Topic Modeling Approaches for Understanding COVID-19 Misinformation Spread in Sub-Saharan Africa

Ezinne Nwankwo[1*], **Chinasa Okolo**[2] and **Cynthia Habonimana**[3]
[1]Duke University
[2]Cornell University
[3]CSU-Long Beach
esn11@duke.edu, chinasa@cs.cornell.edu, cyn.habonimana@gmail.com

Abstract

Since the start of the pandemic, the proliferation of fake news and misinformation has been a constant battle for health officials and policy makers as they work to curb the spread of COVID-19. In areas within the Global South, it can be difficult for officials to keep track of the growth of such false information and even harder to address the real concerns their communities have. In this paper, we present some techniques the AI community can offer to help address this issue. While the topics presented within this paper are not a complete solution, we believe they could complement the work government officials, healthcare workers, and NGOs are currently doing on the ground in Sub-Saharan Africa.

1 Introduction

For the past several months, new research on how to stop the pandemic caused by COVID-19 has appeared almost daily. News outlets are reporting policies around social distancing and daily statistics of the spread of the virus. And, like in many previous global crises, much of the news cycle and the research findings focus on the world's largest economies, such as the United States, the United Kingdom, and China, where the first known cases were discovered. In this paper, we survey the current innovations around stopping the spread of the virus happening in the Global South, such as in Rwanda, Mexico, or India. We provide recommendations for developing machine learning and artificial intelligence techniques to quell the spread of misinformation through mediums such as WhatsApp and Facebook. We also present open areas of research for the "AI for Social Good" community.

2 Misinformation in the Global South

There are also significant differences in how fake news is spread throughout different regions of the world. In other regions where misinformation is spread through memes or satire, misinformation in Africa and the rest of the Global South often takes the form of extreme speech that incites violence or spread racist, misogynous, xenophobic messages [1]. Fake news is often spread with the intent to provide financial or political gain and is often exacerbated by digital illiteracy, lack of knowledge surrounding specific topics, and religious or personal beliefs that are biased towards fake news [2]. To get a sense of what are common themes in the "fake news" being spread across the African continent and in other developing nations, we thought it would be best to discover what ideologies they are rooted in. From our research, we saw three main themes in fake news being spread across Sub-Saharan Africa: distrust of philanthropic organizations, distrust of developed nations, and even distrust of leaders in their own respective countries. We posit that the distrust in these messages stems from the "colonial-era-driven distrust" of international development organizations [3] [4]. A more formal study and analysis is needed to make any concrete conclusions.

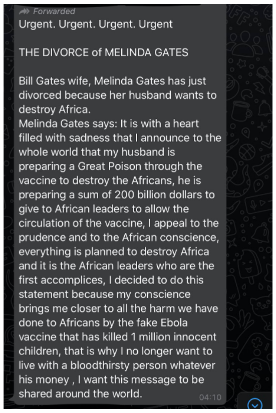
3 Misinformation and Social Media

According to Statista, social media generates 42% of the traffic-related to misinformation and fake news [5]. Twitter, Facebook, and WhatsApp are known to be harbingers of misinformation and the dynamic nature of these platforms not only allows for misinformation to be spread quickly but also very widely. Unfortunately, the distrust of philanthropic organizations such as the World Health Organization (WHO) and the Bill and Melinda Gates Foundation has led to people spreading misinformation and refusing to adhere to policies instituted by these entities. During the current COVID-19 crisis, we have seen widely shared sentiments saying that COVID-19 is a biological warfare weapon being spread by the World Health Organization or COVID-19 is a biological weapon created by China who is trying to purposely spread it in Africa (an example is detailed in Figure 1). The latter sentiments have led to physical harm through xenophobic attacks against Asians around the world and recent news of Africans being falsely accused of spreading coronavirus in China garnered significant media attention and diplomatic intervention.

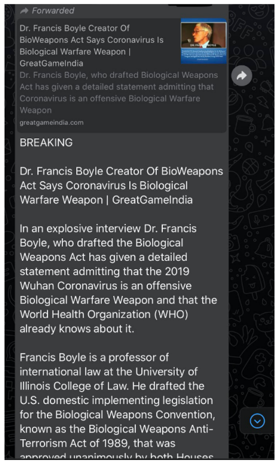
3.1 Addressing Misinformation

Information delivered via WhatsApp is more likely to be perceived as true due to it's higher penetration throughout the Global South and because it appears to come from sources that a recipient is more likely to trust [6]. To combat the rampant misinformation, WhatsApp has now limited the num-

*Contact Author



(a) Message detailing the alleged divorce of Bill and Melinda Gates.



(b) Message proclaiming COVID-19 as a biological weapon.

Figure 1: Examples of fake news spread through WhatsApp.

ber of times messages could be forward within groups to 5 [7]. Since messages on this platform are encrypted, it is unclear how much it has helped to curb misinformation, but these efforts are laudable. WhatsApp has also partnered with the WHO to create WHO Health Alert, a chat-bot service that provides credible information 24/7 about coronavirus [8]. Other platforms have implemented other strategies to fight against misinformation. Facebook and Twitter have begun labeling popular posts as "altered" or "disproven" to indicate that the information located within these posts are untrue. Despite these measures, misinformation spread continues to be rampant throughout Sub-Saharan Africa and the rest of the Global South. To address this, techniques such as LDA or BERT could be used to extrapolate topic modeling on messages for detecting fake news that could be viable, allowing governments and health organizations to identify topics they could focus on.

