



TOPIC MODELLING OF COVID-19 MEDIA COVERAGE IN UGANDA

By Joyce Yanru Jiang & Anne Whitehead
@ Whitehead Communications, 2020

MODELLING OF COVID-19 MEDIA COVERAGE IN UGANDA

INTRODUCTION

Uganda – like the rest of the world – has experienced a large volume of news coverage about the Covid-19 pandemic. Whitehead Communications used Artificial Intelligence (AI) to analyse the dominant topics that emerged within Covid-19 coverage by online news websites in Uganda. We gathered more than 13,000 news articles that included the terms “covid” or “corona” from 13 English language news websites. This analysis included the websites of Uganda’s major print publications Daily Monitor, New Vision, and Red Pepper as well as major online-only Ugandan media houses Chimp Reports, Nile Post, PML Daily, Observer and The Independent¹, and other online-only news publishers including Softpower, The Tower Post, Eagle News, Trumpet News and The Brink News. This was exploratory and experimental research to find out which topics stood out in Ugandan media coverage of Covid-19 according to a machine learning method called LDA Topic Modelling, which employs a Natural Language Processing (NLP) technique.

This methodology alone cannot be considered comprehensive or conclusive, but it offers an initial indication and useful insights into how the subject was covered in the Ugandan press.

Figure 1 below shows of the volume of articles that we collected from each news website. This only represents the online written news media landscape in Uganda, but presents us with an initial sample for further analysis

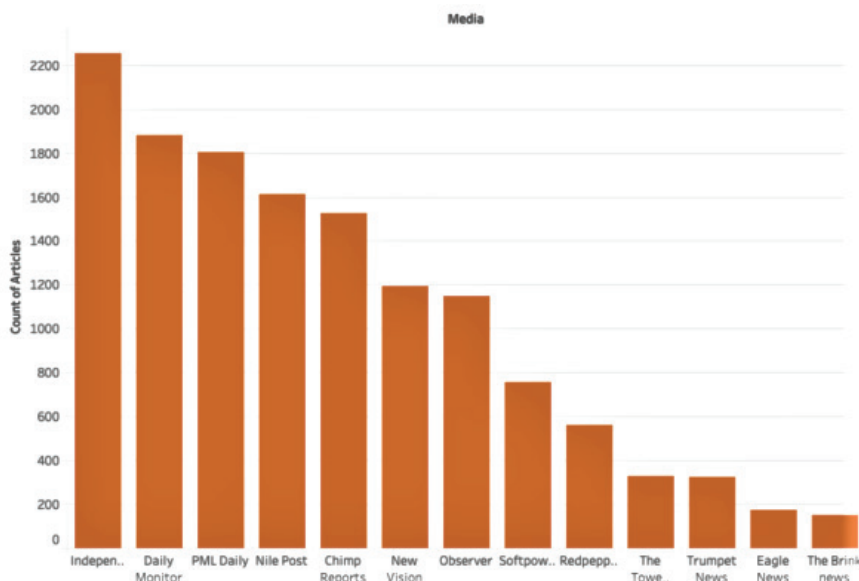


Figure 1 - Volume of articles by media house

¹ The Observer and Independent were both printed but became online-only during the Covid-19 crisis.

OUR PURPOSE

This research aims to explore how Covid-19 has been covered by Ugandan online news websites within the first six months of 2020 using machine learning. We identified top news sites in Uganda that publish in English and applied an LDA Topic Modelling computational technique to discover which topics are being covered related to the coronavirus pandemic. Our research is intended to deliver insights to those who have a special interest in Uganda's media industry, or the country's experience with Covid-19, or those interested in the application of machine learning to media and communications research. This is part of a wider research project by Whitehead Communications exploring the application of multiple new research methods to see the bigger picture, draw correlations and build stronger research-based foundations on which to develop communications strategy.

OVERVIEW OF COVID-19 MEDIA COVERAGE IN UGANDA

The weekly number of articles produced by Ugandan online media increased dramatically in mid-March of 2020, as the first Ugandan case was declared on the 21st of March² and the country went into lockdown in the same week³. Covid-19 related coverage by volume of online articles reached its height in April of 2020, then began to decline. This indicates that media interest in the disease and its impact peaked during the period when lockdown was strictest and cases were just beginning to mount, but before the first Covid-19 death was announced in Uganda on the 23rd of July, 2020⁴. We removed from our dataset any stories republished from foreign media outlets in order to focus only on news produced in Uganda.

According to our dataset, the Daily Monitor published the largest volume of Covid-19 related articles online in April, averaging 150 per week, followed by The Independent (~141/week), then PML Daily (~137/week), Chimp Reports (~114/week), Nile Post (~104/week) and New Vision (~68/week). The volume of articles began to drop again in May. You can see this trend in Figure 2 below.

² Ministry of Health – Uganda @MinofHealth (21 March 2020). Available at: <https://twitter.com/MinofHealthUG/status/1241469504087035904?s=19>

³ Address on the Corona virus (Covid-19) Guidelines on avoiding the Pandemic. Yoweri K. Museveni (18 March 2020). Available at: <https://www.yowerikmuseveni.com/address-corona-virus-covid-19-guidelines-avoiding-pandemic>

⁴ Uganda registers first Covid-19 death. New Vision (23 July 2020). Available at: <https://www.newvision.co.ug/news/1523265/uganda-registers-covid-19-death>

