posting from other countries, location data can be easily manipulated (as discussed above). The reports provide a breakdown of the main cities the data is collected from, with an overabundance of data coming from Rabat in the three reports where location is mentioned (no location data is provided in reports four and five). In report one, the authors briefly note a potential limitation in their location data: "We notice that most of the Data are from the Region of Rabat Salé. This may make us think of the Data collection issues. Maybe the French language or the lack of data from the other region".

Language

Four of the five reports reviewed included clear language data. While in report one, the authors only searched for Arabic, French and English data, in report 2-4 they widen their search to include "all languages", which results in the collection of a small amount of data in Spanish, Catalan, Korean, Indonesian, Thai and Hungarian.

In the first bulletin, concerns about the accuracy of language data are mentioned: "French is the language the most used. This may just mean that the platform is better at generating Data in French than other languages, namely Arabic". There are some other possible explanations for the overabundance of French data; the SMA tool may struggle to recognise Moroccan Arabic, which is the dominant language used in social media discourse (Abdouli et al, 2016). Another explanation could be spelling mistakes, or that some of the Arabic data is written in Arabizi or another form of transliteration (Abdouli et al, 2016). The source of the data is mainly Twitter, and French may simply be the dominant language for discussions on medical issues or for Moroccan users on this platform.

In report three and four, Arabic is noted as the dominant language (49% and 50.6% respectively), suggesting that either the demographic discussing COVID-19 has shifted or, more likely, the authors' approach to using the tool (perhaps using different settings) has shifted over time—presumably to help mitigate these limitations.

Source

These reports clearly display the source of their data in text and graphical form. In the second report, the authors point to a limitation related to this: "As this figure shows, more than 40% of the information is from Twitter and less than 1% from FB, a platform that is widely used in Morocco. This is one of the shortcomings of Talk Walker. The focal point suggests adding pages from FB manually to help get some results from FB." According to social media statistics, there are more than 17 million Facebook users in Morocco and close to 74% of Moroccans who have internet access are registered WhatsApp users (another data source not represented in this data) (Sasu, 2021). As the authors mention, in order to capture data from Facebook, the user must manually add Facebook pages they wish the tool to capture data from. This is a time-consuming process, and another possible introduction of significant researcher bias.

Search approach

While it is unclear what specific keyword searches are used, "themes" are mentioned in all five reports, and these themes are presumably made up of a series of related keyword searches.

In the first report, limitations in the search results are briefly mentioned: "For the minister of health, we get articles related to other ministries, and the minister of health may not have been mentioned in the article". As well as, "For SinoPharm, the sino which means in Spanish 'only' appears in the results". It is not clear whether the identified errors resulted in the problematic data being removed from analysis or whether these limitations are mentioned simply to explain why analysis may be inaccurate.

Review conclusion

The UNICEF report authors do make clear attempts to discuss the limitations present in data collection and analysis, however these mentions are very brief and often do not provide enough information as to *why* the issue may be occurring and *how* they may change their approach to address the problem. The use of data visualisations is very helpful (in particular when representing gender, location and topic data) and may contribute to a larger audience being able to interact with these reports. However, while the UNICEF reports make more attempts to be transparent on data limitations, in comparison to the AIRA reports, they provide far less risk communication guidance (just 2-3 brief dot points in reports 1-3 and none in four or five. The reports are focused on *what* people are talking about, which is interesting from an academic perspective, but this approach may impact the potential for the report to take the leap from interesting data to something actionable that genuinely influences risk communication responses and policy decisions.

Conclusion

One of the main advantages of social media data is that continuous updates allow real-time monitoring of public moods and sentiments. While this is an appealing prospect for researchers, this kind of data is not without its limitations and this has implications for the types of conclusions one can draw from data derived from these platforms. As discussed in this paper, it is our responsibility as scholars to ensure the limitations of any data set are understood and clearly communicated to the audience. By undertaking an assessment of 12 sample social listening reports produced by International NGOs and UN Agencies, it is clear that if these limitations are evident to the researchers, they are not being adequately communicated in the outcomes of this research.