4 Approach

Until now, much of our research has been focused on analyzing a small sample of tweets and messages that have been confirmed as fake news. But we believe that this should be done on a larger scale. News articles, media transcripts, tweets, and search engine queries provide a great opportunity to use natural language processing techniques and sentiment analysis for analyzing text data. To combat misinformation in Africa, an important first step is to understand the common themes and topics proliferating throughout a community. With this knowledge, government officials and health organizations can focus on dissemination factual information about COVID-19. For example, if a topic model indicates that many of the sampled tweets are about a vaccine and have negative sentiment, then governments can tailor their announcements to addressing those concerns.

4.1 BERT and LDA

The two models we are considering for implementing a topic modeling analysis are BERT and LDA. Bidirectional Encoder Representations from Transformers (BERT) bi-directionally trains Transformer, a encoder-decoder used to learn contextual relationships between words, to generate language models [9]. These techniques have allowed BERT to achieve state-of-the-art on language tasks such as sentiment analysis, question answering, and named entity recognition. Latent Dirichlet Allocation (LDA) is a generative probabilistic model that forms explicit representations of text data to group into topics [10]. The backtrack approach that LDA incorporates is extremely useful in generating words related to topics present in documents or shorter forms of text. Both of these models have been widely used in similar applications and have extensive open-source software with pre-trained models.

Related studies incorporating these methodologies have been completed using search engine queries to understand health information needs in Africa [11] and using tweets and news publications to study topic coverage and sentiment dynamics on the Ebola virus [12]. A key difference in these two studies is the subset of text data collected to train the model on. The researchers collected search queries related to HIV or tweets specifically about Ebola to understand what was being discussed in specific communities. For this project, we are specifically looking to study topics in fake news articles shared via WhatsApp or through tweets. However, this subset of the full data of all tweets related to COVID-19 is much harder to acquire and identify.

4.2 Data Collection

As with most machine learning tasks, massive amounts of data are required to train the underlying models on. Collecting data on fake news and misinformation being spread in Africa can be a challenge. In theory, one would first need to collect all articles or tweets related to COVID-19 from Africa and then identify the ones that are fake or false. We suggest that a more practical way of collecting such a dataset is to crowd-source the false messages or tweets. There are already companies doing such work, such as Afara International and Africa Checks, one of which we have contacted for collaboration.

5 Conclusion

Misinformation around COVID-19 and other health-related topics can pose a huge danger to communities all around the globe. When such news is widespread, this could lead to sub-optimal behaviours such as ignoring social distancing requirements or taking medication that may do more harm than good [13]. If governments are able to understand what factors lead to misinformation, this can aid health officials in developing well-informed policies that citizens are more likely to adhere to. In this paper, we argue that this topic is an open area of research and suggest ways that we believe the AI research community can contribute. There is still much work to be done and other avenues of research not discussed in this paper, but we hope that this is a good start.

References

[1] Herman Wasserman and Dani Madrid-Morales. An exploratory study of media trust in kenya, nigeria, and south africa. *African Journalism Study*, 40(1):107–123, August 2019.

[2] David Cheruiyot and Raul Ferrer-Conill. "fact-checking africa" epistemologies, data and the expansion of journalistic discourse. *Digital Journalism*, 6(8):964–975, 2018.

[3] The legacy of colonial medicine in central africa.

[4] Stelios Michalopoulos and Elias Papaioannou. Historical legacies and african development. *Journal of Economic Literature*, 58(1):53–128, March 2020.

[5] Statista. Distribution of traffic sources for fake news in the united states in 2017, 2017.

[6] Raquel Rennó. Whatsapp: The widespread use of whatsapp in political campaigning in the global south, 2018.

[7] WhatsApp Inc. About forwarding limits, 2020.

[8] Adam S Miner, Liliana Laranjo, and A Baki Kocaballi. Chatbots in the fight against the covid-19 pandemic. *npj Digital Medicine*, 3(1):1–4, 2020.

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[10] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[11] Rediet Abebe, Shawndra Hill, Jennifer Wortman Vaughan, Peter M Small, and H Andrew Schwartz. Using search queries to understand health information needs in africa. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 13, pages 3–14, 2019.

[12] Erin Hea-Jin Kim, Yoo Kyung Jeong, Yuyoung Kim, Keun Young Kang, and Min Song. Topic-based content and sentiment analysis of ebola virus on twitter and in the news. *Journal of Information Science*, 42(6):763–781, 2016.

[13] Leonardo Bursztyn, Aakaash Rao, Christopher Roth, and David Yanagizawa-Drott. Misinformation during a pandemic. *University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2020-44.*, April 2020. Available at SSRN: https://ssrn.com/abstract=3580487.