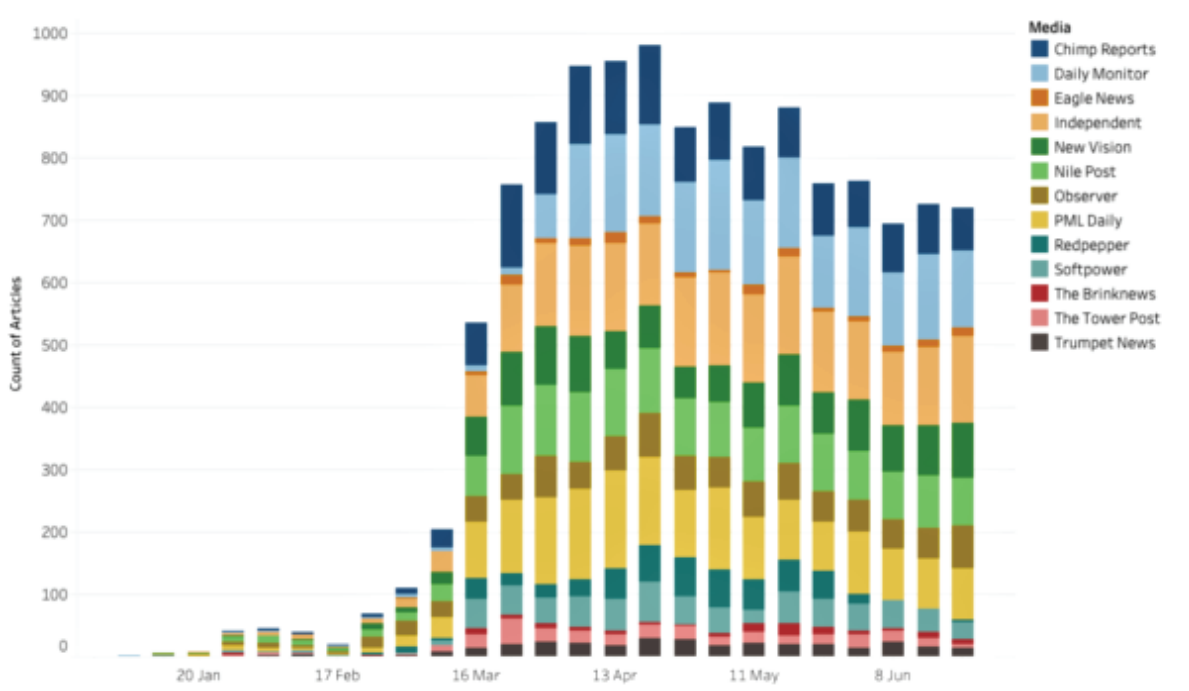
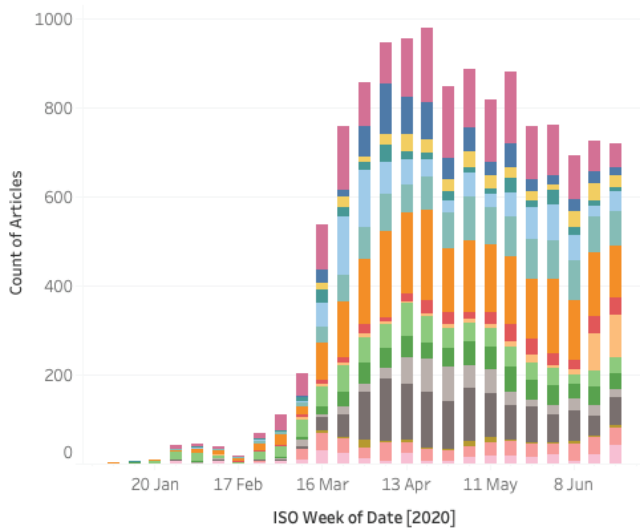


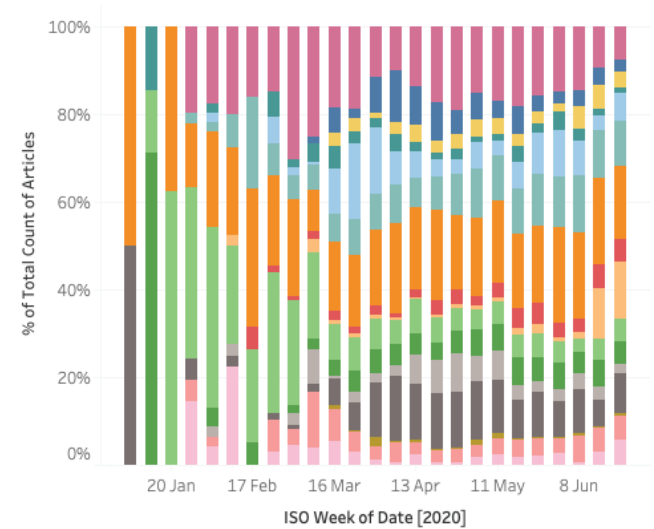
Figure 2 - Weekly volume of articles by media house

TOPIC ANALYSIS

Articles by topic per week



Percentage of topics by weeks



Topics

- Cases & testing
- Court & Justice
- Domestic outbr...
- Editorials & per...
- Electoral politics
- Healthcare
- Police action
- Sports
- Contributions t.
- Culture & religi.
- Economy & fina.
- Education
- Global outbreak.
- Parliament bud.
- Presidential dir.
- Travel restricti..

Figure 3 - Articles by topic

Our analysis identified 16 topics that emerged in Ugandan online news media coverage of Covid-19, as listed below:

- 1.Cases & testing
- 2.Healthcare
- 3.Domestic outbreak (and government enforcement)
- 4.Parliament budgeting
- 5.Travel restrictions
- 6.Global outbreak and international response
- 7.Contributions to Covid-19 budget
- 8.Economy & finance
- 9.Education
- 10.Courts & justice
- 11.Sports
- 12.Culture & religion
- 13.Electoral politics
- 14.Police action
- 15.Editorials & personal stories
- 16.Presidential directives

These topics were identified through an unsupervised algorithm, which we ran several times using different parameters until we found the most optimal result. We checked the articles it grouped together to identify what the topics were. Some topics were straightforward, as the articles shared common themes, such as # 9 Education and # 11 Sports. Others were made up of mixed subtopics under one major topic, such as # 3, which gathered together stories about the domestic outbreak and how the government was responding. We also chose to combine four automatically generated topics into two, since their topics were very similar. The algorithm clustered together groups of stories that differed, but shared a common thread, such as those mentioning the courts system (#10), or those related to different types of restrictions on transport and travel (#5). More details of our methodology and a further examination of a few key topics and how they manifested in news coverage are shared below.

PRESENTATION OF KEY TOPICS

Ugandan media coverage of Covid-19 was similar in some ways to how other countries covered the pandemic,⁵ but our Topic Modelling exercise also highlighted certain narratives that were unique to Uganda's experience, such as electoral politics related to plans for a "scientific election". A few examples are highlighted below to further explore the results of our analysis. The following data visualizations each show a specific topic, with the height of the bars representing the number of articles per week, and the colours breaking down which media houses published those articles, as per the accompanying legend. ⁶

⁵ Communicating COVID-19: weekly updates of main topics in international news coverage of COVID-19. Boston University's AIEM Research Group. Available at: <https://covid19.philemerge.com/>

⁶ Note that data from the Observer was not fully included in the following Topic Modeling analysis due to a technical error limiting initial search results, though we have represented their total articles numbers in Figure 1.