As the aim of these social listening reports is to influence humanitarian policy and risk communication approaches, this deficit risks decisions being inadvertently made on imperfect or misrepresented data. While necessity dictates that humanitarians often make decisions based on imperfect data, it is important the users of the data are aware of the potential deficiencies and can make an informed decision of how data will be used with those limitations in mind. The need for transparency about data does not disappear just because it is collected from a social media platform. If anything, as this is an increasingly important data source that practitioners may be unfamiliar with using, it is even more important to clearly spell out any risks, limitations and concerns and for humanitarian organisations to encourage transparency in their own data and from others. Further research is needed to assess the actions taken as a result of these reports and how practitioners understood and accounted for limitations and what impact the analysis had on policy and programming.

The need for transparency about data does not disappear just because it is collected from a social media platform.

References

Abdouli, A., Chaffai, A., Hassouni, L., Anoun, H., Khalid, R. (2016). Current trends in Moroccan Social Networks, March 2016. *International Journal of Computer Science and Information Security*, 14(3). https://www.researchgate.net/publication/324839604_Current_Trends_in_Moroccan_Social_Networks

Albergotti, R., MacMillan, D., Rusli, E. (2014). Facebook to Pay $19 Billion for WhatsApp. *Wall Street Journal*. pp. A1, A6. ISSN 0099-9660. https://www.wsj.com/articles/facebook-to-buy-whatsapp-for-16-billion-1392847766

Appling, S., Briscoe, E., Ediger, D., Poovey, J., McColl, R. (2014). *Deriving Disaster-Related Information from Social Media*. Georgia Tech Research Institute. https://www.academia.edu/10282640/Deriving_Disaster-Related_Information_from_Social_Media

Berman, H. (2019, 20 June). What's up with WhatsApp? How WhatsApp is used and misused in Africa. *The Economist* https://www.economist.com/middle-east-and-africa/2019/07/18/how-whatsapp-is-used-and-misused-in-africa

Bies, A., Song, Z., Maamouri, M., Grimes, S., Lee, H., Wright, J., Strassel, S., Habash, N., Eskander, R., Rambow, O. (2014). *Transliteration of arabizi into Arabic orthography: developing a parallel annotated arabizi-arabic script sms/chat corpus*. Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP), pp. 93–103. https://aclanthology.org/W14-3612

Bowles, J., Larreguy, H., Liu, S. (2021). Countering formation via WhatsApp: Preliminary evidence from the COVID-19 pandemic in Zimbabwe**.** *PLoS ONE*. 10.14.2020. pp. 1-11. DOI: 10.1371/journal.pone.0240005.

Broniatowski, D., Jamison, A., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S., Dredze, M. (2018). *Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate*. Am J Public Health. DOI:10.2105/AJPH.2018.304567.

Brown, R., Barram, D., Irving, L. (1995). *Falling through the net: a survey of the "have nots" in rural and urban America*. US Department of Commerce, National Telecommunications and Information Administration. https://www.ntia.doc.gov/ntiahome/fallingthru.html

Cassa, C., Chunara, R., Mandl, K., Brownstein, J (2013). Twitter as a Sentinel in Emergency Situations: Lessons from the Boston Marathon Explosions *PLoS Currents*, no. JUNE. DOI: currents.dis.ad70cd1c8bc585e9470046cde334ee4b.

Chanely, S., Benjamin, P., Mechae, P. (2021). *Finding the Signal through the Noise: A landscape review and framework to enhance the effective use of digital social listening for Immunization demand generation*, Health Enabled with technical guidance and support from Gavi, the Vaccine Alliance, UNICEF and WHO. https://www.gavi.org/sites/default/files/2021-06/Finding-the-Signal-Through-the-Noise.pdf

Cohen, S. E. (2013). *Lessons Learned: Social Media and Hurricane Sandy*. Virtual Social Media Working Group and DHS First Responders Group. https://www.dhs.gov/sites/default/files/publications/Lessons%20Learned%20Social%20Media%20and%20Hurricane%20Sandy.pdf>

Darwish, K. (2014). *A Panoramic Survey of Natural Language Processing in the Arab World*. Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP), pp.217–224., https://eds.b.ebscohost.com/eds/detail/detail?vid=40&sid=63fecaea-5a6a-4c76-a561-abb6e3ed2c12%40pdc-v-sessmgr02&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=149459802&db=aci

Dashtian, H., Murthy, D. (2021). CML-COVID: A large-scale covid-19 twitter dataset with latent topics, sentiment and location information. *Academia Letters*, Article 314. https://doi.org/10.20935/AL314.

Davies, W. (2020, 2 July). What's wrong with WhatsApp? *The Guardian online.* https://www.theguardian.com/technology/2020/jul/02/whatsapp-groups-conspiracy-theories-disinformation-democracy

Dotto, C., Cubbon, S. (2021). Disinformation exports: How foreign anti-vaccine narratives reached West African communities online. *First Draft.* https://firstdraftnews.org/long-form-article/foreign-anti-vaccine-disinformation-reaches-west-africa/

Duarte, N., Llanso, E., Loup, A. (2018). *Mixed Messages? The Limits of Automated Social Media Content Analysis*. 2018 Conference on Fairness, Accountability, and Transparency. https://cdt.org/wp-content/uploads/2017/12/FAT-conference-draft-2018.pdf

Ethnologue. (2021). Country language profiles for: Kenya, Nigeria, South Africa, Ivory Coast, Burkina Faso, Senegal, Democratic Republic of Congo, Niger, and Mali. https://www.ethnologue.com/browse/countries>

Gilmore, B., Ndejjo, R., Tchetchia, R., de Claro, V., Mago, E., Diallo, A., Lopes, C., Bhattacharyya, S. (2020). Community engagement for COVID-19 prevention and control: a rapid evidence synthesis. *BMJ Global Health* 2020;5**:**e003188. https://gh.bmj.com/content/5/10/e003188

Goh, Dion Hoe-Lian., Chua, Alton Y. K., Shi, Hanyu., Wei, Wenju., Wang, Haiyan., Lim, Ee-peng. (2017). *An analysis of rumor and counter-rumor messages in social media, Digital libraries: Data, information, and knowledge for digital lives*. 19th International Conference on Asia-Pacific Digital Libraries, ICADL 2017, Proceedings, pp. 256-266. Research Collection School Of Information Systems. https://ink.library.smu.edu.sg/sis_research/3875

Goncalves, A. (2017). *Social media analytics strategy: using data to optimize business performance.* APress, https://eds.a.ebscohost.com/eds/detail/detail?vid=1&sid=4

Guellil, I., Saâdane, H., Azouaou, F., Gueni. B., Nouvel, D. (2021). Arabic natural language processing: An overview. *Journal of King Saud University - Computer and Information Sciences*, 33(5): pp. 497-507. https://doi.org/10.1016/j.jksuci.2019.02.006

Hampton, K.. Rainie, L., Weixu, L., Dwyer, M., Shin, I., Purcell, K. (2014). *Social Media and the 'Spiral of Silence'*. Pew Research Center. https://www.pewresearch.org/internet/2014/08/26/social-media-and-the-spiral-of-silence/

Hargittai, E. (2015) Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites. *Annals of the American Academy of Political and* **Social** *Science*. 659(1): pp. 63-76. DOI: 10.1177/0002716215570866

Hirschberg, J., Manning, C. (2015). Advances in Natural Language Processing. *Science 261.* American Association for the Advancement of Science, Stanford University. https://cs224d.stanford.edu/papers/advances.pdf.