Domestic Outbreak and government response

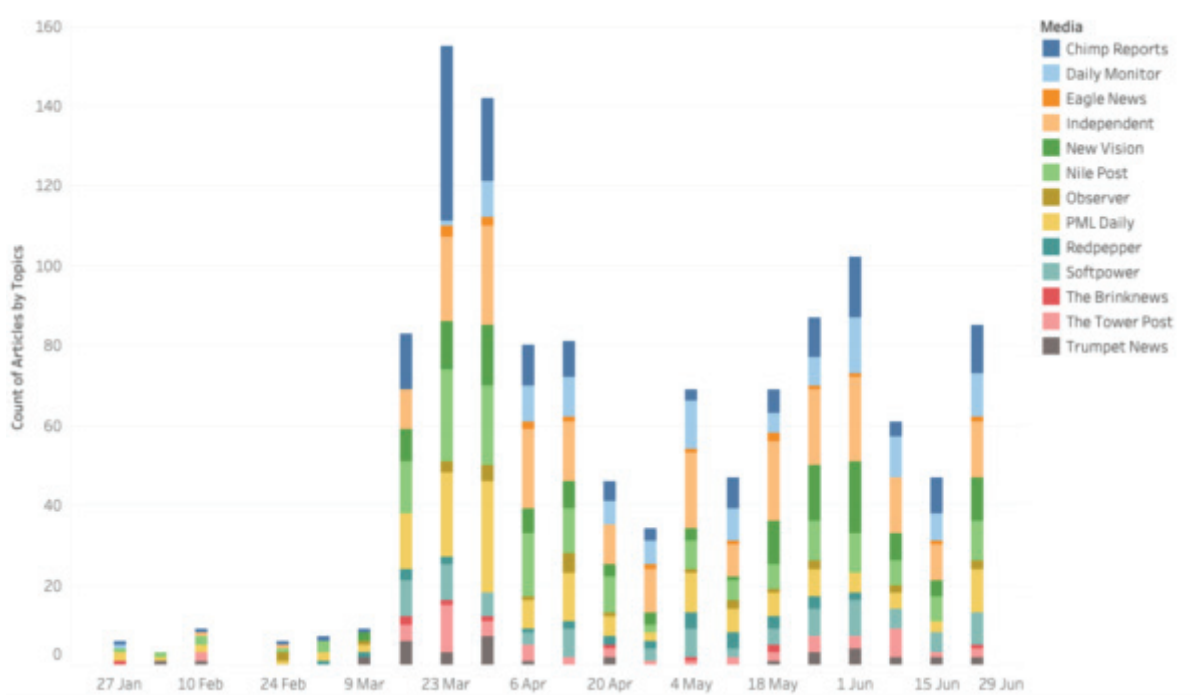


Figure 4 - Domestic outbreak

Uganda registered its first case of Covid-19 on the 21st of March. Similarly, media coverage of the domestic outbreak spiked that same week to its highest level. The spikes in this topic at the beginning June can be attributed to the inclusion of articles about repatriating citizens from abroad and reviews of other lockdown measures.

Topic	Narratives	Sample Headlines	
Domestic outbreak and government response	Public health directives issued & Push for social distancing & Repatriation & Economic response	3/18/2020	CoronaVirus: Museveni Bans Wedding, Political Rallies, Washing Corpses, Bars, Schools, Churches
		3/30/2020	Museveni orders lockdown in latest coronavirus measures
		4/5/2020	COVID-19: Health Ministry asks Security to “Intervene” as Jogging crowds Pile on Roads
		4/9/2020	VIDEO: Museveni challenges Ugandans on indoor exercising
		5/6/2020	Govt Proposes Stringent Guidelines for Boda Bodas, Taxis After Lockdown
		6/3/2020	We are Ready to Receive Ugandans Stranded Abroad- Foreign Affairs Minister Sam Kuteesa
		6/25/2020	World leaders urge governments on COVID-19, stimulating economies
		6/12/2020	Bars, restaurants petition Kadaga to push for their re-opening