Hoffmann, C., Bublitz, W. (2017). *Pragmatics of Social Media*. De Gruyter Mouton, Berlin. https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1595340&authtype=sso&custid=deakin&site=eds-live&scope=site

Hossain, L., Kam, D., Kong, F., Wigand, R.T., Bossomaier, T. (2016). Social media in Ebola outbreak *Epidemiol Infect*. 144(10): pp. 2136-43. DOI: 10.1017/S095026881600039X.

Hou, Z., Tong, Y., Du, F., Lu, L., Zhao, S., Yu. K., Piatek, S.J., Larson, H.J., Lin, L. (2021). Assessing COVID-19 Vaccine Hesitancy, Confidence, and Public Engagement: A Global Social Listening Study. *Journal of Medical Internet Research*. 23(6): pp. e27632. https://eds.a.ebscohost.com/eds/detail/detail?vid=4&sid=8a33421d-e516-4d8c-ac59-9c83af16a7f8%40sessionmgr4007&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#db=mdc&AN=34061757

Johansson, F., Kaati, L., Sahlgren, M. (2016). Detecting Linguistic Markers of Violent Extremism in Online Environments. *Combating Violent Extremism and Radicalization in the Digital Era*. pp. 374–90. https://www.foi.se/download/18.3bca00611589ae7987820d/1480076542059/ FOI-S--5452--SE.pdf

Kenworth, S. (2019). *How a Russian Troll Farm Aims to Upset the Current Hegemonic, Unipolar World Order*. DOI: 10.13140/RG.2.2.32736.74248. https://www.researchgate.net/publication/333092238_How_a_Russian_Troll_Farm_Aims_to_Upset_the_Current_Hegemonic_Unipolar_World_Order

Kietzmann, J.H., Hermkens, K., McCathy, I.P., Silvestre, B.S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*. 54(3): pp. 241-25. https://www.sciencedirect.com/science/article/abs/pii/S0007681311000061

Lazer, D.M.J., Baum, M.A., Benkler, Y., Berinsky, A.J., Greenhill, K.M., Menczer, F. (2018). The science of fake news. *Science*. 359(6380): pp. 1094–1096. https://doi.org/10.1126/science.aao2998

Landers, C. (2017). *The Digital Divide: Issues, Recommendations and Research*. Hauppauge, New York, Nova Science Publishers, Inc. https://eds.a.ebscohost.com/eds/detail/detail?vid=4&sid=4ace7791-db24-43d3-97b7

Liebermann, M. (2013). *Social: why our brains are wired to connect*. Oxford, Oxford University Press. https://eds.b.ebscohost.com/eds/detail/detail?vid=32&sid=63fecaea-5a6a-4c76-a561-abb6e3ed2c12%40pdc-v-sessmgr02&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=deakin.b3460813&db=cat00097a

Lindsay, B. (2011). *Social Media and Disasters: Current Uses, Future Options, and Policy Considerations*, Congressional Research Service report for Congress. https://sgp.fas.org/crs/homesec/R41987.pdf

Mancosou, M., Vegetti, F. (2020). What You Can Scrape and What Is Right to Scrape: A Proposal for a Tool to Collect Public Facebook Data. *Journal of Social Media and Society*. 6(3). https://doi.org/10.1177/2056305120940703

Markham, A., Buchanan, E. (2012). *Association for Internet Researchers' Ethics Working Committee—Ethical decision-making and Internet research*: Version 2.0. Association of Internet Researchers https://aoir.org/reports/ethics2.pdf

Murthy, D., Gross, A.J. (2016). Social media processes in disasters: Implications of emergent technology use, *Social Science Research*. DOI: 10.1016/j.ssresearch.2016.09.015.