Global outbreak and international response

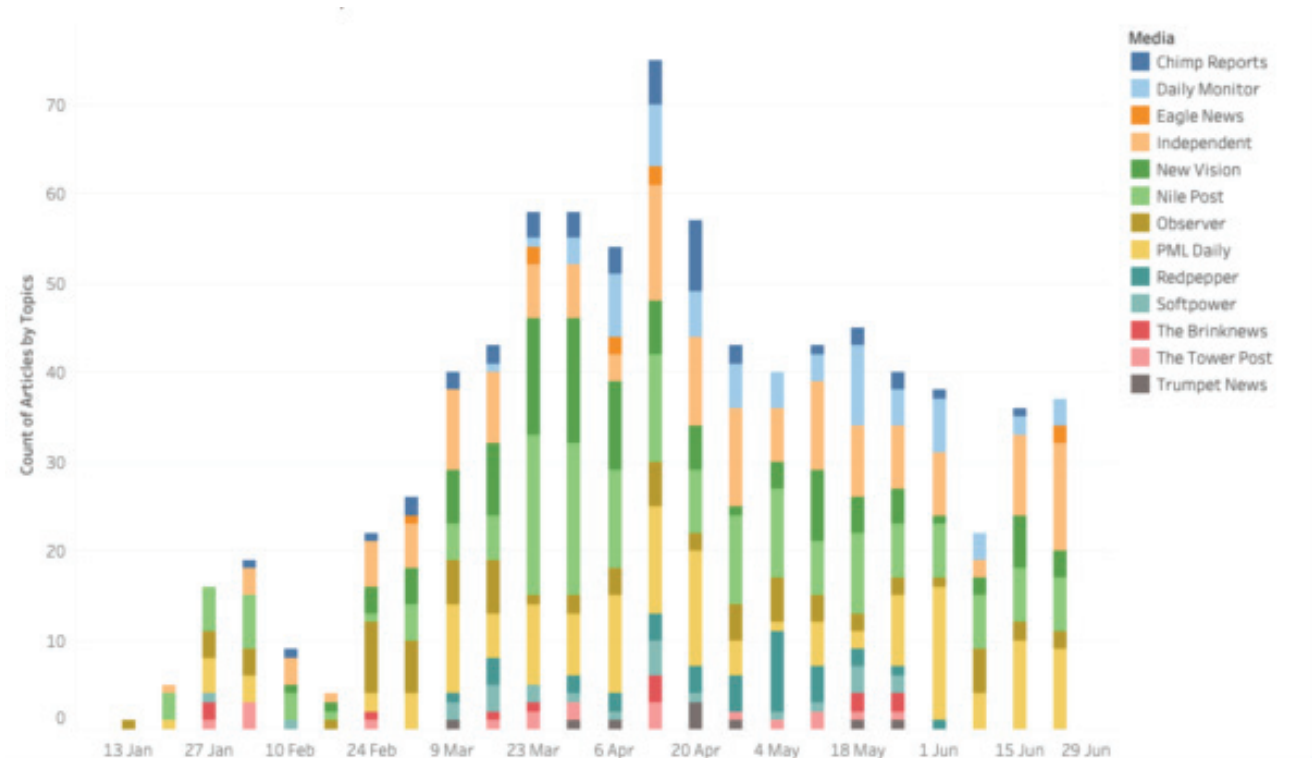


Figure 5 - Global outbreak and international response

Ugandan media covered a range of topics related to the global pandemic and international response. This included updates on the outbreak in various countries, as well as statements by world leaders and international aid contributed to Uganda.

Topic	Narratives	Sample Headlines
Global outbreak and International response	Global statements & Outbreaks in other countries & International aid to Uganda	1/31/2020 Coronavirus declared a global health emergency by World Health Organization
		2/6/2020 Coronavirus: Museveni says Ugandans in China shouldn't be evacuated
		3/25/2020 Spain's coronavirus death toll jumps 514 in 24 hours
		4/15/2020 Boris Johnson 'discharged from hospital'
		4/17/2020 US's suspension of funding WHO is a threat to world's COVID-19 response
		4/17/2020 COVID-19 CRISIS: US govt gives Uganda Shs14b to boost Covid-19 fight
		5/19/2020 Museveni, Canadian Prime Minister Trudeau discuss repatriation plan for Canadians stranded in Uganda
		5/28/2020 World leaders urge governments on COVID-19, stimulating economies
		6/12/2020 After Burundi president's death, African leaders response to Covid-19 comes under scrutiny
		6/29/2020 China refutes Pompeo's anti-China statements over China-Africa ties

Cases & testing

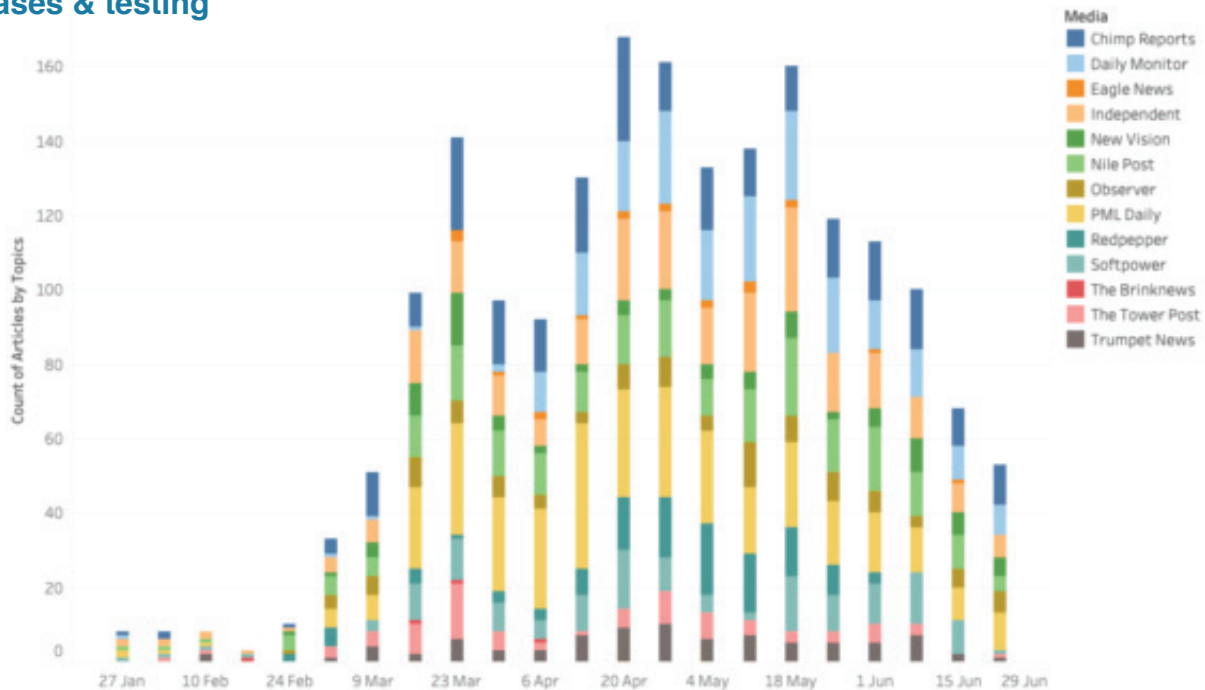


Figure 6 - Cases & testing

Reports on testing and cases were covered fairly highly from late March (when the first Ugandan case was reported) through to early June, with a few spikes along the way. Reporting on this subject declined through June, despite case numbers rising in the same period. A major sub-topic that appeared under the “cases & testing” topic was reference to truck drivers. This is because Uganda implemented a focused regimen at border points testing truck drivers from border countries such as Tanzania and Kenya, and therefore more positive cases were initially declared among truck drivers. There was also a rise in testing and cases related stories in mid-May when numbers were rising in Uganda and the East African Region.

Topic	Narratives	Sample Headlines
Cases & testing	Updates on cases & Test on truck drivers	3/21/2020 BREAKING: Uganda has confirmed first coronavirus case
		4/21/2020 How truck drivers are putting Uganda's Covid-19 fight to the test
		4/24/2020 Kenyan Truck Driver Who Tested Positive for Coronavirus Intercepted in Kisoro
		5/16/2020 No new Community case as 43 Truck Drivers test Positive for COVID-19
		5/20/2020 Uganda Records Highest One-Day Rise in COVID-19 Cases as 21 Truck Drivers Test Positive
		5/22/2020 South Sudan Vice President Machar, wife, information minister test positive for COVID-19
		6/5/2020 Developing Story: OPM, Finance Ministry Drivers test Positive for COVID-19
		6/14/2020 Uganda: Covid-19 Cases Rise to 696 As 11 New Infections Are Registered

Electoral Politics

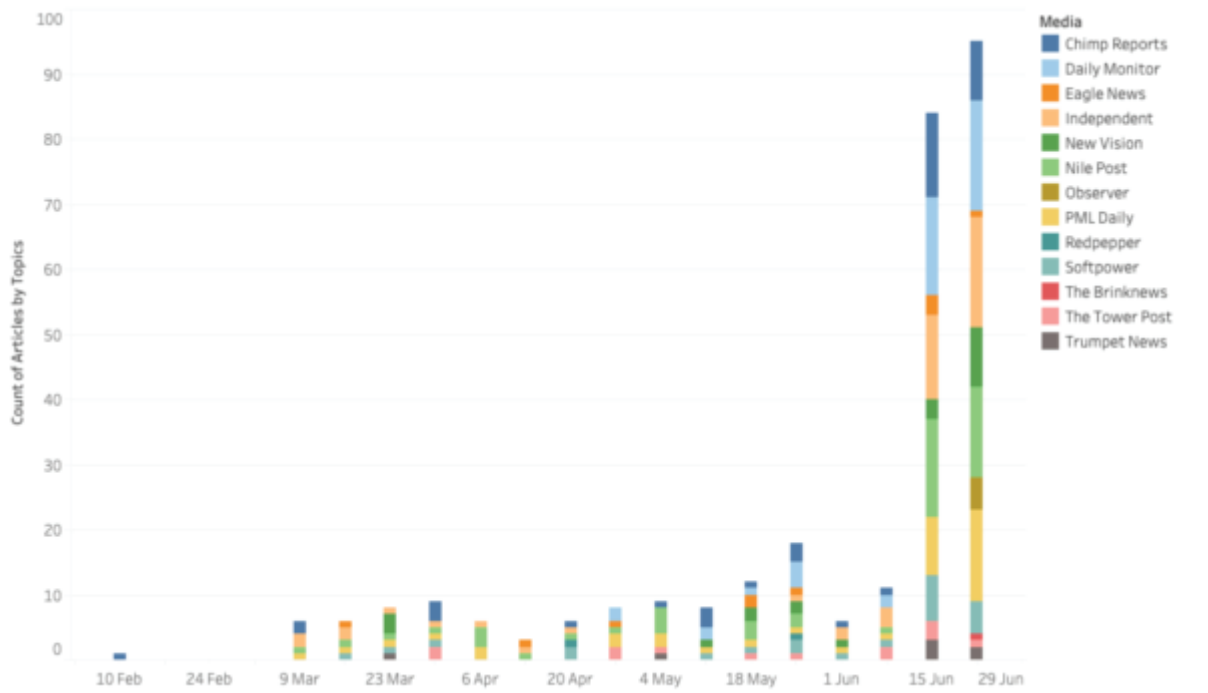


Figure 7 - Electoral politics

Uganda’s next election is scheduled for February 18th, 2021. “Electoral politics” is one topic that was picked up by our LDA Model and coverage on this subject increased dramatically in mid-June. This was likely due to an announcement at that time laying out the roadmap for electoral campaigns, which has come to be known as a “scientific election”. The concept is that, due to public health measures meant to manage the Covid-19 pandemic, candidates will be restricted from holding public rallies or other meetings that may contradict social distancing directives and will instead depend on media to persuade the voters through the campaign period. This topic was heavily covered and extensively debated in the second half of June.

Topic	Narratives	Sample Headlines	
Electoral politics	Scientific election announced, debated & Campaign plans	4/30/2020	#COVID-19: NRM to postpone its primaries
		5/29/2020	Is COVID-19 a stumbling block along 2021 electoral process?
		6/16/2020	2021 Elections: EC Bans Mass Rallies, Candidates to Campaign Using TV, Radio and Digital Media
		6/20/2020	UCC Speaks Out on Scientific Campaigns, Says Media Coverage Might Not Match Voter Distribution
		6/21/2020	The 2021 scientific election: Rigged process or blessing?
		6/29/2020	The Besigye-Bobi Wine alliance

Courts & justice

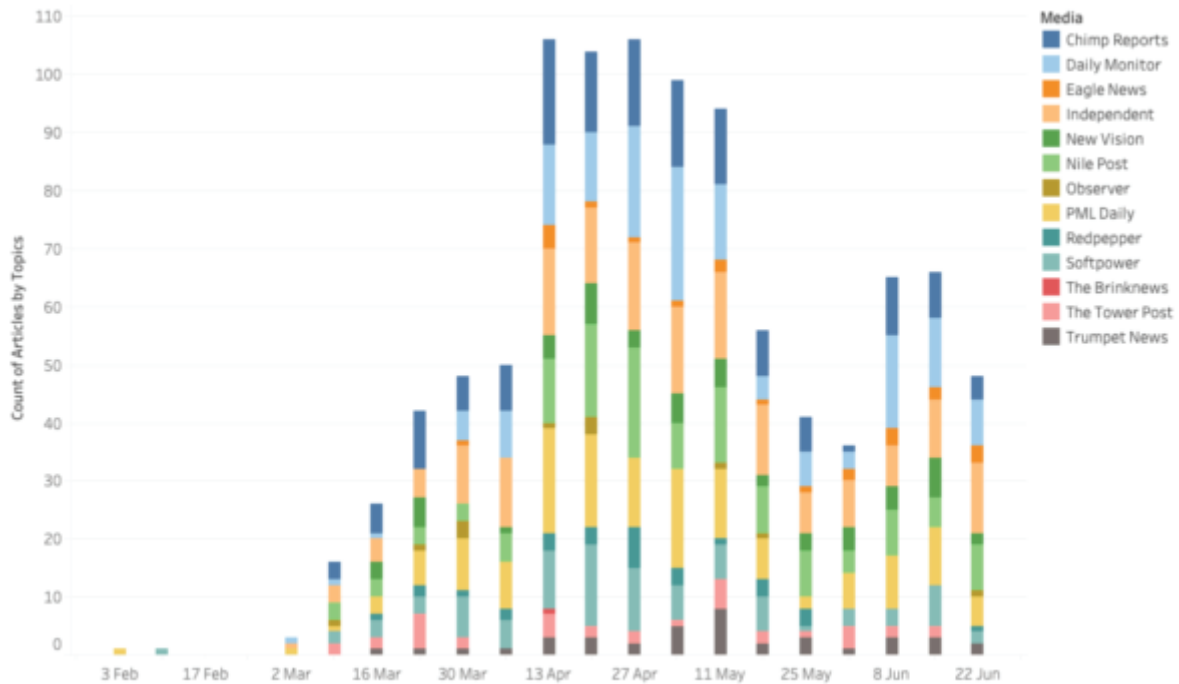


Figure 8 - Courts & justice

Stories involving the justice system were grouped together, which included coverage of how certain court cases were suspended and lawyers advocating for returning to court in April, then a story about socialite Bad Black suing over compensation for a Covid-19 PSA in June.