Noelle-Neumann, E. (1974). The Spiral of Silence: A Theory of Public Opinion. *Journal of Communication*. 24(2): pp. 43-51. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1460-2466.1974.tb00367.x

Noveck, B., Button, J., Gambrell, D., Paulson. L., Spada, P., Subramanian, L. (2021). *The Power of Virtual Communities*. The Gov Lab. https://virtual-communities.thegovlab.org/files/DTR_report_en_EN.pdf

Nutbeam, D. (2021). The vital role of meaningful community engagement in responding to the COVID-19 pandemic. *Public health research and practice*. DOI: https://doi.org/10.17061/phrp3112101.

Obar, J., Wildman S. (2015). *Social media definition and governance challenge*, University of Ontario Institute of Technology and Michigan State University, working paper. https://d1wqtxts1xzle7.cloudfront.net/38260771/Obar_Wildman_GOSM_TP_Editorial_July_22_2015-with-cover-page-v2.pdf

Oh, O., Agrawal, M., Rao, R. (2013). Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets During Social Crises. *Management Information Systems Quarterly*. 37(2): pp. 407-426. https://aisel.aisnet.org/misq/vol37/iss2/7/

O'Mathuna, D., Siriwardhana, S. (2017). Research ethics and evidence for humanitarian health. *The Lancet*. 390(10109):pp. 2228-9. DOI: https://doi.org/10.1016/S0140-6736(17)31276-X

Orfan, S. (2020). Political participation of Afghan Youths on Facebook: A case study of Northeastern Afghanistan. *Cogent Social Sciences*. 7(1) https://www.tandfonline.com/doi/full/10.1080/23311886.2020.1857916

Pasquetto, I., Jahani, E. (2020). *Mobile Instant Messengers (MIMs) in Developing Countries*, Shorenstein Center on Media, Politics and Public Policy, Harvard Kennedy School. https://shorensteincenter.org/misinformation-on-mims

Perrin, A., Anderson, M. (2019). *Share of US adults using social media, including Facebook, is mostly unchanged since 2018*, Pew Research Center. https://pewrsr.ch/2VxJuJ3

Plexico-Sinclair, H., Vestrheim, K. (2020). *Engaging Communities during a Pandemic: Experiences of Community Engagement during the COVID-19 Response in Camps and Out-of-Camp Settings*, Norwegian Refugee Council. https://reliefweb.int/sites/reliefweb.int/files/resources/engaging-communities-during-a-pandemic.pdf

Ragnedda, M., Muschert, G. (2013). *The Digital Divide : The Internet and Social Inequality in International Perspective*, Routledge Advances in Sociology Series, Taylor & Francis Group. https://ebookcentral-proquest-com.ezproxy-f.deakin.edu.au/lib/deakin/detail.action?docID=1221501

Rai, A. (2019). Digital Divide: How Do Women in South Asia Respond?. *International Journal of Digital Literacy and Digital Competence*. 10(1). DOI: 10.4018/IJDLDC.2019010101

Ravn, S., Barnwell, A., Neves, B. (2019). What Is "Publicly Available Data"? Exploring Blurred Public–Private Boundaries and Ethical Practices Through a Case Study on Instagram. *Journal of Empirical research on Human research ethics*. 15(1-2): pp. 40-45. DOI: https://doi.org/10.1177/1556264619850736

Remy, E. (2019). *How public and private Twitter users in the U.S. compare - and why it might matter for your research*. Pew Research Center. https://medium.com/pew-research-center-decoded/how-public-and-private-twitter-users-in-the-u-s-d536ce2a41b3

Rosler, P., Mainka, C., Schwenk, J. (2018). *More is Less: On the End-to-End Security of Group Chats in Signal, WhatsApp, and Threema*. 2018 IEEE European Symposium on Security and Privacy. https://eds.b.ebscohost.com/eds/detail/detail?vid=13&sid=63fecaea-5a6a-4c76-a561-abb6e3ed2c12%40pdc-v-sessmgr02&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=edseee.8406614&db=edseee