Topic	Narratives	Sample Headlines	
Courts & justice	Disruptions to the justice system & Legal actions	3/19/2020	Court Sitings Suspended Over Coronavirus
		4/7/2020	Covid-19: Prisons identify 2,000 inmates for release
		4/14/2020	Tumukunde Bail Postponed Over Covid19 Lock down
		5/5/2020	Court business goes digital as lockdown bites
		5/9/2020	Court permits lawyers to move during lockdown
		5/16/2020	Digital migration: Is Covid-19 a turning point for the Judiciary?
		6/14/2020	Bad Black threatens to sue Health Ministry for shs550m over Covid-19 deal

Culture & religion

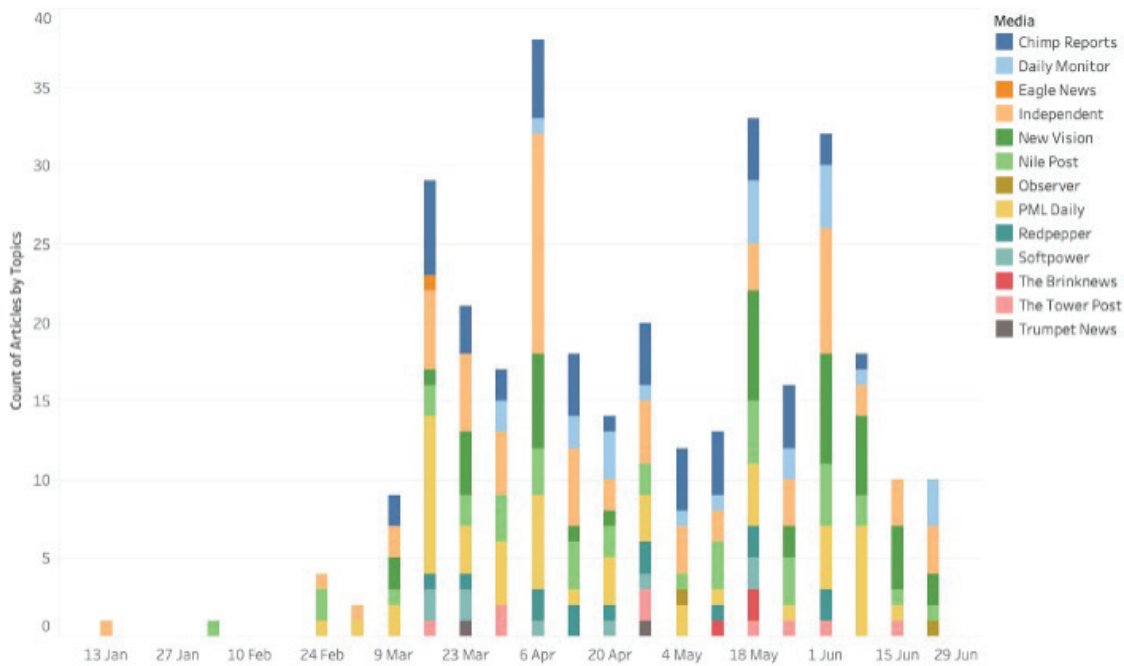


Figure 9 - Culture and religion

This topic grouped together articles about how Ugandans observed various cultural and religious activities during the Covid-19 lockdown, including: Easter, the Kabaka’s birthday, Ramadan and Martyr’s Day.

Topic	Narratives	Sample Headlines
Culture & religion	Guidelines & Easter & Kabaka’s birthday & Ramadan & Martyr’s Day	3/14/2020 Coronavirus: Archbishop Kazimba Issues Strict Guidelines for Church of Uganda
		4/1/2020 COVID-19 pandemic: How different Christian denominations will celebrate Easter
		4/9/2020 All Saints Cathedral: Prepare for Unusual Easter Celebration
		4/13/2020 UCC Speaks Out on Scientific Campaigns, Says Media Coverage Might Not Match Voter Distribution
		5/23/2020 Kabaka Celebrates 65th Birthday, Awards Kawanga Semwogerere
		6/2/2020 “Pray from home, God is everywhere”- Museveni tells Muslims in Eid message
		6/14/2020 Only 60 people invited for martyrs day Celebrations

Contributions to Covid-19 budget

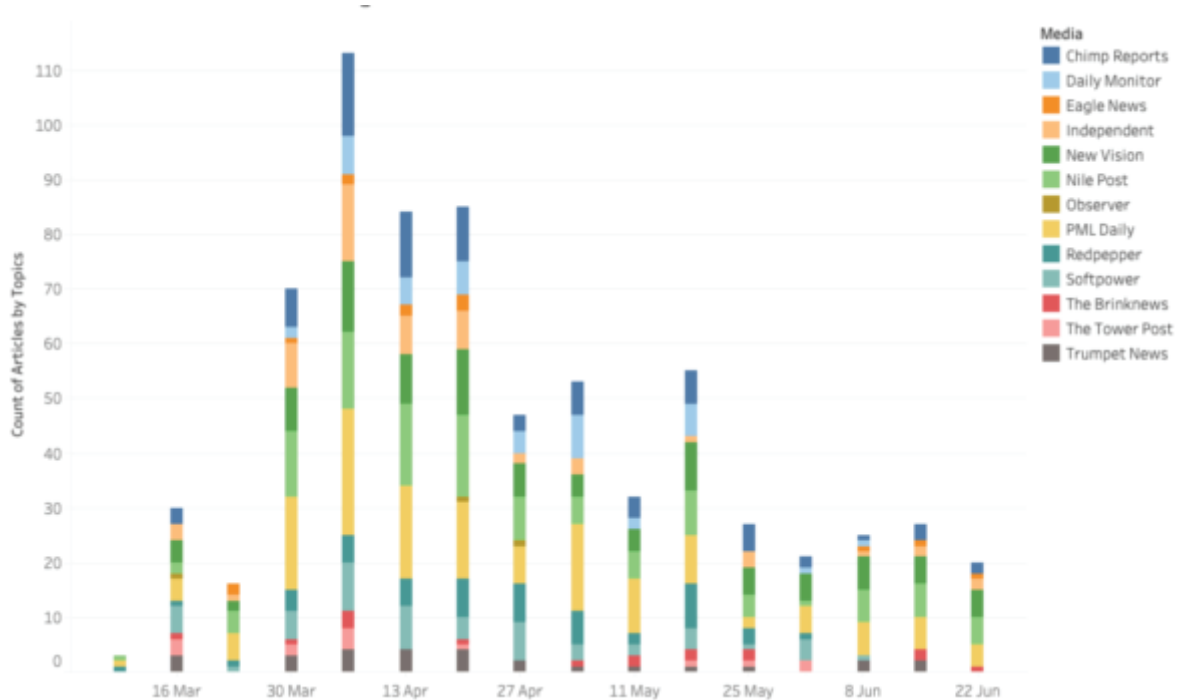


Figure 10 – Contributions to Covid-19 budget

A lot of news was generated in Uganda around which companies and other organizations were contributing to the Covid-19 National Task Force’s budget to handle the crisis. The vast majority of articles were positive PR pieces for the donors. There was also some coverage of directives restricting donations and a few pieces criticizing the fundraising exercise, including: a call for citizens to donate 10,000 from their salaries, Speaker Rebecca Kadaga’s plan to develop a sanitizer product, questions over accountability and plans to distribute face masks.

Topic	Narratives	Sample Headlines	
Contributions to Covid-19 budget	Donations to the Covid-19 National Task Force & Donation restrictions & Criticism of Kadaga & Facemask distribution & calls for transparency	3/17/2020	I won't be insulted! Kadaga hits back at Medical Association over quack US doctor claims
		4/3/2020	COVID-19 CRISIS: People Power beats Museveni's directive, distributes food to needy families
		4/10/2020	Total, Stanbic bank, UBL donate shs120M fuel towards Covid-19 response
		4/24/2020	COVID-19: MTN Heeds to Museveni's Call with 3 Pickups after he Trashed Data Donation
		5/17/2020	Uproar as government asks salaried workers to donate Shs10, 000 each
		5/19/2020	Call for transparency in war against COVID-19
		5/20/2020	Coca-Cola Beverages Africa Hands Over UGX 1.3 Billion To the COVID-19 National Task Force
		5/20/2020	Minister Anite cautions on political color masks