Ross, P., Zaidi, N. (2019). Limited by our limitations. *Perspect Med Educ*. 8(4): pp. 261–264. DOI: 10.1007/s40037-019-00530-x

Salvatore, C., Biffignandi, S., Bianchi, A. (2020). Social media and twitter data quality for new social indicators. *Social Indicators Research*. 156(2-3): pp. 601-630. https://eds.b.ebscohost.com/eds/detail/detail?vid=35&sid=63fecaea-5a6a-4c76-a561-abb6e3ed2c12%40pdc-v-sessmgr02&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#db=her&AN=151632112

Sasu, D. (2021). *Number of active social media users in Morocco 2017-2021*, Morocco Country report. https://www.statista.com/statistics/1172771/number-of-social-media-users-morocco/

Simon, T., Goldberg, A., Aharonson-Daniel, L., Leykin, D., Adini, B. (2014). Twitter in the Cross Fire—The Use of Social Media in the Westgate Mall Terror Attack in Kenya. *PLoS ONE*. 9(8): e104136. DOI:10.1371/journal.pone.0104136.

Simon, T., Avishay, G., Bruria, A. (2015). Socializing in emergencies—A review of the use of social media in emergency situations. *International Journal of Information Management*. 35: pp. 609–619. https://www.academia.edu/20379786/Socializing_in_emergencies_A_review_of_the_use_of_social_media_in_emergency_situations

Snider, M. (2018, 16 February). Robert Mueller investigation: What is a Russian troll farm? *USA Today* https://www.usatoday.com/story/tech/news/2018/02/16/robert-muellerinvestigation-what-russian-troll-farm/346159002/

Statista Analytics. (2021). *Most popular social networks worldwide as of July 2021, ranked by number of active users*, Q3 2008 to Q2 2021 industry report. Statista Research Department. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

Stewart, M., Arnold, C. (2018). Defining Social Listening: Recognizing an Emerging Dimension of Listening. *International Journal of Listening*. 32: pp. 85–100. DOI: https://doi.org/10.1080/10904018.2017.1330656.

Talafha, B., Abuammar, A., Al-Ayyoub, M. (2021). ATAR: Attention-based LSTM for Arabizi transliteration. *International Journal of Electrical & Computer Engineering*. 11(3): pp. 2327-2334 https://eds.a.ebscohost.com/eds/detail/detail?vid=29&sid=4ace7791-db24-43d3-97b7-66badfb05b47%40sessionmgr4008&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=148685260&db=aci

UNICEF. (2021). *UNICEF mission statement*. https://www.unicef.org/about-us/mission-statement

Utomo, M., Adji, T., Ardiyanto, I. (2018). *Geolocation prediction in social media data using text analysis: A review*. 2018 International Conference on Information and Communications Technology (ICOIACT). DOI: 10.1109/ICOIACT.2018.8350674.

Uyheng, J., Carley, K. (2020). Bots and online hate during the COVID-19 pandemic: case studies in the United States and the Philippines. *Journal of Computational Social Science*. 3(2): pp. 445-46. https://eds.b.ebscohost.com/eds/detail/detail?vid=20&sid=63fecaea-5a6a-4c76-a561-abb6e3ed2c12%40pdc-vsessmgr02&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=edssjs.36F3AF94&db=edssjs

Vashisha, N., Zubiaga, A. (2021). Online Multilingual Hate Speech Detection: Experimenting with Hindi and English Social Media. *Information*. 12(1). (https://eds.a.ebscohost.com/eds/detail/detail?vid=24&sid=4ace7791-db24-43d3-97b7-66badfb05b47%40sessionmgr4008&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=148316730&db=lls

Vieweg, S., Hughes, A., Starbird, K., Panel, L. (2010). *Microblogging During Two Natural Hazards Events: What Twitter May Contribute to Situational Awareness*. CHI 2010: Crisis Informatics April 10–15, 2010. http://www.pensivepuffin.com/dwmcphd/syllabi/insc547_wi13/papers/microblog/vieweg.et.al.TwitterAwareness.CHI10.pdf

Walker, T. (2017). *Humanitarian futures for messaging apps: understanding the opportunities and risks for humanitarian action*, ICRC. https://reliefweb.int/sites/reliefweb.int/files/resources/4299_002_Humanitarian-Futures-for-Messaging-Apps_WEB_.pdf

Wiegmann, M., Kerston, J., Senaratne, H., Potthast, M., Klan, F., Stein, B. (2021). Opportunities and risks of disaster data from social media: a systematic review of incident information. *Natural Hazards and Earth System Sciences*. 21: pp. 1431-1444. https://eds.b.ebscohost.com/eds/detail/detail?vid=42&sid=63fecaea-5a6a-4c76-a561-abb6e3ed2c12%40pdc-v-sessmgr02&bdata=JmF1dGh0eXBlPXNzbyZjdXN0aWQ9ZGVha2luJnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=edsdoj.1a8e1c70ba6f4675a6353c1ed092632c&db=edsdoj

Winhill, L. (2018). *How Africa Tweets 2018*, Portland Communications. www.portlandcommuications.com/pdf/how-africa-tweets-2018.pdf

Wojcik, S., Messing, S., Smith, A., Rainie, L., Hitlin, P. (2018). *Bots in the Twittersphere*. Pew Research Center. https://www.pewresearch.org/internet/2018/04/09/bots-in-the-twittersphere

Wojcik, S., Adam, H. (2019). *Sizing up Twitter users*. Pew Research Center. https://pewrsr.ch/2VUkzj4

World Health Organization. (2020). *COVID-19 Global Risk Communication and Community Engagement Strategy*. https://www.who.int/teams/risk-communication/the-collective-service

World Health Organization. (2020). *Landmark alliance launches in Africa to fight COVID-19 misinformation*. https://www.afro.who.int/news/landmark-alliance-launches-africa-fight-covid-19-misinformation>

Yang, K., Pierri, F., Hui, P., Axelrod, A., Torres-Lugo, C., Bryden, J., Menczer, F. (2021). *The COVID-19 Infodemic: Twitter versus Facebook, Big Data and society*. Sage Journals. https://doi.org/10.1177/20539517211013861

Zaracostas, J. (2020). How to fight an infodemic. *Lancet*. 29(395): p. 676. DOI: 10.1016/S0140-6736(20)30461-X

Zimmer, M. (2010). But the data is already public: on the ethics of research in Facebook, Ethics and Information Technology. *Ethics and Information Technology*. 12(4): pp. 313-325. DOI: 10.1007/s10676-010-9227-5

Annex 1: Social Listening reports analysed

Author: Africa Infodemic Response Alliance		
1	*COVID-19 Infodemic Trends in the African Region*	17 May 2021
2	*COVID-19 Infodemic Trends in the African Region*	24 May 2021
3	*COVID-19 Infodemic Trends in the African Region*	5 July 2021
4	*COVID-19 Infodemic Trends in the African Region*	12 July 2021
5	*COVID-19 Infodemic Trends in the African Region*	19 July 2021
6	*COVID-19 Infodemic Trends in the African Region*	17 August 2021
7	*COVID-19 Infodemic Trends in the African Region*	13 September 2021

Author: UNICEF Morocco		
1	*Social Listening report on COVID-19 Vaccination in Morocco #1*	15-21 June 2021
2	*Social Listening report on COVID-19 Vaccination in Morocco #2*	22-28 July 2021
3	*Social Listening report on COVID-19 Vaccination in Morocco #3*	30 July-5 August 2021
4	*Social Listening report on COVID-19 Vaccination in Morocco #4*	6-12 August 2021
5	*Social Listening report on COVID-19 Vaccination in Morocco #5*	2-9 September 2021