Coverage by media houses

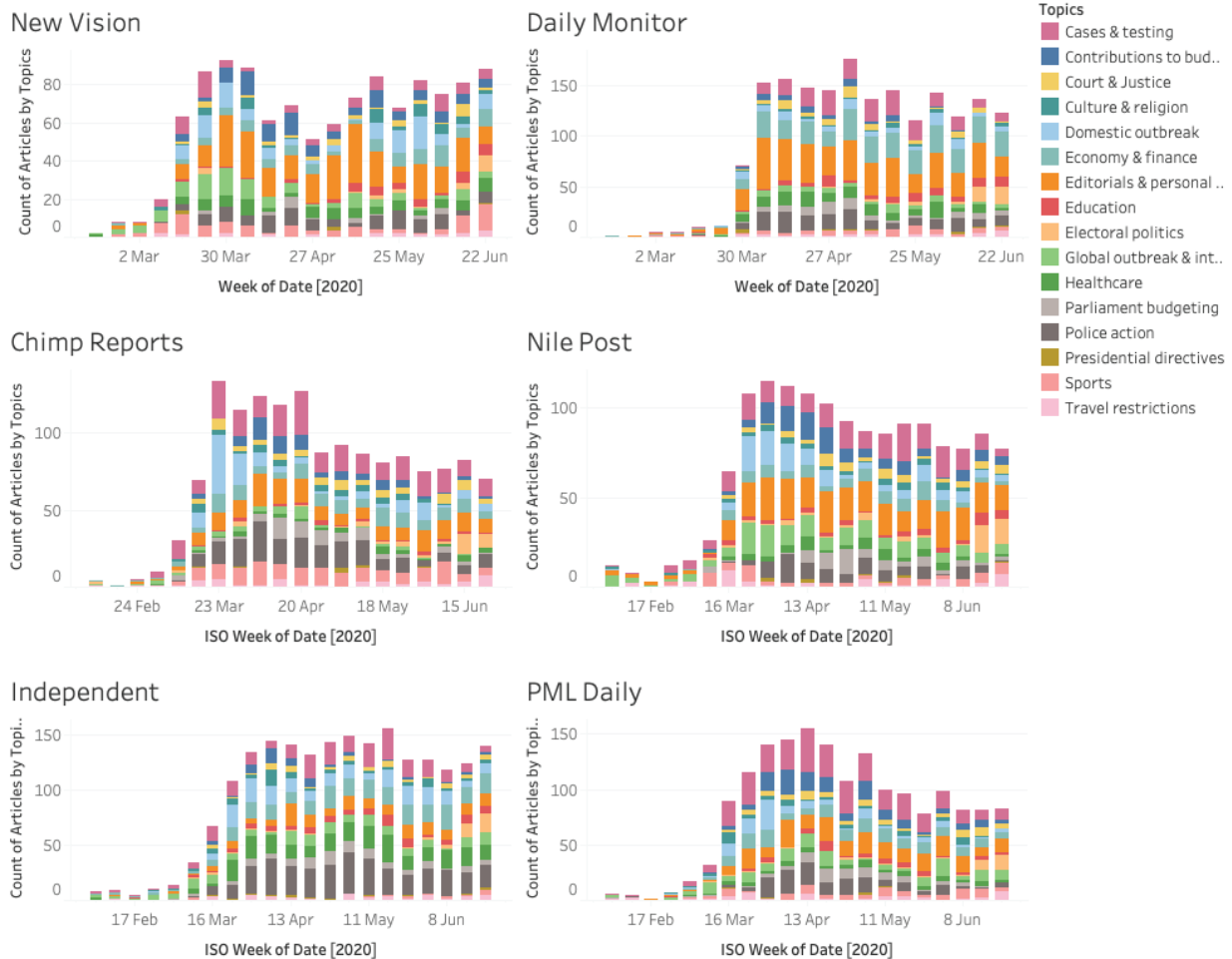


Figure 11 – Comparison of coverage by media houses

The graphs above show how each media house covered Covid-19 over the first six months of 2020, indicating the volume of articles (height of bars) and the topics included (colours). Uganda's two major daily publications New Vision and Daily Monitor both published more editorial content, while online-only publications like Chimp Reports and PML Daily covered more updates on cases and testing. Most media houses saw a peak in coverage during April, with the exception of New Vision (it was restructuring its staff at that time⁷, which may have had an effect on production). The Independent had the most articles overall and included more articles about police action related to Covid-19.

⁷ Media Crisis! Giants Vision Group, NMG announce huge salary cuts (30 April 2020). PML Daily. Available at: <https://www.pmldaily.com/news/2020/04/media-crisis-giants-vision-group-nmg-announce-huge-salary-cuts.html>

METHODOLOGY

We used the **Latent Dirichlet allocation (LDA) Topic Modelling** technique to categorize news based on topics and assign topics to each article. The following are the main steps in our research.

***Data collection:** First, using local media knowledge, we identified the Ugandan news websites likely to produce the highest volume of articles. Then we developed a tailored scraping script in Python for each news website. On each website, we searched the capital-insensitive stem keywords “covid” and “corona” to gather articles related to Covid-19 news (this also included extended varieties of these keywords such as “coronavirus” and “Covid-19”). The data we scraped ranged in dates from January to mid July. Since we did not get comprehensive data from July, this month was excluded from our analysis.

***Data cleaning:** After gathering 10,427 articles including the keyword “covid”, we combined this dataset with 9,927 articles including the keyword “corona”, then we conducted data cleaning and got 12,750 in our final dataset for LDA modelling. The steps we took for data cleaning included:

- Removed duplicated articles
- Only kept articles whose headlines or article body text contained Covid-related keywords (ex. COVID, corona, lockdown, outbreak, curfew, reopen, facemask, respiratory, WHO, hospital, etc.). The keywords were determined in a discretionary way that may lead to some Type I and Type II errors. However, we believe our result is mathematically reliable.
- Removed articles that contain keyword “Toyota” (to exclude articles that were picked up in our data about the Toyota Corona car model) and “coronation”
- Remove articles that contain AFP (Agence France-Presse) and Xinhua (Chinese news agency) and do not have any Uganda-related keywords. Both AFP and Xinhua are international news agencies that sell news to Ugandan local media. Though the Ugandan audience is still exposed to this content, we chose to remove international articles unrelated to Uganda in order to focus on news produced locally.

***Challenges with specific media houses:** the Observer was not included in our topic model because its initial search results did not return more than 100 articles at a time. We were later in touch with Observer and they updated their website functionality to allow for unlimited search results, so that allowed us to measure the overall volume of articles they

published on this topic, but unfortunately we were not able to run the algorithm again so their content was not fully included in our topic model. This is because, once we added additional articles, the order of data changed, which affected the results of modelling. We also faced a challenge with Nile Post, in that their dates were formatted as “X time ago” instead of the exact date. We reached out to Nile Post and they were able to change the format of dates on their website for us in time to include these results in our LDA Topic Model. Finally, we noticed a gap in our data from the Red Pepper, but were able to confirm with this media house that their website was inactive at the time.

***LDA Topic Modelling:** We ran our algorithm in many variations, changing specifications of the number of topics between 15 to 22, and found that we received the most optimal results when we set the number of topic clusters to 18. Two pairs of topics seemed redundant, so we grouped them together, which left us with 16 unique topics. Our algorithm generated two clusters that were a mix of personal stories, editorials and tips for coping with the effects of Covid-19, which we amalgamated into topic # 15. Also, stories to do with declaring the number of cases and efforts to conduct testing were amalgamated into topic # 1.

***Topic analysis:** Our first trial for topic analysis was to check keywords for each cluster. We applied text analysis of article and headline contents using Python to develop word clouds for each cluster. Since some topics comprised of mixed subtopics, we were unable to tell the actual topic from analysing keywords themselves. The Figure 2 below is a demonstration of this challenge.



Figure 12 - Word cloud: text analysis of topic cluster

We then reviewed around 15 of the top matching sample articles in each cluster to further analyse and determine topic categories. We went over the article headlines together with our team manually again to confirm topic subjects and name each one.

FUTURE RESEARCH

Whitehead Communications hopes to extend our use of the LDA Topic Modelling technique by experimenting with different parameters on different projects to optimize our results. This novel machine learning method allows us to conduct big data analysis of news coverage, which can then be assessed and correlated with other research methods to offer insights into the nature and impact of media.

African news is commonly excluded or marginalized in academic research due to the lack of organized archives providing a convenient dataset. Ugandan news websites are less likely to offer APIs for developers, and we had to tailor scraping methods for each website, which was more labour intensive. However, we had the advantage of close relationships with media owners who helped us navigate their websites. We wish to challenge the status quo by applying cutting edge methods to Ugandan research and ensure that Uganda is included in the emergence of this type of research globally. We encourage more local media houses to cooperate with international news archives, or develop APIs for easier data collection, in order to make historical articles more accessible for researchers and scholars. We also encourage researchers in African countries and internationally to continue including African content in global analysis, particularly when training algorithms or otherwise contributing to a more inclusive global body of knowledge.

ACKNOWLEDGEMENTS

We must express our gratitude to everyone who offered their assistance to this project, from our own research team in Uganda, including Owilla Mercy Abiro and Norman Angel, to the editors and owners of various Ugandan media houses who offered feedback on our preliminary data and updated their websites to facilitate our research. We are also grateful to Ugandan PR practitioners and media researchers who offered their recommendations and moral support. Finally, we would like to recognize the AIEM research group at Boston University who took interest in our project when we realized that we were doing similar research and whom we hope to continue collaborating with in the coming months. We share this exploratory research in good faith and remain open to learning and collaborating further with anyone interested.